



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Making Mobile Authentication Practical: User-Centric Design for Deployable Security

Florentin Putz

DISSERTATION

submitted in fulfillment of the requirements
for the degree of Doctor of Engineering (Dr.-Ing.)

First examiner: Prof. Dr.-Ing. Matthias Hollick

Co-examiner: Univ.-Prof. Dr. René Mayrhofer

Department of Computer Science
Technische Universität Darmstadt

Darmstadt, 2026
D17



Florentin Putz, *Making Mobile Authentication Practical: User-Centric Design for Deployable Security*, Doctoral Thesis, Technical University of Darmstadt, 2026.

Secure Mobile Networking Lab
Department of Computer Science
Technical University of Darmstadt
Year of publication on TUpriints: 2026
Date of defense: July 9, 2026

Please cite this work as:

DOI: <https://doi.org/10.26083/tuda-8185>

URN: <https://nbn-resolving.org/urn:nbn:de:tuda-tuda-157637>

In Copyright

<https://rightsstatements.org/page/InC/1.0/>

Abstract

Authentication is one of the most user-facing components of digital security, yet many authentication measures remain difficult to use or deploy. Complex passwords and manual key verification often cause user frustration, circumvention, and systemic vulnerabilities. In this dissertation, we investigate how to make mobile authentication protocols more practical by elevating usability and deployability to first-class design goals alongside traditional security. We take a user-centric perspective on three connected scenarios that progressively expose different practical barriers to authentication: everyday web login, mobile service use under infrastructure disruption, and decentralized authentication for end-to-end encrypted messengers.

First, we examine FIDO2, the passwordless web authentication standard underlying passkeys. In a lab study ($N = 87$), we compare biometric platform authentication on smartphones with authentication through an external NFC security key. The results show that both options provide high usability, though participants preferred biometric platform authentication. We identify critical adoption barriers, such as multi-device usage constraints, and audit FIDO2's extensibility, revealing that current web browser limitations severely hinder custom deployments.

Second, we investigate a deployment context in which conventional authentication assumptions break down: mobile service use when centralized infrastructure fails, such as during natural disasters that create locally connected but globally disconnected network islands. Our survey ($N = 857$) shows that respondents prioritized familiar general-purpose applications over dedicated crisis tools, with messenger applications ranked highest overall. These findings identify messenger-based communication as a major user priority during infrastructure disruptions, motivating the need for messenger authentication that does not depend on globally reachable, centralized backend services.

Finally, building on this result, we introduce PairSonic, a decentralized authentication ceremony for end-to-end encrypted messengers. PairSonic enables nearby users to authenticate each other simultaneously without requiring trusted third parties, public-key infrastructure, or Internet access. In a lab study ($N = 45$), participants significantly preferred PairSonic over the state of the art because it reduced manual verification effort and scaled better to groups of users. Because PairSonic relies on an acoustic out-of-band channel, we systematically evaluate acoustic data transmission across diverse commodity smartphones. We identify multipath propagation and hardware variations as major reliability challenges and propose robust design patterns for practical deployment.

Overall, this dissertation demonstrates that strong authentication becomes practical when cryptographic design is integrated with human interaction and deployment constraints.

Zusammenfassung

Authentifizierung gehört zu den sichtbarsten Komponenten digitaler Sicherheit. Viele Verfahren sind jedoch wenig benutzungsfreundlich, schwer praktisch einsetzbar, oder von Infrastruktur abhängig, die in kritischen Situationen ausfallen kann. Komplexe Passwörter und manuelle Schlüsselverifikationen führen häufig zu Frustration, Umgehungsverhalten und systemischen Schwachstellen. Diese Dissertation untersucht daher, wie mobile Authentifizierungsprotokolle praxistauglicher gestaltet werden können, indem Benutzungsfreundlichkeit und praktische Nutzbarkeit neben Sicherheit als gleichrangige Entwurfsziele behandelt werden. Im Mittelpunkt stehen drei miteinander verbundene Szenarien: alltägliche Web-Logins, mobile Dienste bei Infrastrukturausfällen und dezentrale Authentifizierung für Ende-zu-Ende-verschlüsselte Messenger.

Zunächst untersuchen wir FIDO2, den passwortlosen Web-Authentifizierungsstandard für Passkeys. In einer Laborstudie (N = 87) vergleichen wir biometrische Plattformauthentifizierung auf Smartphones mit der Authentifizierung über externe NFC-Tokens. Beide Verfahren erwiesen sich als gut benutzbar, wobei biometrische Plattformauthentifizierung insgesamt bevorzugt wurde. Zugleich zeigen sich zentrale Hürden für die Verbreitung, insbesondere bei der geräteübergreifenden Nutzung. Eine Analyse der Erweiterbarkeit von FIDO2 verdeutlicht zudem, dass aktuelle Browserbeschränkungen den Spielraum zur Anpassung an Einsatzszenarien erheblich einschränken.

Anschließend betrachten wir die Nutzung mobiler Dienste bei Ausfällen zentraler Infrastruktur, etwa während Naturkatastrophen, die lokal verbundene, aber global getrennte Netzwerkinseln entstehen lassen. Eine Umfrage (N = 857) zeigt, dass vertraute Apps gegenüber speziellen Krisen-Apps bevorzugt werden; Messenger-Apps wurden dabei insgesamt am höchsten priorisiert. Dies verdeutlicht den Bedarf an Messenger-Authentifizierung, die ohne zentralisierte Backend-Dienste auskommt.

Darauf aufbauend stellen wir PairSonic vor, ein dezentrales Authentifizierungsprotokoll für Ende-zu-Ende-verschlüsselte Messenger. PairSonic ermöglicht Personen vor Ort eine gegenseitige Authentifizierung, ohne vertrauenswürdige Dritte, Public-Key-Infrastruktur oder Internetzugang vorauszusetzen. In einer Laborstudie (N = 45) wurde PairSonic signifikant gegenüber dem Stand der Technik bevorzugt, da es den Aufwand manueller Verifikation reduziert und besser auf Gruppen skaliert. Da PairSonic einen akustischen Out-of-Band-Kanal nutzt, untersuchen wir außerdem die akustische Datenübertragung auf verschiedenen handelsüblichen Smartphones. Dabei identifizieren wir Mehrwegeausbreitung und Hardwareunterschiede als zentrale Zuverlässigkeitsprobleme und leiten robuste Entwurfsmuster für den praktischen Einsatz ab.

Insgesamt zeigt diese Dissertation, dass starke Authentifizierungsmaßnahmen praxistauglich werden, wenn kryptografisches Design, menschliche Interaktion und reale Einsatzbedingungen gemeinsam berücksichtigt werden.

Acknowledgments

This dissertation was written during my time as a researcher at the Secure Mobile Networking Lab at the Technical University of Darmstadt. First and foremost, my sincere thanks go to my advisor Prof. Dr.-Ing. Matthias Hollick for his support, for the trust he placed in me, and for fostering a highly independent way of working, allowing my ideas to flourish and thrive. I am also grateful to Univ.-Prof. Dr. René Mayrhofer, head of the Institute of Networks and Security at the Johannes Kepler University Linz, Austria, for agreeing to serve as co-examiner.

I would like to extend my heartfelt thanks to all my colleagues at SEEMOO for contributing to a constructive, welcoming atmosphere that encouraged discussion and joint problem-solving. Special thanks to Nils Rollshausen, Jakob Link, Lucas Becker, David Breuer, Leon Janzen, and Frank Hessel for giving helpful feedback on this dissertation, and to Eicke Hauck, Julian Euler, Valentin Conrad, and Lars Almon for their excellent administrative support. Thanks also to Ursula Paeckel and all the other hidden champions who kept our university running behind the scenes.

I am grateful to the more than 60 students and student assistants whom I supervised in labs and seminars, as well as in their bachelor's and master's theses. This experience has taught me a lot, and many of them left a mark on my research with their own contributions.

I am also grateful to the research project emergenCITY¹ for funding my research and for bringing me into contact with so many great colleagues through interdisciplinary discussions and collaborations: Dr. Steffen Haesler, Jan-Philipp Muttach, Dr. Nadja Thiessen, Tobias Gebhard, Dr. Lars Baumgärtner, Markus Sommer, Yi Wang, Jonathan Sattler, Burak Yilmaz, and many more. I would also like to thank Lea Holaus for illustrating my work with Figures 6 and 8. Special thanks to Dr. Anne Hofmeister, Katharina Kleinschnitger, Uta Drews, Gloria Philipp, Angela Zimmerling, and Janila Dierks for their invaluable organizational support.

I would like to thank the many interesting researchers I met along the way, both within TU Darmstadt and through collaborations with colleagues at the Johannes Kepler University Linz, the University of Cambridge, Fraunhofer AISEC, and the University of Strathclyde. These interdisciplinary and international collaborations have broadened my perspective, shaping the direction and quality of my research. I am equally grateful to all study participants whose time, feedback, and valuable data provided the foundation for my usability evaluations.

Thanks to my family and friends, especially to my brother, for their love and support!

This dissertation is dedicated to all who care about making security measures more accessible and easier to use for the general public.

¹ This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

Contents

Abstract	iii
Acknowledgments	v
List of Publications	vii
Open Science and Replicability	ix
Collaborations and My Contribution	x
1 Introduction	1
1.1 Application Scenarios	5
1.2 Thesis Outline	7
2 Web Authentication	9
2.1 Our Contributions	11
2.2 Related Work	20
3 Foundations for Decentralized Authentication	24
3.1 Our Contributions	25
3.2 Related Work	30
4 Group Messenger Authentication	33
4.1 Our Contributions	35
4.2 Related Work	43
5 Conclusions	48
5.1 Findings	48
5.2 Future Work	49
Bibliography	52
List of Figures	79
List of Tables	79
List of Acronyms	80
A FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones	82
B Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions	99
C The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity	116
D Sounds Good? Fast and Secure Contact Exchange in Groups	139
E PairSonic: Helping Groups Securely Exchange Contact Information	184
F Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices	188
Erklärung zur Dissertationsschrift	220

List of Publications

During the course of writing this dissertation, I co-authored several papers and articles that I list below. Asterisks (*) indicate equal contributions.

PUBLICATIONS INCLUDED IN THIS THESIS

- [A] L. WÜRSCHING*, F. PUTZ*, S. HAESLER, M. HOLLICK. “FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones.” In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Apr. 23–28, 2023). CHI ’23. Hamburg, Germany: ACM, Apr. 2023. ISBN: 978-1-4503-9421-5. DOI: [10.1145/3544548.3580993](https://doi.org/10.1145/3544548.3580993). CORE Rank A*. Best Paper Award. Appendix A.
- [B] F. PUTZ, S. SCHÖN, M. HOLLICK. “Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions.” In: *Emerging Technologies for Authorization and Authentication*. ETAA ’21. Cham, Switzerland: Springer International Publishing, 2021, pp. 17–32. ISBN: 978-3-030-93746-1. DOI: [10.1007/978-3-030-93747-8_2](https://doi.org/10.1007/978-3-030-93747-8_2). Appendix B.
- [C] L. JANZEN, F. PUTZ, M.-A. KAUFHOLD, K. STRAUB, M. HOLLICK. “The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity.” In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Apr. 26–May 1, 2025). CHI ’25. Yokohama, Japan: ACM, Apr. 2025. DOI: [10.1145/3706598.3714324](https://doi.org/10.1145/3706598.3714324). CORE Rank A*. Honorable Mention Award. Appendix C.
- [D] F. PUTZ, S. HAESLER, M. HOLLICK. “Sounds Good? Fast and Secure Contact Exchange in Groups.” In: *Proceedings of the ACM on Human-Computer Interaction* 8. CSCW ’24, Nov. 2024. DOI: [10.1145/3686964](https://doi.org/10.1145/3686964). CORE Rank A. Best Paper Award. Appendix D.
- [E] F. PUTZ, S. HAESLER, T. VÖLKL, M. GEHRING, N. ROLLSHAUSEN, M. HOLLICK. “Pair-Sonic: Helping Groups Securely Exchange Contact Information.” In: *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work & Social Computing* (Nov. 9–13, 2024). CSCW ’24. San José, Costa Rica: ACM, Nov. 2024, pp. 69–71. DOI: [10.1145/3678884.3681818](https://doi.org/10.1145/3678884.3681818). Appendix E.
- [F] F. PUTZ, P. FORTMANN, J. FRANK, C. HAUGWITZ, M. KUPNIK, M. HOLLICK. “Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices.” In: *ACM Transactions on Internet of Things* 7.1, Feb. 2026. DOI: [10.1145/3779439](https://doi.org/10.1145/3779439). Appendix F.

FURTHER PUBLICATIONS

- [G] F. PUTZ, F. ÁLVAREZ, J. CLASSEN. “Acoustic Integrity Codes: Secure Device Pairing Using Short-Range Acoustic Communication.” In: *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks* (July 8–10, 2020). WiSec ’20. ACM, July 2020, pp. 31–41. DOI: [10.1145/3395351.3399420](https://doi.org/10.1145/3395351.3399420). CORE Rank B.
- [H] S. HAESLER, R. MOGK, F. PUTZ, K. T. LOGAN, N. THIESSEN, K. KLEINSCHNITZER, L. BAUMGÄRTNER, J.-P. STROSCHE, C. REUTER, M. KNODT, M. HOLLICK. “Connected Self-Organized Citizens in Crises: An Interdisciplinary Resilience Concept for Neighborhoods.” In: *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (Oct. 23–27, 2021). CSCW ’21. ACM, 2021, pp. 62–66. DOI: [10.1145/3462204.3481749](https://doi.org/10.1145/3462204.3481749).
- [I] A. HEINRICH, S. KROLLMANN, F. PUTZ, M. HOLLICK. *Smartphones with UWB: Evaluating the Accuracy and Reliability of UWB Ranging*. Technical Report. 2023. DOI: [10.48550/arXiv.2303.11220](https://doi.org/10.48550/arXiv.2303.11220).
- [J] M. SCHECK, F. PUTZ, F. HESSEL, H. LEINWEBER, J. CRYSTALL, M. HOLLICK. *DEMO: Secure Bootstrapping of Smart Speakers Using Acoustic Communication*. Demo presented at the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec 2023). Guildford, Surrey, United Kingdom, 2023. DOI: [10.26083/tuprints-00024180](https://doi.org/10.26083/tuprints-00024180).
- [K] E. LEIERZOPF, M. ROLAND, R. MAYRHOFER, F. PUTZ. “A Large-Scale Data Collection and Evaluation Framework for Android Device Security Attributes.” In: *IDIMT-2023: New Challenges for ICT and Management: 31st Interdisciplinary Information Management Talks* (Sept. 6–8, 2023). Hradec Králové, Czech Republic, Sept. 2023, pp. 63–72. DOI: [10.35011/IDIMT-2023-63](https://doi.org/10.35011/IDIMT-2023-63).
- [L] H. LEINWEBER*, J. CRYSTALL*, F. HESSEL, F. PUTZ, M. HOLLICK. “Introducing FreeSpeaker - A Modular Smart Home Hub Prototyping Platform.” In: *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking* (Oct. 2–5, 2023). MobiCom ’23. Madrid, Spain: ACM, Oct. 2023. DOI: [10.1145/3570361.3614080](https://doi.org/10.1145/3570361.3614080).
- [M] E. LEIERZOPF, R. MAYRHOFER, M. ROLAND, W. STUDIER, L. DEAN, M. SEIFFERT, F. PUTZ, L. BECKER, D. R. THOMAS. “A Data-Driven Evaluation of the Current Security State of Android Devices.” In: *2024 IEEE Conference on Communications and Network Security* (Sept. 30–Oct. 3, 2024). CNS ’24. Taipei, Taiwan: IEEE, 2024, pp. 1–9. DOI: [10.1109/CNS62487.2024.10735682](https://doi.org/10.1109/CNS62487.2024.10735682).
- [N] J. LINK, F. PUTZ, M. HOLLICK. “Pair-Fi: Integrity Code Protected Secure Device Pairing via SDR-Enabled Wi-Fi Chips on Smartphones.” In: *Proceedings of the 19th ACM Conference on Security and Privacy in Wireless and Mobile Networks* (June 30–July 3, 2026). WiSec ’26. Saarbrücken, Germany: ACM, 2026. DOI: [10.1145/3765613.3811679](https://doi.org/10.1145/3765613.3811679). CORE Rank B. Best Paper Award.
- [O] M. HENKEL, S. HAESLER, F. PUTZ, C. REUTER. “Is It There Yet? Designing a User Interface For Messenger Apps That Use Disruption-Tolerant Networking in Crises.” In: *Proceedings of the ACM on Human-Computer Interaction* 10. CSCW ’26, Oct. 2026. DOI: [10.1145/3816955](https://doi.org/10.1145/3816955). CORE Rank A.

Open Science and Replicability

The full versions of all papers described in this dissertation are freely available on arXiv, at the ACM Digital Library, or through dedicated project websites at <https://fputz.net>. All datasets, including user study results and experimental measurements, are permanently archived on Zenodo with DOIs for reliable long-term access. The source code of all software implementations is publicly available on GitHub at <https://github.com/seemoo-lab/>, with repositories clearly linked in each publication.

Our work in [F] specifically advances reproducibility standards for acoustic communication research by providing complete experimental methodology and a reference dataset of over 11 000 smartphone transmissions, addressing a domain where unrealistic testing conditions have historically impeded progress. By releasing all artifacts under permissive licenses, we enable replication and reuse by researchers and practitioners, ultimately accelerating the deployment of more usable authentication systems.

Open Science. Boxes like this appear throughout the dissertation to highlight contributions to open science and replicability.

DATASETS

- [DS1] L. WÜRSCHING, F. PUTZ, S. HAESLER, M. HOLLICK. *Lab Study Dataset: FIDO2 Platform and Roaming Authentication on Smartphones*. Zenodo, Jan. 2023. DOI: [10.5281/zenodo.7572697](https://doi.org/10.5281/zenodo.7572697).
- [DS2] L. JANZEN, F. PUTZ, K. STRAUB, M.-A. KAUFHOLD, M. HOLLICK. *Survey Dataset: The User Perspective on Island-Ready 6G Communication*. Zenodo, Feb. 2025. DOI: [10.5281/zenodo.14812894](https://doi.org/10.5281/zenodo.14812894).
- [DS3] F. PUTZ, S. HAESLER, M. HOLLICK. *Lab Study Dataset: Fast and Secure Contact Exchange in Groups*. Zenodo, Aug. 2024. DOI: [10.5281/zenodo.13324112](https://doi.org/10.5281/zenodo.13324112).
- [DS4] F. PUTZ, P. FORTMANN, J. FRANK, C. HAUGWITZ, M. KUPNIK, M. HOLLICK. *Replication Package: Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices*. Zenodo, Nov. 2025. DOI: [10.5281/zenodo.17661991](https://doi.org/10.5281/zenodo.17661991).

SOFTWARE

- [SW1] L. WÜRSCHING, F. PUTZ, S. HAESLER, M. HOLLICK. *FIDO2 The Smartphone*. 2023. URL: <https://github.com/seemoo-lab/fido2-the-smartphone>.
- [SW2] F. PUTZ, S. SCHÖN, M. HOLLICK. *Bring Your Own FIDO2 Extensions*. 2021. URL: <https://github.com/seemoo-lab/fido2ext>.
- [SW3] F. PUTZ, S. HAESLER, T. VÖLKL, M. GEHRING, N. ROLLSHAUSEN, M. HOLLICK. *Pair-Sonic*. 2024. URL: <https://github.com/seemoo-lab/pairsonic>.

Collaborations and My Contribution

This dissertation is based on six scientific publications that I co-authored with my advisor Matthias Hollick and several collaborators. I am grateful to all co-authors for their significant contributions to these productive and pleasant collaborations. Across these publications, my contributions span the conceptual development of research ideas, empirical study design, quantitative and qualitative analysis, software implementation, artifact release, manuscript preparation, and project coordination.

In the following, I summarize my specific contributions to each publication.

Chapter 2 is based on joint work with Leon Janzen (born Würsching), Steffen Haesler, and Matthias Hollick [A], as well as Steffen Schön and Matthias Hollick [B].

[A] is partially based on the master's thesis of Leon Janzen, which Matthias Hollick and I co-supervised. Leon Janzen and I contributed equally as co-first authors and jointly led the overall research design, project coordination, and manuscript preparation. My main contributions were the lab study's quantitative analysis, statistical implementation in R, and preparation of the dataset and replication package [DS1]. Leon Janzen was mainly responsible for participant recruitment, in-person lab sessions, qualitative analysis, and implementing the prototype used in the study. Together, Leon Janzen and I drafted most of the manuscript and coordinated revisions with the co-authors. Steffen Haesler and I led the user study design and methods. He also ensured data quality, helped interpret the empirical results, and supported writing the methods section. Matthias Hollick advised the project and provided feedback throughout the research and writing process.

For [B], I served as the lead author and was responsible for the primary research design, project coordination, and manuscript preparation. My main contributions were the survey of existing FIDO2 extensions and their authenticator compatibility, the analysis of FIDO2 extension support in popular web browsers, the source code inspection and implementation of our custom extension in the Chromium web browser, and the design of all figures and tables. The exploratory work that contributed to this publication was conducted by Steffen Schön as part of his master's thesis, which Matthias Hollick and I co-supervised. Steffen Schön implemented the custom extensions for the hardware tokens, the Python library, and the relying party, and provided feedback on the manuscript. Matthias Hollick advised the project and provided feedback throughout the research and writing process. I prepared the open-source release of the implementation and the corresponding documentation [SW2].

Chapter 3 is based on joint work with Leon Janzen, Marc-André Kaufhold, Kolja Straub, and Matthias Hollick [C]. As lead author, Leon Janzen defined the research questions and led the study design, evaluation, and writing process. He defined the conceptual model of island connectivity, derived the smartphone service typology, and led the discussion. My main contributions were the quantitative evaluation and the statistical analysis of the

study in R, including the design of the corresponding figures and tables. I also provided feedback and led the publication of the study dataset [DS2]. Marc-André Kaufhold contributed to the review of related work and the corresponding sections, and provided feedback on the questionnaire and manuscript. As part of his master's thesis, Kolja Straub co-designed and conducted the study and provided feedback on the manuscript. Matthias Hollick advised the project, contributed to its conceptual development, and provided feedback throughout the research and writing process.

Chapter 4 is based on joint work with Steffen Haesler and Matthias Hollick [D], alongside a demo paper at the same conference with additional contributions from Thomas Völkl, Maximilian Gehring, and Nils Rollshausen [E]. It also includes joint work with Philipp Fortmann, Jan Frank, Christoph Haugwitz, Mario Kupnik, and Matthias Hollick [F].

In [D], I served as the lead author and was responsible for the paper's overall research design, management, study execution, data analysis, and writing process. Steffen Haesler advised on the research design and contributed to the qualitative data analysis and the methods section. Matthias Hollick advised the project and provided feedback throughout the research and writing process. I prepared the first complete manuscript draft and coordinated the revision process with the co-authors.

For our demonstrator implementation [E], I conceptualized the core idea of PairSonic and led the design and development of the PairSonic Android app. Steffen Haesler, Thomas Völkl, Maximilian Gehring, and Nils Rollshausen supported me with the software design and implementation.¹ Matthias Hollick advised the project and provided feedback throughout the research and writing process. I prepared the manuscript, the replication package [DS3], and the open-source release of PairSonic [SW3].

[F] is partially based on the master's theses of Philipp Fortmann and Jan Frank, both of which were jointly supervised by Matthias Hollick and me. As lead author, I led the research design, management, and writing process. I developed the core idea of the paper, led the evaluation of the audio transmissions, developed the corresponding analysis methodology, implemented the MATLAB analysis pipeline, and performed the statistical visualization of the results in R. As part of his master's thesis, Philipp Fortmann executed the majority of the experiments, conducted the systematic literature study, and contacted authors of prior work to request implementations. As part of his master's thesis, Jan Frank developed the three re-implementations and contributed the discussion of the corresponding implementation challenges. Christoph Haugwitz and Mario Kupnik supported the measurements in the anechoic chamber and provided feedback on the experimental setup. Matthias Hollick advised the project and provided feedback throughout the research and writing process. I drafted the manuscript, integrated co-author feedback during the revision process, and prepared the replication package [DS4] to make the experimental methodology and analysis reproducible.

¹ The PairSonic app grew out of an excellent student project by Maximilian Gehring and was further developed with substantial contributions from my student assistant Thomas Völkl, both of whom I supervised.

Introduction

Security experts have built robust cryptographic foundations for the digital world. Protocols like Transport Layer Security (TLS) solve fundamental problems of secure communication across global networks, without requiring trust in intermediate network nodes [Res18]. These technical achievements protect billions of users daily, operating at unprecedented scale: from instantaneous payments and e-commerce to secure messaging and telemedicine, all conveniently accessible from users' smartphones.

Yet this cryptographic promise often fails at the human interface. Concerning evidence emerges from studies of real-world security tools: when researchers examined the Signal smartphone messenger in 2018, they discovered that 75% of participants could not locate the contact verification screen necessary to establish authentic end-to-end encrypted (E2EE) chats [Vaz+18]. A 2023 follow-up found similar usability barriers also in other popular messengers [FK23]. Similar patterns appear across dozens of controlled studies on encrypted email, passwords, and multi-factor authentication [Gut+22; RIB22]. These observations are not isolated incidents but highlight a broader misalignment between cryptographic design and human interaction. Security mechanisms compromise their protective goals when they impose significant usability barriers. This human interface failure becomes more pronounced as computing shifts from controlled desktop environments to ubiquitous mobile devices [Sta26]. While early cryptographic protocols assumed users with keyboards and full screens, today's security interactions occur primarily through *smartphones*: devices with fundamentally different constraints and usage patterns.

As security experts, we share responsibility when protocols with strong theoretical guarantees fail in practice; such outcomes represent systemic failures rather than user shortcomings. The decades-long struggle to deploy Pretty Good Privacy (PGP) exemplifies this gap. Despite praise for its strong cryptographic foundation, end-to-end encrypted email fails to gain widespread adoption due to usability constraints [RS19; WT99].

This pattern points to a fundamental challenge in our discipline. Historically, our analyses have centered on cryptographic robustness—understandably, given our primary expertise—while practical concerns of user interaction and system deployment have received less attention [Gut+22; Las+24; Naq20]. This gap was characterized by Bonneau et al. in their seminal paper on web authentication:

"In our experience, security experts focus more on security but less on usability and practical issues related to deployment." — Bonneau et al. [Bon+12]

While our community's rigorous analysis of security properties is essential, the practical bottleneck for security often lies elsewhere: in the design of measures that users can practically and willingly adopt.

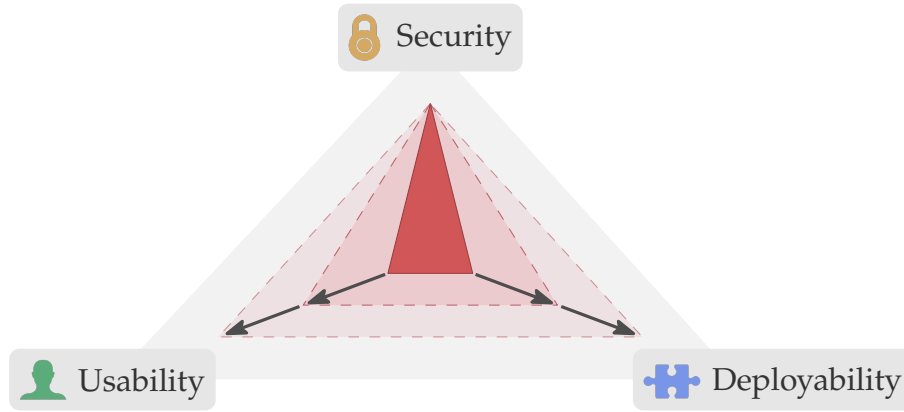


Figure 1: We aim to increase the usability and deployability of authentication systems. Existing security measures often lack in these two aspects, as shown by Bonneau et al. [Bon+12] for the case of web authentication.

To confront this disciplinary challenge, this dissertation focuses on the security component where human interaction is most direct and failure is most visible: *authentication*. Almost every type of online service requires user accounts, making authentication both the most user-facing security component and a central topic in the usable security community [Dro23; Rib22]. Whether through passwords, two-factor authentication, or cryptographic proofs, authentication typically constitutes users’ first security interaction with a service. Surprisingly, many protocols gloss over this step and simply assume that authentic key material exists. *How* this authentication is achieved is rarely specified in practice:

“[T]he problem of distributing and managing keys is one of the really difficult parts of cryptography, for which we have only partial solutions.”

— Ferguson, Schneier, and Kohno [FSK10]

This oversight creates a dangerous gap: strong authentication underpins nearly all security protocols, yet its human implementation remains the weakest link [Ruo+19]. We address this gap through a holistic approach that elevates all three core properties to first-class design goals (illustrated in Figure 1). Alongside traditional *security* (the system’s resilience against technical attacks), we prioritize *usability* (the ease with which humans can use the system correctly and willingly) and *deployability* (the feasibility of adopting a system that is compatible with existing hardware and software, with minimal reliance on trust and network infrastructure). Guided by this user-centric perspective, we design and evaluate authentication schemes that satisfy the operational constraints of both end users and service providers, thereby strengthening practical security.

Usable Authentication. Some security experts might argue that usability concerns fall outside their responsibility, belonging instead to user experience (UX) specialists who join later in the development cycle. However, when an authentication scheme relies on human cooperation—such as choosing a password, approving a push notification, or safeguarding a hardware token—its security boundary expands to include human behavior [Ell07]. When designing such systems without considering how real people actually use (or

attempt to circumvent) them, the overall security of the system often *degrades* [WT99]. Therefore, usability is an important concern also from a security perspective.

Ignoring usability creates new attack vectors that cryptography alone cannot solve. When cryptographic keys must be associated with and managed by users, key establishment transforms from a primarily technological problem into a complex socio-technical challenge. Attackers exploit this vulnerability relentlessly: social engineering and phishing succeed not by breaking the underlying cryptography but by manipulating confused or overloaded users. These attacks belong to the most prevalent attack types according to the Federal Bureau of Investigation (FBI) [Fed25] and the German Federal Office for Information Security (BSI) [BSI25]. In 2024 alone, 68% of confirmed data breaches involved human error or misuse [Hyl+24], showing that the human element is often the most vulnerable component in a security chain.

Regulatory bodies now formally recognize this reality. The National Institute of Standards and Technology (NIST) explicitly mandates: *“Evaluating the usability of authentication is critical, as poor usability often results in coping mechanisms and unintended workarounds that can ultimately degrade the effectiveness of security controls.”* (US NIST SP 800-63B-4 §8.1). Consequently, the usable security field has seen remarkable growth, both in academia [DTO23; Gut+22; RIB22; Sur+24] and industry [Naq20].

Deployable Authentication. A practical authentication system must not only be secure and usable, it must also be deployable. From the user’s perspective, it should be compatible with existing smartphones and operating systems, without requiring specialized hardware, non-standard software, or even the purchase of new devices [Bon+12]. Accessibility for users with visual or physical impairments is also essential for broad deployment [Zim+18]. Service providers, in turn, must be able to adopt the system with existing equipment and infrastructure at acceptable cost, including ongoing maintenance and support [Las+24]. To lower barriers to entry, the system should be mature and support gradual migration from legacy architectures [Bon+15].

Deployability also benefits from architectural resilience. Authentication systems become fragile when they are tightly coupled to external dependencies such as centralized trust infrastructure. This coupling not only threatens availability during outages but also increases operational risks and maintenance costs [Las+24]. To mitigate this, designs should favor decentralization so authentication remains possible when external infrastructure fails or is unreachable. For example, an instant messaging system should let two nearby users authenticate each other, even while hiking without mobile Internet coverage.

Although deployability is hard to quantify because individual requirements vary by application scenario [Bon+15], the overarching goal is reliable operation with minimal operational costs for both service providers and users [Las+24]. When we neglect deployability, security measures remain largely theoretical or achieve only partial adoption [Aeb+20]. To illustrate this, almost every password alternative examined by Bonneau et al. demonstrated superior security properties but scored poorly on deployability [Bon+12], which explains why passwords remain omnipresent despite their well-documented flaws.

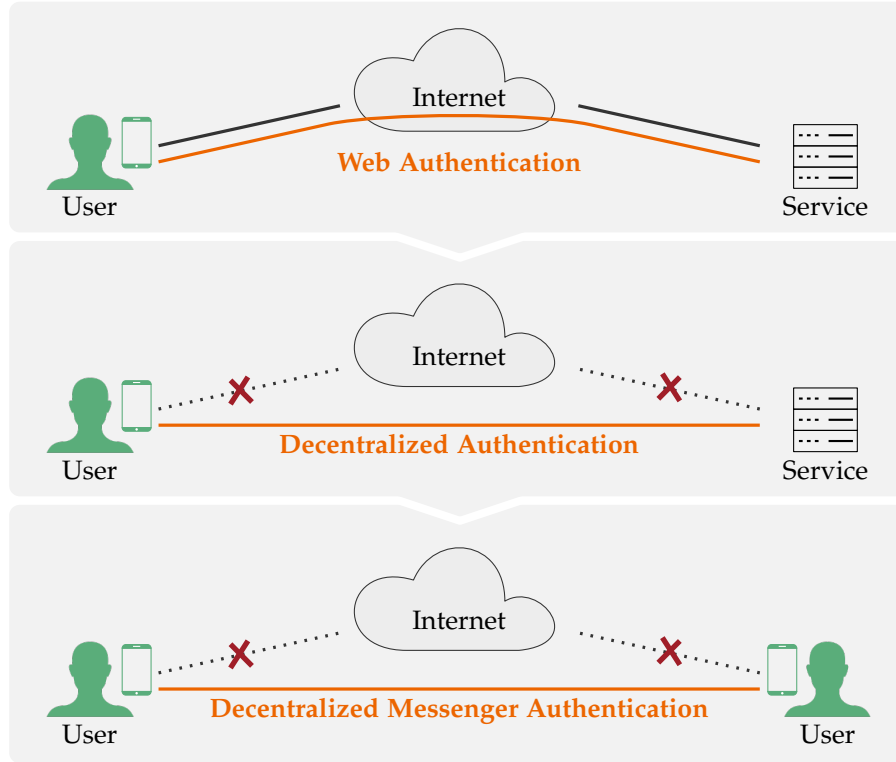


Figure 2: Overview of the authentication scenarios from the user’s perspective, illustrating the progression from classical web authentication to resilient, decentralized messaging.

Goal. To address these systemic challenges, the goal of this dissertation is to advance the practicality of mobile authentication protocols by treating usability and deployability as first-class design goals. Given the dominance of smartphones as the primary platform for Internet access [Sta26], we focus specifically on mobile devices, recognizing that they have not merely replaced desktop computers but fundamentally altered how users expect to interact with security systems. To this end, we examine three application scenarios that lack practical authentication solutions in terms of usability and deployability. For each scenario, we empirically study user perceptions to identify the precise user requirements and shortcomings that hinder real-world adoption. Guided by these insights, we then design a novel authentication protocol, implement a smartphone prototype, and evaluate it in a lab study, showing significant user preference compared to the state of the art.

To facilitate scrutiny and support future research, all papers, datasets, implementation source code, and evaluation scripts associated with this dissertation are publicly available. This transparency enables replication and demonstrates that the systemic issues highlighted in this introduction are solvable through rigorous, user-centric design.

In the following, we first introduce each studied scenario at a high level, then outline our contributions towards making these applications more user-friendly and deployable.

1.1 Application Scenarios

Mobile authentication occurs in two fundamental contexts: *user-to-service communication* and *user-to-user communication*, each presenting distinct usability and deployability challenges. Because these challenges become especially visible when infrastructure is degraded, we also examine how such constraints shape user priorities, revealing insights that guide our authentication research.

In this dissertation, we study three interconnected scenarios, depicted in Figure 2. First, we consider the classical problem of web authentication at online services (user-to-service communication). Second, we study which smartphone services users prioritize when network infrastructure is limited, such as during natural disasters. This analysis identifies the application contexts in which resilient authentication is most urgently needed and thereby motivates our subsequent focus on decentralized messenger authentication. Third, we examine E2EE instant messaging (user-to-user communication), proposing a decentralized authentication solution that is particularly effective for groups of users wanting to communicate securely without relying on centralized trust infrastructure.

In the following, we briefly describe each scenario, highlighting the associated usability and deployability challenges. Each scenario gives rise to one research question that structures the contributions of this dissertation.

Web Authentication. Most online services require user accounts, meaning users must authenticate to prove their identity to the service provider. Strong authentication is crucial to prevent unauthorized access to sensitive *information* (e. g., medical records, financial data, location history), but also prevent unauthorized *actions* (e. g., payments). Without strong authentication, malicious access to cloud storage, email accounts, or calendars can expose users to financial loss as well as legal and privacy risks [Tem+25b].

Password authentication currently dominates due to excellent deployability but suffers from severe security vulnerabilities: adversaries regularly compromise passwords through phishing attacks, such as by creating convincing fake websites that trick users into entering credentials, or through breaches of online services. The website “Have I Been Pwned?” [Hun26] documents over 17 billion exposed credentials from more than 950 data breaches, many of which can be abused across websites due to password reuse.

Despite numerous attempts to replace passwords, proposed alternatives have failed to achieve widespread adoption, often because they were difficult to use or deploy. The world’s largest technology companies recently formed an alliance to develop FIDO2 (branded as “passkeys”), a password replacement based on public-key cryptography that resists phishing, keylogging, and data breaches [FID24a]. However, its practical adoption depends not only on these security properties, but also on whether users and service providers can use and deploy it effectively. This motivates the first research question:

RQ1: *What usability and deployability barriers hinder the adoption of FIDO2 authentication on smartphones?*

Decentralized Authentication. Current authentication solutions often rely on centralized infrastructure such as certificate authorities, identity providers, and cloud services. When this infrastructure becomes unreachable, users lose access to important smartphone services, making centralization a significant *deployability* risk.

Natural disasters and other crises can damage network infrastructure and create “*island networks*”: self-contained local networks that are disconnected from the global Internet. In such challenging environments, communication and other smartphone services may still be useful within the local network, but only if they can operate independently of remote infrastructure. However, current online service deployments and cellular networks largely do not support this mode of operation.

The first step towards designing authentication systems for these conditions is to identify user requirements. Specifically, we must determine which services users prioritize when infrastructure fails, since different services require distinct forms of authentication: web services require user-to-service authentication, whereas peer communication requires user-to-user authentication. This motivates the second research question:

RQ2: *Which smartphone services do users prioritize when network infrastructure fails, and what do these priorities imply for resilient authentication design?*

Decentralized Messenger Authentication. Smartphone messaging applications like WhatsApp, Telegram, and Signal provide E2EE communication, promising that when users communicate one-to-one or in groups, nobody else can read their messages, not even the messaging service provider. However, this security guarantee depends on users verifying their contacts’ identities through an *authentication ceremony*, which typically involves physical meetings and manual verification tasks.

Unfortunately, current authentication ceremonies are difficult to use and scale poorly beyond pairwise verification [AS23]. This creates a particular challenge for group conversations, where users may need to establish trust among multiple participants rather than verify only a single contact. As a result, users rarely complete this crucial step, leaving conversations vulnerable to active machine-in-the-middle (MitM) attacks. Recent reports of law enforcement agencies compromising service providers demonstrate that active attacks are not merely theoretical, highlighting urgent privacy risks [FK23].

If we want to strengthen the security of messaging applications, we must design a practical and user-friendly authentication ceremony that can support multiple users simultaneously. From a deployability perspective, such ceremonies must minimize external dependencies, functioning without trusted third parties, public-key infrastructures (PKIs), pre-shared secrets, or Internet connectivity. This infrastructure independence becomes crucial in island connectivity scenarios where centralized services may be unreachable. This motivates the third research question:

RQ3: *How can authentication ceremonies for messengers be made usable and scalable for groups while remaining deployable without centralized infrastructure?*

1.2 Thesis Outline

The remaining chapters describe our contributions to make each communication scenario presented in § 1.1 more practical by improving usability and deployability. We briefly summarize the chapters as follows:

Chapter 2. We address RQ1 by investigating FIDO2 in terms of usability [A] and deployability [B]. We present the first controlled lab study (N = 87) comparing FIDO2 *platform authentication* (Touch ID) versus *roaming authentication* (NFC YubiKey) on smartphones, to determine their practical strengths and weaknesses as perceived by end users [A]. Most participants expressed willingness to adopt passwordless authentication, but our analysis reveals that users prioritize usability and security differently depending on the type of online service. We identify critical adoption barriers including missing support for account delegation and multi-device usage patterns, which the industry has since begun addressing through cloud-synced *passkeys*.

Furthermore, in [B] we evaluate FIDO2's extensibility for custom authentication deployments. As the process for creating FIDO2 extensions remains largely undocumented, we demonstrate how to design and implement tailor-made authentication leveraging the existing FIDO2 ecosystem through Chromium-based browsers, the `python-fido2` client, and two off-the-shelf security tokens. We present the first comprehensive compatibility audit of FIDO2 extensions, surveying all 21 publicly known extensions by analyzing source code from 10 mainstream FIDO2 clients and authenticators, identifying limitations in current web browsers that inhibit innovative extension development.

We maintain a project website at <https://fputz.net/project/fido2-usability> with comprehensive explanations and all research artifacts [DS1; SW1; SW2]. The work in [A] received the CHI 2023 Best Paper Award (top 1%).

Chapter 3. We address RQ2 by investigating which smartphone services users prioritize when network infrastructure fails, establishing which services need resilient authentication solutions. Current smartphone apps depend on remote servers and fail when infrastructure becomes unreachable, which is a critical deployability concern during emergencies that damage networks and create isolated connectivity islands.

We present the first empirical study (N = 857) examining which smartphone apps citizens prioritize in areas with only local connectivity [C]. In both normal and local-only connectivity scenarios, citizens rank messenger apps highest out of 23 app categories. We also identify apps used for crisis-specific tasks, as disasters often leave only local connectivity available. Our findings guide network operators in prioritizing services for isolated areas and identify stakeholders needed to deploy island connectivity in practice. Crucially, the dominance of messaging applications motivates our subsequent focus on authentication for user-to-user communication.

We make our research artifacts available online [DS2]. The work in [C] received the CHI 2025 Honorable Mention Award (top 5%).

Chapter 4. We address RQ3 by introducing PairSonic, an innovative authentication ceremony for E2EE messaging designed to improve the usability and scalability of current approaches [D; E]. PairSonic achieves superior deployability by eliminating dependencies on trusted third parties, additional infrastructure, or Internet connectivity, making it suitable for island network scenarios. Our protocol is faster and simpler than existing solutions, particularly for larger groups, as it reduces the number of manual verification actions from $O(n^2)$ pair-wise comparisons to a single group-wise action.

We implement a smartphone prototype and conduct a lab study ($N = 45$) to compare PairSonic against the state of the art [Far+13], demonstrating significantly higher preference and usability scores [D; E]. To achieve infrastructure independence, we design a novel communication stack enabling secure device-to-device communication through an acoustic out-of-band (OOB) channel.

Unfortunately, our study showed that the acoustic channel in our prototype was sensitive to practical usage conditions: transmission failures occurred especially when participants positioned their devices unfavorably or when background noise interfered with the signal. Therefore, in [F] we systematically evaluate ad-hoc acoustic communication between nearby smartphones to identify robust solutions. We contribute a repeatable evaluation method to test reliability under realistic conditions, including user handling effects. We review 31 acoustic communication studies, re-implement promising candidates, and evaluate eight representative schemes through over 11 000 smartphone transmissions across 5 smartphone models, 6 transmission environments, and 8 realistic noise conditions. Our findings reveal significant challenges in practical usage: while many schemes perform well in controlled settings, they struggle to generalize across different smartphone hardware, varying transmission distances, and realistic background noise.

We maintain a project website at <https://fputz.net/project/pairsonic> with comprehensive explanations and all research artifacts [DS3; DS4; SW3]. A demo video showcasing PairSonic’s acoustic pairing is available on YouTube at <https://youtu.be/e1AMYDLWN0E>. The work in [D] received the CSCW 2024 Best Paper Award (top 1%).

Chapter 5 concludes this thesis by reflecting on the gap between cryptographic theory and human reality, and suggesting directions for future work in usable authentication.

Web Authentication

Online services often require user accounts, especially when they handle private or confidential information such as banking records, tax returns, or medical insurance. In these cases, authentication is necessary to prevent unauthorized access to sensitive *information* (e.g., medical records, financial data, location history) and unauthorized *actions* (e.g., online shopping orders and payments). Without strong authentication, users face risks such as financial loss from fraudulent purchases, as well as targeted espionage or extortion when attackers gain access to cloud storage, email, or calendars. Even seemingly harmless accounts, such as social media accounts, carry serious risks because impersonation can cause substantial reputational damage.

The most common form of web authentication remains password-based authentication, which is easy to deploy. Unfortunately, passwords have several serious security weaknesses that exploit users as the weakest link. First, adversaries regularly steal passwords through phishing attacks by creating convincing copies of legitimate websites that trick users into entering their credentials. Second, because managing unique passwords is burdensome, users often reuse them across services, allowing adversaries who compromise one account to access others [Das+14; FH07; Was+16]. Third, server-side data breaches can expose large numbers of credentials at once. The website “*Have I Been Pwned?*” (created by Troy Hunt) lists more than 17 billion exposed credentials from over 950 separate breaches [Hun26], many of them available to criminals. Despite numerous efforts to replace passwords, alternatives have failed to achieve widespread adoption because they often provide poorer usability or deployability [Bon+12].

In response to these weaknesses, an industry consortium of more than 250 technology companies and browser vendors developed the Fast IDentity Online 2 (FIDO2) standards for passwordless authentication [FID21; Wor21], now commonly known as *passkeys* [FID24a]. These standards are supported by all major web browsers and operating systems [Dev26]. Rather than relying on a shared secret such as a password, FIDO2 uses public-key cryptography: during registration, an authenticator generates a key pair for a *relying party* (e.g., a website); during login, it proves possession of the private key by signing a server challenge. As a result, FIDO2 is phishing-resistant and avoids long-standing problems of password-based authentication, including password reuse and the burden of memorizing high-entropy secrets [DTH06; IWS04; Yan+04].

FIDO2 supports two authentication variants with fundamentally different user interactions. In *roaming authentication*, private keys are stored on an external authenticator, such as a YubiKey [Yub20], which connects to the client device via Universal Serial Bus (USB), NFC, or Bluetooth Low Energy (BLE) [FID21]. In *platform authentication*, the smartphone’s integrated secure element stores the private keys, protected by a local authentication mechanism using the smartphone’s unlock mechanism, such as Apple Touch ID [App20a].

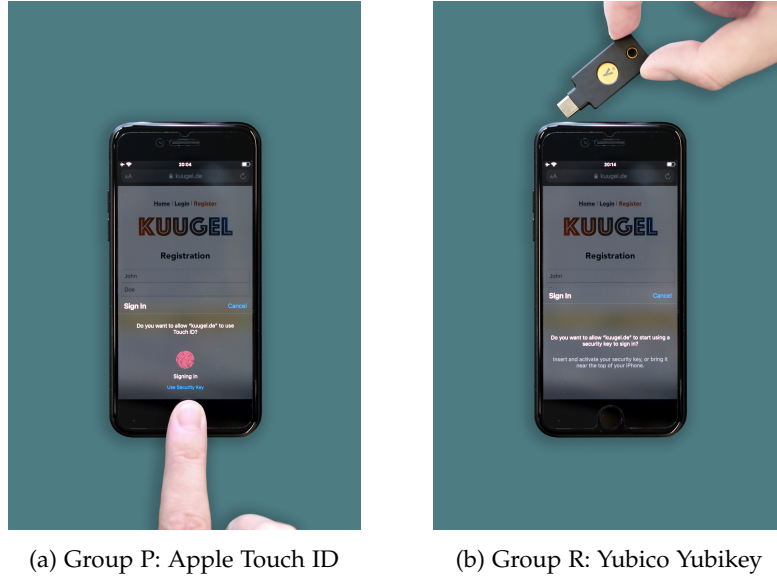


Figure 3: User interactions with the smartphone during the experiment. Participants either used Apple Touch ID (platform authenticator) or a Yubico YubiKey (roaming authenticator).

Both approaches replace knowledge-based authentication with possession-based authentication and may incorporate biometrics. However, they differ substantially in how users experience setup, login, portability, and recovery in everyday use.

Despite broad FIDO2 support on recent iPhones and Android devices [App24a; FID19], previous research has focused primarily on desktop clients [Lya+20]. This leaves an important gap: smartphones are now a common access device for online services, but findings from desktop settings do not necessarily transfer to mobile use. It therefore remains unclear how usable FIDO2 authentication is on smartphones and whether practical deployability barriers hinder its widespread adoption.

At the same time, usable security research has recognized that *tailorability* is necessary to address the wide range of real-world usage contexts [RIB22]. For FIDO2, authentication requirements vary across general-purpose use, internal company deployments, and research prototypes for entirely new use cases. As the FIDO2 ecosystem grows, new authentication approaches may require the standard to adapt, for example by exchanging additional data between authenticators and the relying parties. Yet current infrastructure severely limits the practical extensibility of FIDO2, making it difficult for developers and researchers to design custom extensions [Kli+20; Yeo+23].

In this chapter, we address both dimensions of our first research question:

RQ1: *What usability and deployability barriers hinder the adoption of FIDO2 authentication on smartphones?*

First, we evaluate the usability of FIDO2 authentication on smartphones through a lab study of platform and roaming authentication (cf. § 2.1.1). Second, with respect to deployability, we analyze and practically demonstrate FIDO2’s extension capabilities for custom deployments (cf. § 2.1.2), an area that remains largely undocumented [Kli+20]. Finally, we contextualize these efforts through related work (cf. § 2.2).

2.1 Our Contributions

2.1.1 Evaluating FIDO2 Platform and Roaming Authentication on Smartphones

This section summarizes the main findings of our publication on the strengths and weaknesses of FIDO2 passwordless authentication on smartphones, included in Appendix A:

- [A] L. WÜRSCHING*, F. PUTZ*, S. HAESLER, M. HOLLICK. “FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones.” In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Apr. 23–28, 2023). CHI ’23. Hamburg, Germany: ACM, Apr. 2023. ISBN: 978-1-4503-9421-5. DOI: [10.1145/3544548.3580993](https://doi.org/10.1145/3544548.3580993). CORE Rank A*. Best Paper Award. Appendix A.

We conducted the first lab study (N = 87) comparing FIDO2’s two smartphone authentication methods to examine how users respond to the shift from knowledge-based authentication to possession- and biometric-based factors. After completing practical web authentication tasks on an iPhone and a follow-up questionnaire, participants rated both methods as highly usable, though they generally preferred biometric platform authentication. We further identify adoption barriers that prevent FIDO2 from replacing passwords, including limited support for common practices such as account sharing and multi-device use. In the following, we summarize our main findings in more detail.

Study Design. We conducted an IRB-approved in-person lab study (N = 87) comparing platform and roaming passwordless authentication on smartphones (see Figure 3), combined with a post-task questionnaire. Participants completed practical web authentication tasks on an iPhone SE (2nd gen) using a purpose-built mock website that supported FIDO2 registration and login. Our study used a between-groups design with random assignment: one group authenticated with Touch ID as a platform authenticator, while the other used a YubiKey 5C NFC as a roaming authenticator. Sessions were conducted one-on-one in neutral meeting rooms with standardized hardware.

At the start of each study, we obtained consent, followed by Touch ID fingerprint enrollment for the platform group. Afterwards, all participants watched a standardized five-minute instruction video to ensure consistent information delivery. Participants then registered and logged into the mock website using their assigned method, i. e., fingerprint or YubiKey. After completing the tasks, they filled out a questionnaire covering usability, acceptance, likelihood of adoption across 11 account types, and open-ended reflections on strengths, weaknesses, and adoption barriers. We also collected four demographic measures and four control variables, including the affinity for technology interaction (ATI) [FAW19] and prior two-factor authentication (2FA) experience. In addition, we conducted a pilot study (N = 6) to validate the setup and materials.

Of the 87 participants, 63% identified as male and 37% as female. The sample was relatively highly educated, with 56% holding at least a Bachelor’s degree. Roughly half were students. Ages ranged from 18 to 69 years, and 34% were older than 30.

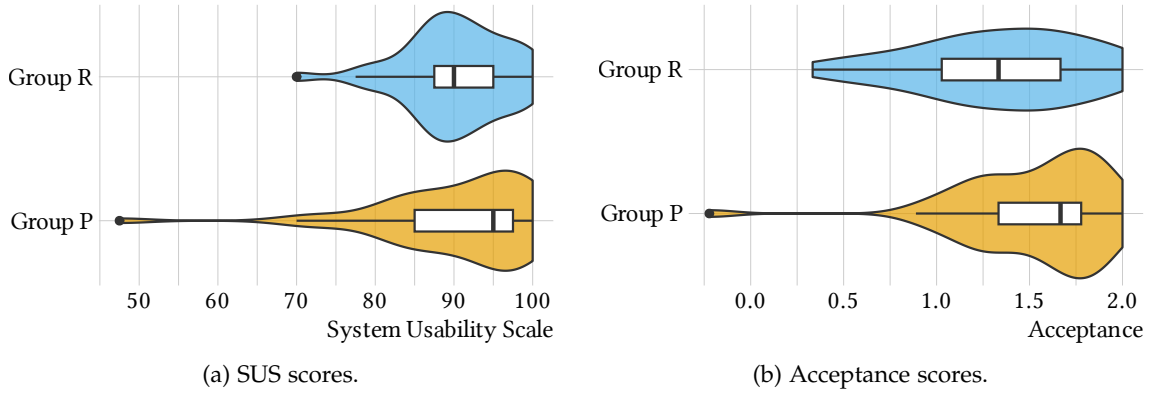


Figure 4: Comparison of participants' SUS and acceptance scores for roaming authentication (Group R) and platform authentication (Group P). Violin plots show density estimates of the distributions; boxplots show quartiles, medians, and outliers.

Usability and Acceptance. After the authentication tasks, participants in both groups rated FIDO2 on smartphones as highly usable. System Usability Scale (SUS) [Bro96] scores were near the top of the scale (cf. Figure 4a) and did not differ significantly between Group R (Mdn = 90) and Group P (Mdn = 95), $W = 835.5$, $p = .351$, $r = -.10$. Because SUS scores were not normally distributed and each group had one outlier, we used the non-parametric Wilcoxon-Mann-Whitney test to compare central tendencies.

Acceptance scores, measured using the scale by van der Laan et al. [VHD97], were also positive but differed significantly between Group R (Mdn = 1.3) and Group P (Mdn = 1.7) according to the Wilcoxon-Mann-Whitney test, $W = 691$, $p = .031$, $r = -.23$ (cf. Figure 4b). This indicates that both methods offered high usability, but platform authentication achieved higher acceptance on smartphones.

Participants' open-ended questionnaire responses about the strengths and weaknesses of their assigned authentication method further support this pattern (cf. Table 1). Many described the removal of password-related effort as a key advantage of passwordless methods [Ur+15]. Participants in both groups described the studied methods as "intuitive", "simple", and "easy to use, even for non-technical people", while especially Group P participants emphasized speed. Most participants also commented on the methods' security benefits. This is encouraging because a new authentication method must be perceived as secure and trustworthy, not merely be secure, to succeed [Lya+20].

Adoption Barriers. About 18% of participants said they were unlikely or very unlikely to adopt FIDO2. Yet this group still reported a respectable median SUS score of 93.8, showing that good usability alone does not guarantee adoption. Instead, participants pointed to practical barriers that affect adoption of FIDO2 in everyday life (cf. Table 1).

Across both Group P and Group R, participants found smartphone-based registration and login usable, but raised concerns about account delegation, authenticator revocation, and account recovery. In particular, they wanted clear ways to regain access after losing an authenticator and to revoke stolen devices or security keys. They also noted that passwordless authentication makes account sharing more difficult.

Table 1: Perceived strengths (*left*) and weaknesses (*right*) of FIDO2, as reported by participants.

	Strength	Plat.	Roam.		Weakness	Plat.	Roam.
Both	Secure	69 %	69 %	Both	Revocation/Recovery	11 %	26 %
	No password memorization issues	47 %	52 %		Complicated for lay users	13 %	5 %
	Easy to use/Intuitive	38 %	33 %		Privacy concerns	4 %	10 %
	No passwords	16 %	14 %		Website compatibility	4 %	2 %
	Easier than passwords	16 %	12 %		Coerced authentication	0 %	5 %
	No password creation	7 %	5 %		Account sharing	4 %	0 %
	No password updates	4 %	2 %		Unfamiliarity	4 %	0 %
Platform	Fast	47 %	12 %	Platform	Use on multiple clients	27 %	5 %
	Always available	11 %	2 %		Technical problems	20 %	5 %
	Privacy	11 %	2 %		Technology mistrust	13 %	0 %
	Easy to setup	4 %	2 %		Empty battery	7 %	0 %
Roaming	One solution for many accounts	4 %	10 %	Roaming	Biometric security	4 %	0 %
	Good for lay users	2 %	5 %		Loss/Destruction	20 %	81 %
	Security less smartphone-reliant	0 %	2 %		Something to carry	2 %	36 %
					Theft	4 %	21 %
					Costs	0 %	12 %
					Cumbersome	2 %	7 %

Note: Numbers are in-group percentages (%) for Group P and Group R.

Availability was another common concern. Unlike passwords, which remain usable as long as they are remembered, passwordless methods require access to specific devices. Participants worried about being locked out if a smartphone battery died, a fingerprint reader failed, or a roaming authenticator such as a YubiKey was forgotten, lost, stolen, or damaged. For roaming authenticators, some also objected to buying and carrying additional hardware. For platform authentication, some questioned how easily they could access accounts from other devices.

Adoption also depended strongly on account type. Participants weighed authentication methods differently depending on account sensitivity, usage frequency, and mobility. For sensitive accounts, participants showed two opposing patterns. Some viewed stronger security as a reason to adopt passwordless authentication for important accounts, while rejecting it for less sensitive accounts where availability mattered more. Others preferred passwords for sensitive accounts because they trusted their own memory more than the technology, despite describing passwordless authentication as easy, fast, and generally secure. This tension suggests that perceived security alone does not determine adoption; trust in the mechanism also matters.

Usage frequency and mobility also shaped preferences: roaming authentication could feel too burdensome for frequently used accounts, while platform setup could seem unjustified for rarely used accounts. Overall, passwordless authentication was most attractive when its convenience and security benefits clearly outweighed setup effort, physical effort, and the risk of temporary unavailability.

Passkeys. Because the FIDO2 ecosystem evolved after our 2021 study, we briefly relate these findings to the later introduction of *passkeys*. In May 2022, Apple, Google, and

Microsoft jointly announced expanded support for a common passwordless sign-in model based on the FIDO2 standards [App22b]. Today, passkeys are supported across all major operating systems and browsers; as of May 2025, 48% of the world’s top 100 websites support passkeys [FID25]. Passkeys do not introduce a separate authentication standard, but provide a user-facing deployment model and brand name for FIDO2/WebAuthn credentials [FID24a]. These credentials can either remain device-bound on a specific authenticator, such as a hardware security key, or they can be synchronized across a user’s devices via cloud infrastructure like Apple’s iCloud Keychain [App26], Google Password Manager [Goob], and Microsoft Accounts [Mica]. Passkeys therefore maintain the distinction between platform and roaming authenticators while broadening their deployment. Smartphones, tablets, and laptops commonly serve as platform authenticators that synchronize passkeys across devices, while external authenticators such as YubiKeys remain relevant as device-bound roaming authenticators.

Synced passkeys address several weaknesses of platform authentication identified in our research, most importantly by decoupling authentication from a single smartphone. Because these credentials are available across multiple clients, users can authenticate seamlessly on secondary devices [App26]. Users can also employ their smartphone as a bridging roaming authenticator to sign in across competing operating systems, such as using an Android phone to access a Windows laptop. Furthermore, synchronization and backup mechanisms simplify account recovery after device loss, reducing access concerns that were particularly salient for roaming authenticators [Goob; Hol24].

Despite these advances, passkeys have not yet achieved universal portability. Their management still depends on the user’s chosen passkey provider, whether integrated into the operating system, browser, or a third-party credential manager [Hol25]. Thus, although the ecosystem is broader and less tied to a single vendor than early Apple-centric deployments suggested, full portability and usability still depend on cross-platform provider support [Goo26; Hol24].

Early Reception. Our study [A] has shown early community recognition and scholarly uptake (cf. Table 2). Several usability challenges identified in the study, particularly multi-device access and recovery after device loss, were subsequently addressed by the FIDO Alliance through synced passkeys, which are now widely deployed. Without claiming direct influence on standardization or vendor deployment, this alignment suggests that our study captured practically significant barriers to passwordless authentication adoption at an important transitional moment.

Open Science. To support future research, we provide artifacts that allow others to verify and reproduce our findings. In a Zenodo repository [DS1], we publish our pseudonymized dataset with 22 variables for each of the 87 participants. These variables include usability and acceptance scores, adoption likelihood for 11 account types, and nine control variables, including privacy concerns and ATI. We also publish our R Markdown source code to replicate all statistical figures, tables, and quantitative results. Finally, we provide the source code of our mock website [SW1] to facilitate future work.

Table 2: Early reception and impact of paper [A].

Reception Type	Evidence
Community Recognition	CHI 2023 Best Paper Award (top 1%).
Scholarly Uptake	Cited across several strands of subsequent research: • Follow-up studies on FIDO2 usability: [Ame+25; Hol+25; RGP26; Riv+26] • Empirical analyses of FIDO2 deployments: [Daf+25; Kuc+23; MAI25; WAR25] • Analyses of synced passkeys: [BG25; Isl25; Isl+25; Lee+25] • Literature reviews on FIDO2 & passkeys: [APX24; Ble+25; Ble25; Mat+25] • Broader authentication practice: [Amf24; Kli+24; Ort26; Ort+25]
Practical Relevance	Findings on multi-device access, recovery, and authenticator trade-offs align with later FIDO guidance and deployments of synced passkeys by major platforms (e. g., Apple, Google, and Microsoft).

2.1.2 Custom FIDO2 Authentication via Extensions

Driven by the widespread adoption of passkeys, FIDO2 has become the leading standard for strong public-key authentication on the Internet. Its design imposes few restrictions on authenticators, but emerging use cases may require the standard to adapt, such as by exchanging additional data between authenticators and relying parties.

To support such adaptation, the FIDO2 standards define an extension interface. Extensible web standards, such as X.509 and Transport Layer Security (TLS), have historically helped advance the Internet by allowing custom protocol behavior without changing the core specification. Because standards organizations such as the Internet Engineering Task Force (IETF) and the World Wide Web Consortium (W3C) cannot anticipate every future use case, extensions provide flexibility while reducing fragmentation risk.

Potential of FIDO2 Extensions. FIDO2 extensions can add new features or mitigate shortcomings of the standard. For example, while synced passkeys have improved consumer account recovery, hardware-bound credentials in roaming authenticators still lack efficient recovery options [CPD19; Far+20; Lya+20; A]. Yubico attempted to address this issue through an extension for automatic backup credential registration [LN23]. Similarly, extensions such as hmac-secret enable new use cases, such as deriving encryption keys for password managers directly from a FIDO2 authenticator.

Beyond standard-track features, extensions are valuable for prototyping and enterprise tailoring. Developers and researchers can use them to demonstrate novel authentication designs within the existing FIDO2 ecosystem and on real-world browsers, while companies can adapt the protocol for internal deployments.

Deployment Bottleneck of FIDO2 Extensions. Despite this potential, current infrastructure severely limits the practical extensibility of FIDO2. Prior research reported difficulties extending FIDO2, particularly due to unclear browser support and implementation paths [Kli+20; Yeo+23]. As gatekeepers through the WebAuthn API, web browsers

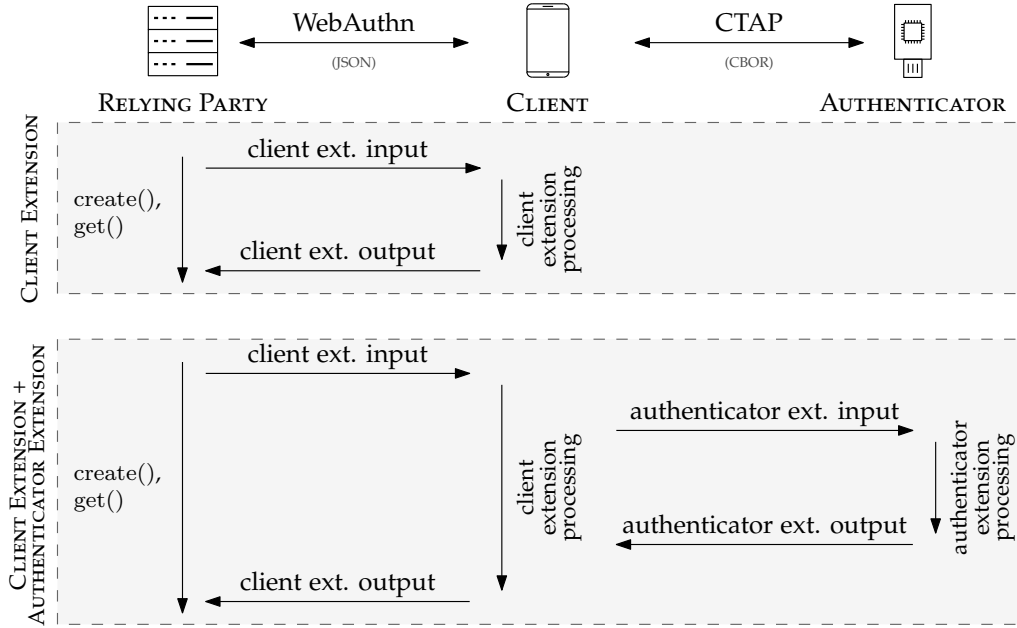


Figure 5: Protocol flow for FIDO2 extensions. Extensions are always processed by the client (*top*) and may additionally be processed by the authenticator (*bottom*).

currently do not support third-party FIDO2 extensions. Even passive extensions that require no browser-side processing and merely need to be forwarded to the authenticator are actively blocked. To the best of our knowledge, this has prevented prior research from practically extending FIDO2 through the mechanisms defined by the standards. This creates a gap between the architectural role of extensions and their practical deployability for third-party experimentation and prototyping.

This gap motivates a systematic analysis of FIDO2 extensions. Such an analysis should document the design space of existing extensions, clarify implementation paths across browsers, libraries, and authenticators, and identify changes that would make custom extensions easier to prototype and deploy without weakening security or interoperability.

This thesis has contributed significantly to the systematic analysis of the FIDO2 extension ecosystem through the following publication, which can be found in Appendix B:

- [B] F. PUTZ, S. SCHÖN, M. HOLLICK. “Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions.” In: *Emerging Technologies for Authorization and Authentication*. ETAA ’21. Cham, Switzerland: Springer International Publishing, 2021, pp. 17–32. ISBN: 978-3-030-93746-1. DOI: [10.1007/978-3-030-93747-8_2](https://doi.org/10.1007/978-3-030-93747-8_2). Appendix B.

Regrettably, public information on custom FIDO2 deployments is scarce, so we for the first time document the process of designing and implementing FIDO2 extensions in web browsers, software libraries, and hardware tokens. We also survey published FIDO2 extensions by manually inspecting the source code of major browsers and authenticators, and propose browser modifications to facilitate custom extension deployment.

Existing Extensions. Our survey provides a snapshot of the FIDO2 extension landscape as of 2021 (cf. Table 3). Most known extensions were defined in the FIDO2 standards themselves, i. e., in WebAuthn [Wor19; Wor21] and in the Client to Authenticator Protocol (CTAP) [FID21]. These extensions add functionality such as legacy Universal 2nd Factor (U2F) compatibility, credential protection policies, application data storage, and symmetric-key derivation via `hmac-secret` for use cases such as disk encryption or password managers [FID21].

Beyond the standard extensions, we identified three notable non-standard extensions. Yubico’s recovery extension proposes registering backup authenticators to improve account recovery [Fry+20; LN23]. Google’s `caBLE` mechanism enables a smartphone to act as a roaming authenticator via BLE, although it is implemented as a browser-specific transport mechanism rather than as a normal FIDO2 extension with a public specification [Moo20]. Finally, Google’s `googleLegacyAppIdSupport` extension provides limited backward compatibility with older Google-specific U2F deployments [Kre21]. Overall, our survey shows that extensions were already used both to evolve the standard and to address deployment-specific requirements.

Ecosystem Compatibility. To determine the practical deployability of FIDO2 extensions, we analyzed support in three layers of the ecosystem: web browsers, client libraries, and authenticators (cf. Table 3). Unfortunately, there is no generic interface that reveals supported extensions, as the WebAuthn standard permits clients to silently ignore unknown extensions. Therefore, we determined support by manually inspecting the source code of major web browsers, clients, and authenticators.

Web browsers are the primary gatekeepers for FIDO2 traffic. We inspected the major browser engines available in 2021: Chromium 93 (Chrome/Edge), Gecko 91 (Firefox), and WebKit 611 (Safari). The results showed highly uneven support. Chromium implemented by far the broadest extension support, including several WebAuthn and CTAP extensions as well as Google’s non-standard `caBLE` mechanism. Firefox and Safari supported only a small subset, mainly extensions required for compatibility with legacy U2F credentials. Thus, extension support was already fragmented and incomplete at the browser layer.

For developers bypassing the browser constraint, client libraries offer more flexibility. Yubico’s `python-fido2` [Yubb] library provided the most robust support for custom integrations, handling six different extensions. Similarly, Yubico’s C library, `libfido2` [Yuba], supported four extensions and provided helpful tooling for prototyping.

We analyzed hardware tokens by inspecting open-source firmware and querying devices via the `authenticatorGetInfo` CTAP2 command. Roaming authenticators, including the YubiKey 5 series [Yub20], the C-based SoloKey [Sol18], and Google’s Rust-based OpenSK [Gooa], predominantly supported only the `credProtect` and `hmac-secret` extensions. On the platform side, Microsoft’s Windows Hello platform authenticator supported three FIDO2 extensions via its Win32 API [Micb].

In summary, the 2021 ecosystem showed a clear mismatch between the architectural role of extensions in FIDO2 and their practical deployability. A growing set of standardized and vendor-specific extensions had been proposed, yet support across browsers, libraries, and authenticators remained sparse and inconsistent.

Table 3: Compatibility of known FIDO2 extensions as of July 2021. The table provides a snapshot of their specification status and implementation support.

Identifier	Chrome	Edge	Firefox	Safari	libfido2	python-fido2	Windows Hello	YubiKey 5	SoloKey	OpenSK	Specification	Reference	IANA Registration?
appid	✓	✓	✓	✓	✗	✓	-	-	-	-	WebAuthn2 §10.1	[Wor21]	✓
appidExclude	✓	✓	✗	✗	✗	✗	-	-	-	-	WebAuthn2 §10.2	[Wor21]	✗
authnSel	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.4	[Wor19]	✓
biometricPerfBounds	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.9	[Wor19]	✗
credBlob	✓	✓	✗	✗	✓	✓	✓	✗	✗	✗	CTAP2 §12.2	[FID21]	✗
credProps	✓	✓	✗	✗	✗	✗	-	-	-	-	WebAuthn2 §10.4	[Wor21]	✗
credProtect	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	CTAP2 §12.1	[FID21]	✗
exts	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.5	[Wor19]	✓
hmac-secret	✓	✓	✓	✗	✓	✓	✗	✓	✓	✓	CTAP2 §12.5	[FID21]	✗
largeBlob	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.5	[Wor19]	✗
largeBlobKey	✓	✓	✗	✗	✓	✓	✗	✗	✗	✗	CTAP2 §12.3	[FID21]	✗
loc	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.7	[Wor19]	✓
minPinLength	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	CTAP2 §12.4	[FID21]	✗
txAuthSimple	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.2	[Wor19]	✓
txAuthGeneric	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.3	[Wor19]	✓
uvi	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn1 §10.6	[Wor19]	✓
uvm	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	WebAuthn2 §10.3	[Wor21]	✓
caBLE	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗		[Moo20]	✗
googleLegacyAppIdSupport	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗		[Kre21]	✗
recovery	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗		[LN23]	✗

Designing and Implementing Extensions. To illustrate the practical extensibility of FIDO2, we designed and implemented a proof-of-concept extension across the full stack, including the relying party, client, and authenticator. Our example, the greeter extension, is intentionally simple: it takes a string input and returns a greeting string, while exercising extension inputs and outputs in both registration and authentication.

We implemented greeter in a prototype web application, in Chromium, in Yubico’s python-fido2 client library [Yubb], and in the open-source authenticators Solo 1 [Sol18] and OpenSK [Gooa]. This end-to-end prototype shows that custom FIDO2 extensions are feasible in practice, but harder to realize than the protocol design suggests. In particular, Chromium does not expose extension handling through a modular interface, and the required implementation steps are largely undocumented. We therefore derived them manually from existing extensions. This required extending the WebIDL and Mojo definitions for extension inputs and outputs, adapting conversions between JavaScript-facing and internal representations, adding the extension to Chromium’s WebAuthn

request processing, and including it in CTAP request construction. The response path required analogous changes to propagate authenticator output back to JavaScript.

In contrast, `python-fido2` provides a more modular basis for custom extensions. On the authenticator side, extension support in Solo 1 and OpenSK is not modularized; for OpenSK, the implementation approach was again undocumented and had to be inferred from source code. Overall, our implementation experience shows that the main obstacles lie not in the protocol itself, but rather in the surrounding software ecosystem.

Extension Pass-Through. WebAuthn extensions are optional: clients and authenticators may ignore them, and relying parties must therefore tolerate any subset of requested extensions being unsupported. To improve compatibility with new or unknown extensions, the WebAuthn specification allows clients to pass extension inputs directly to the authenticator and return extension outputs to the relying party without additional client-side processing, using defined conversion rules between JSON and CBOR [Wor21].

In practice, however, our inspection showed that no major browser supported extension pass-through. This is a substantial barrier for deploying and evaluating custom FIDO2 extensions on the web, because unrecognized extensions are dropped even when the authenticator could process them. According to our survey, only Google as a browser vendor was able to implement their own custom extensions inside their web browser. As of 2026, to the best of our knowledge, custom third-party extensions are still not deployable through standard web browsers. As a result, the extensibility envisioned by WebAuthn remains only weakly realized in current browser implementations.

A legitimate concern is that custom extensions could expose fine-grained authenticator properties and thereby increase the risk of fingerprinting or discrimination between authenticators. Chromium’s refusal to implement the standard `uvi` and `uvm` extensions illustrates this concern [Lan20]. Even so, a more balanced approach would be to support pass-through under user control, for example through an explicit browser setting or permission prompt. This would make it easier to experiment with new FIDO2 functionality while preserving user choice and privacy.

Outlook. Custom FIDO2 extensions currently remain most realistic in controlled environments, such as internal deployments where browser configurations and authentication policies can be managed centrally. For broader third-party experimentation, browsers would need to support extension pass-through or expose custom FIDO2 functionality through existing browser extension mechanisms. Client libraries and authenticators would likewise benefit from more modular, better-documented extension support. Improving this infrastructure would help researchers and developers explore new use cases without weakening the security and interoperability guarantees of the FIDO2 standard.

Early Reception. Our work [B] has received early scholarly uptake in studies on FIDO2 deployment, surveys, and extensions (cf. Table 4) and identified limited browser support for custom FIDO2 extensions as a practical deployment barrier. This barrier remains visible in later work, where extensions are used to support new use cases [Bav25; Riv+25] and workarounds such as WebAuthn API hooking [Yeo+23] or out-of-band

Table 4: Early reception and impact of paper [B].

Reception Type	Evidence
Scholarly Uptake	Cited across several strands of subsequent research: • Empirical analyses of FIDO2 deployments: [Kuc23; Kuc+23; Las+24] • FIDO2 surveys and position papers: [APX24; BU22; Mat+25] • Extending and improving FIDO2: [AE22; Yeo+23]
Practical Relevance	Our findings on missing extension pass-through remain relevant: later research using FIDO2 for new authentication concepts [Dan+24; Hac+22; Yeo+23] continues to report deployment difficulties due to limited browser support.

communication [Dan+24; Hac+22] are needed. This suggests that missing extension pass-through remains a relevant constraint for implementing novel FIDO2 applications.

Open Science. We publish documentation and supplementary material on how to implement custom FIDO2 extensions on GitHub [SW2]. Using the greeter extension as an example, we describe how to implement extensions across the FIDO2 stack: on the relying party (website), on the client (Chromium web browser and Python client), and on the authenticator (Solo 1 hardware token). For the relying party, we provide a complete demo web app written in Vue.js. For the remaining components, we provide patches that add our greeter extension to the respective projects.

2.2 Related Work

This section situates our work in research on passwords, FIDO2, and related authentication approaches. We distinguish earlier work motivating our analysis from subsequent work that further clarifies the adoption challenges identified in this dissertation.

2.2.1 Towards Passwordless Web Authentication

Passwords have long been the default mechanism for web authentication, but their limitations are well known [Bon+12]. Secure passwords are difficult to memorize [Yan+04], password reuse enables large-scale compromise across services [IWS04], and password-based authentication remains highly susceptible to phishing attacks [DTH06]. Password managers mitigate some of these problems by helping users generate and store unique secrets, yet they do not remove phishing risks [OR20; Sil+14; SJ14] and may still expose credentials to malicious client-side scripts or browser extensions after autofill [Gau+25]. More broadly, prior work has shown that many proposed password replacements struggle to achieve widespread adoption because improvements in security often came at the expense of usability and deployability [Bon+12].

To strengthen password-based authentication, websites commonly add a second factor such as SMS or email verification codes (OTP), time-based one-time passwords (TOTPs),

or hardware tokens [Gil25; JSS25; QHD18; ULC19]. However, widely deployed OTP-based approaches remain vulnerable to SMS rerouting [Pee+22] and real-time phishing attacks, where attackers relay user-submitted codes to the legitimate site [Ulq+21; ULC19]. Empirical studies further show that 2FA deployments are often inconsistent and difficult to understand, during both enrollment and everyday use [Ger+23; LBB23]. Recovery and backup practices remain a recurring problem as well: many services offer insecure fallback mechanisms, and many users lack robust plans after losing access to their second factor [Höl+24; Kli+23; Tie+25].

2.2.2 FIDO2 Passwordless Authentication

The limitations outlined above motivate the move towards passwordless authentication based on public-key cryptography rather than shared secrets. Compared with passwords and common OTP-based 2FA, such approaches promise stronger phishing resistance while reducing the burden of memorizing and managing secrets. FIDO2 is the most prominent realization of this direction and aims to improve security, deployability, and usability on the web [FID24a]. The following sections therefore focus on related work on FIDO2, especially smartphone usability and remaining adoption barriers.

Roaming Authentication. Prior usability research on roaming authenticators has mostly studied external security keys, first for U2F [FID17] (FIDO2’s predecessor) and later for FIDO2 [Wor21]. Across these studies, authentication with security keys was often fast and usable after enrollment [Ree+19; Rey+18], but adoption remained constrained by surrounding practicalities. Users sometimes perceived security keys less favorably than familiar OTP-based 2FA methods [CPD19], adoption remained limited outside specialized contexts [Col+18], and acceptance did not always follow usability alone [DDC18; Das+18]. At the same time, large-scale enterprise experience demonstrated the practical viability of security keys in managed environments [Lan+17].

For FIDO2 roaming authenticators specifically, prior work reports both encouraging usability results and persistent practical barriers. Lyastani et al. found YubiKey-based authentication to be more usable and more accepted than passwords on desktop systems [Lya+20], while other studies highlight friction related to token form factor, connectivity, price, setup, and backup [CPD19; DDC18; Far+20; Rey+18]. A related line of work explored smartphones as roaming authenticators and found mixed results: users recognized the security benefits of FIDO2, but sometimes still preferred passwords [OUA20] or worried about device availability, especially battery depletion [Owe+21; Ras21]. Overall, prior work indicates that roaming authenticators can be highly usable once in place, but that their long-term acceptance is shaped less by the authentication act itself than by surrounding practicalities such as transport, setup, and backup. To the best of our knowledge, their usability on smartphone clients has not been studied before.

Platform Authentication. Platform authenticators integrate FIDO2 authentication directly into the user’s device and typically rely on local PIN or biometric verification. Compared with roaming authenticators, they avoid carrying and connecting a separate token and therefore appear well suited to smartphone-based authentication. Prior work

emphasizes their convenience and familiarity, especially when they build on established device-unlocking mechanisms such as fingerprint or face recognition [Boo22].

Empirical studies of FIDO2 platform authentication on smartphones remain limited. Prior interview-based work on smartphone platform authentication examined first-time use and identified opportunities to improve website flows and onboarding [Oog+20]. Earlier work showed that users can hold misconceptions about platform authentication and proposed client-side notifications to mitigate them [Las+21]. Research beyond smartphones points to similar themes: users often perceive biometrics as usable, but adoption depends on hardware availability, user understanding, and deployment context [Boo22; Far+22; Mec+24]. Existing studies therefore indicate strong usability potential, but they mostly examine platform authentication in isolation rather than directly comparing platform and roaming FIDO2 authentication on smartphones.

Deployment and Adoption Barriers. Beyond the authentication interaction, FIDO2 adoption is shaped by organizational, technical, and procedural factors. Studies in enterprise and commercial contexts, published after our lab study, report that adoption is often slowed by rollout complexity, missing integrations into common server stacks, regulatory requirements, and recovery concerns [Hol+25; Kep+23; Las+24].

Several works besides this dissertation identify account recovery, backup authenticators, and multi-device access as major open challenges for passwordless systems [Alq+20; Aro+22; Ble+25; Kun+21]. Proposed solutions include enrolling multiple authenticators at once [Aro+22; Fry+20], peer-assisted recovery [BSU25], and delegation mechanisms for accounts shared by multiple users [FGM22; Tay23; YHV24]. Since our study, synced passkeys have made these lifecycle issues more visible: they improve usability and cross-device availability, but also introduce new dependencies on credential managers as trusted third parties [Ble+25; BG25]. Related work therefore explores cross-platform synchronization [Set+24] or alternatives to cloud-based key syncing [BSU25; SS25], while security analyses highlight remaining risks such as compromised passkeys [Isl+25], compromised clients [Kuc+23], malicious browser extensions or scripts [YS24], and interpersonal threat models involving routine physical device access [Daf+25].

Finally, recent work suggests that adoption depends strongly on communication and interaction design. Improving users' technical understanding of FIDO2 does not necessarily increase adoption intentions; risk communication and self-efficacy may matter more [Ame+25]. Field evidence further shows that well-timed enrollment prompts can significantly increase passkey adoption [RGP26]. Overall, the main barriers to widespread FIDO2 adoption no longer lie primarily in the authentication interaction itself, but in deployment complexity, user interaction design, and unresolved lifecycle workflows such as recovery, sharing, and multi-device use [Ble+25; Hol+25; Las+24; Mat+25].

Summary. Related work has studied passwords, 2FA, and the usability of FIDO2 authenticators, but mostly focuses on roaming authenticators, platform authenticators in isolation, or deployment challenges in specific settings. Little work directly compares both FIDO2 authentication options on smartphones in a controlled user study. Our work addresses this gap by comparing both options on smartphones and by identifying adoption barriers beyond the authentication interaction itself. Subsequent research further

reinforces the relevance of these broader lifecycle and deployment issues, especially recovery, synchronization, delegation, and organizational adoption.

2.2.3 FIDO2 Extensibility

Extension mechanisms allow security protocols to evolve without fragmenting their ecosystems, as illustrated by X.509 certificate extensions [Boe+08] and TLS handshake extensions [Eas11]. FIDO2 follows this approach. Its two protocol layers, WebAuthn and CTAP2, define an extension framework for use cases beyond the core authentication flow [FID21; Wor21]. Extensions may involve only the client or both the client and authenticator, and may apply to registration, authentication, or both. For interoperability, extensions are identified by strings that can be registered in the IANA “WebAuthn Extension Identifiers” registry [IAN], established by RFC 8809 [HMJ20].

The FIDO2 standards already include several extensions, and CTAP 2.1 mandates some under specific conditions, including hmac-secret and credProtect [FID21]. Although clients can query these via the `authenticatorGetInfo` command, public data on their practical deployment remains scarce compared to mature ecosystems like X.509 and TLS. These gaps motivate our survey of existing FIDO2 extensions and their support across browsers, client libraries, and authenticators.

Academic Use of FIDO2 Extensions. Besides our work, Klieme et al. documented their experience implementing a custom FIDO2 extension for continuous authentication [Kli+20]. They criticized the lack of supporting material and reference implementations, the error-prone nature of the implementation process, and the difficulty of determining browser support for both existing and custom extensions. Because of these challenges, they implemented their extension only on the relying party and authenticator, but not in the browser. Our work complements theirs by documenting implementation across the full FIDO2 stack, including the browser, and by providing the first survey of extension support across browsers, software libraries, and authenticators (cf. Table 3).

Since our 2021 analysis of FIDO2 extensions, several studies have used extensions to address gaps in the standard and enable new use cases. Hackenjos et al. implemented a malware-resistant web authentication scheme based on FIDO2 using the `txAuthSimple` extension [Hac+22]. Yeoh et al. integrated anonymous credentials and eID solutions via a custom extension [Yeo+23]. Dang et al. designed a custom extension for anonymous authentication [Dan+24]. Recent work has also used the `hmac-secret` extension to derive symmetric secrets for end-to-end encrypted cloud services [Bav25] and EAP-based Wi-Fi authentication [Riv+25]. Together, these studies show that extensions can adapt FIDO2 to new application domains, although browser support still limits practical deployment.

Foundations for Decentralized Authentication

Modern smartphones rely heavily on centralized network connectivity. Unfortunately, power outages, natural disasters, and similar crises can damage network infrastructure, disconnecting buildings, neighborhoods, or even entire cities from the global Internet. This disconnection severely disrupts how mobile devices and apps operate: services that seem routine in everyday life suddenly fail when users need them most.

A major source of this fragility is the centralized design of many smartphone services. While static content can be replicated to local edge servers, many apps and cellular networks still depend on remote backend services for authentication, coordination, or message routing. As a result, otherwise available local services become unusable when backend connectivity is lost, and users lose access to important services such as digital payments, social media, and cloud-hosted documents. The dependency becomes especially striking in messaging apps: even nearby users cannot communicate if their messages must be routed through remote servers. Centralized dependencies therefore constitute a major *deployability* risk.

Case Study: Natural Disasters. As a tragic example of network fragility, consider the 2021 flood disaster in central Europe [Bir+26]. Extreme flash flooding caused at least 180 deaths in Germany alone, particularly in the Ahrweiler district, and severely damaged local infrastructure [Moh+23]. Streets and bridges carrying critical fiber links were destroyed, cutting off backbone connectivity for both mobile and fixed networks, while flooded electrical substations disrupted power to distribution nodes and cell towers. Although some towers had battery backups, floodwaters destroyed many fiber connections to the core network, leaving mobile and fixed Internet access unavailable in parts of the district for weeks [Kok+22]. This outage severely hampered crisis response: residents struggled to access weather forecasts and emergency updates, inform relatives about their safety, and coordinate rescue and recovery efforts through messaging groups or social media platforms [Fek21; Thi+23]. The Ahrweiler disaster shows how strongly emergency coordination depends on applications tied to centralized infrastructure.

Case Study: Extended Blackouts. A similar pattern emerged during the January 2026 blackout in southwest Berlin, which was caused by an arson attack on critical power infrastructure. The outage affected over 45 000 households and 2 000 businesses, with full restoration taking several days [BDW26]. It also cascaded into telecommunications infrastructure: fixed Internet relied on powered customer equipment, while mobile Internet sites failed as backup power was exhausted [Bri26; BNe22]. As in Ahrweiler, the loss of connectivity cut users off from important smartphone services and critical information. This example shows that centralized infrastructure dependencies remain an important *deployability* risk even in urban environments.

Decentralized Local-Only Operation. One way to improve resilience to disruptions that isolate a region from the global Internet is to support *local-only operation* during outages. In such a mode, nearby users could continue to communicate with one another and access services hosted within the affected region, even when remote cloud services and the global Internet are unreachable. However, current cellular architectures and online service deployments lack support for such decentralized fallback modes. As a result, services that could in principle be useful within a local network island remain unavailable because their dependencies are outside the affected region.

To make communication systems *island-ready*, developers, authorities, and mobile network operators (MNOs) need a clear understanding of which services users prioritize under local-only connectivity. This prioritization is important for deployability: local compute, storage, energy, and network resources are limited, and not every application can realistically be supported with the same quality of service during an outage. User priorities can therefore guide which services should be adapted first, which dependencies must be decentralized, and where authentication support is most critical.

This chapter provides the user-centered foundation for the dissertation’s later work on decentralized authentication. We first summarize our empirical study on application priorities during network disruptions. We then use these findings to show why decentralized authentication becomes a major deployability challenge for island-ready systems, especially when users prioritize services such as messaging that depend on secure user-to-user communication.

This chapter thereby addresses the second research question:

RQ2: *Which smartphone services do users prioritize when network infrastructure fails, and what do these priorities imply for resilient authentication design?*

3.1 Our Contributions

To the best of our knowledge, user preferences during network infrastructure failures remain underexplored: the research community lacks empirical evidence on which applications citizens prioritize when a region loses global Internet access but retains local-only connectivity. We address this gap with the following publication, located in Appendix C:

- [C] L. JANZEN, F. PUTZ, M.-A. KAUFHOLD, K. STRAUB, M. HOLLICK. “The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity.” In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Apr. 26–May 1, 2025). CHI ’25. Yokohama, Japan: ACM, Apr. 2025. doi: [10.1145/3706598.3714324](https://doi.org/10.1145/3706598.3714324). CORE Rank A*. Honorable Mention Award. Appendix C.

Specifically, we present findings from a survey of 857 adult smartphone users across major German cities. The study examines how participants prioritize smartphone application categories in everyday use and under a hypothetical crisis scenario with local-only

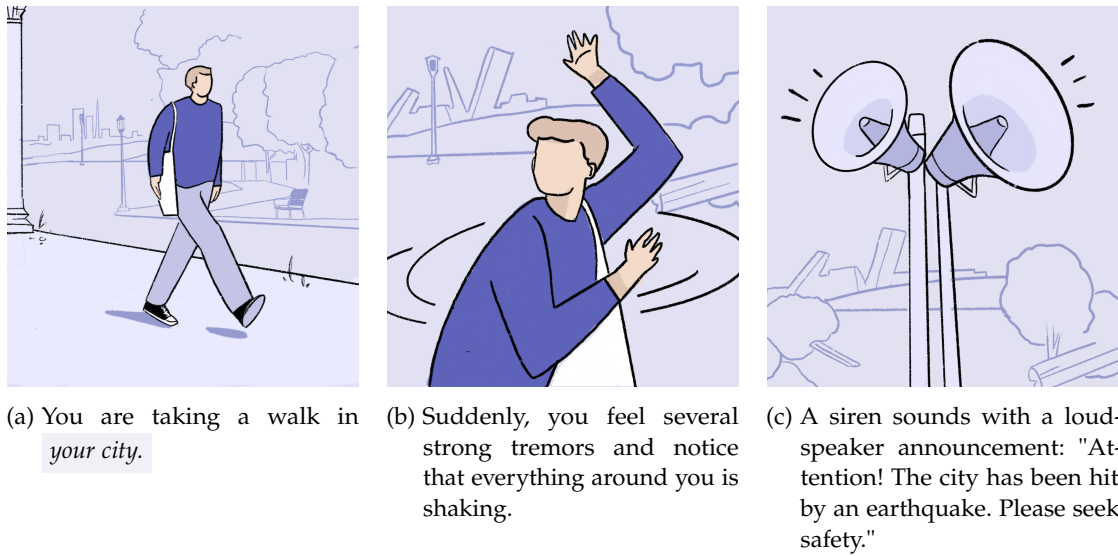


Figure 6: Comic board used to introduce participants to the study's assumed urban crisis scenario. The first panel was individualized by replacing *your city* with the participant's city.

connectivity. Our results indicate a shift in app demand, with users frequently favoring general-purpose apps over dedicated crisis apps. Notably, citizens rank messenger applications highest across both normal and local-only operations, demonstrating a sustained need for interpersonal communication and digital coordination even when communication is limited to the affected region.

These findings provide a user-centric basis for prioritizing services in island-ready cellular systems. They also highlight why messaging is a particularly important case for local-only operation: if communication and collaboration remain central in crisis scenarios, then resilient systems must support secure user-to-user authentication without relying on globally reachable identity infrastructure.

Study Design. To gather representative user preferences, we conducted an IRB-approved online survey ($N = 857$) targeting adult smartphone users in major German cities. Following a pilot study ($N = 29$), we recruited participants through an ISO-certified panel provider, which compensated each participant. The final sample represents smartphone users from all 16 German federal states and 79 of the 82 major cities with more than 100 000 residents at that time.

Our survey included nine demographic questions and four control variables, including the affinity for technology interaction (ATI) scale [FAW19] and prior experience with crises or Internet outages. To explain the assumed urban crisis scenario and local-only connectivity, we used comicboarding [Mor+07], a visual method for breaking down complex information. We chose this method because the concept of isolated local networks might otherwise be difficult for non-experts to understand. We analyzed the quantitative data in R 4.4.1 using an alpha level of 0.05 for statistical significance.

App Demand under Local-Only Connectivity. During the survey, participants ranked 23 application categories according to their daily usage frequency using positional voting.

Table 5: App demand under normal and local-only operation, determined by the Borda count. Dotted arrows indicate rank changes ≥ 2 ; bold blue arrows indicate **strong trends**.

(a) Normal Operation		(b) Local-Only Operation	
App Category		App Category	
1	Messengers	1	Messengers
2	Social Media	2	News
3	Web Browsers	3	Crisis
4	Weather	4	Social Media
5	News	5	Weather
6	Shopping	6	Web Browsers
7	Navigation	7	Navigation
8	Entertainment	8	Entertainment
9	Music	9	Music
10	Finance	10	Utilities
11	Utilities	11	Reference
12	Reference	12	Finance
13	Productivity	13	Food & Drink
14	Games	14	Games
15	Food & Drink	15	Productivity
16	Travel	16	Books
17	Health & Fitness	17	Health & Fitness
18	Education	18	Medical
19	Books	19	Shopping
20	Sports	20	Education
21	Crisis	21	Travel
22	Medical	22	Creativity
23	Creativity	23	Sports

We then introduced a hypothetical crisis scenario with local-only connectivity: an earthquake that disconnects the city from the global Internet (cf. Figure 6). A supplementary short video explained that communication would be limited to users within the city and that digital content would only be available from locally hosted servers. Participants then repeated the ranking for this scenario, assuming the crisis had lasted several days.

Table 5 shows the resulting shifts in app demand. Interestingly, messenger apps ranked highest in both normal and local-only operation, indicating a persistent need for communication even when users can only reach others within the same city. Under local-only connectivity, news and crisis applications became the next most important categories and gained substantial importance compared to normal operation. In contrast, users assigned lower priority to social media and web browsers, and no longer considered shopping and travel applications central priorities.

App Demand for Crisis-Specific Use Cases. Local-only connectivity often arises during crises such as flash floods, earthquakes, or other large-scale disruptions, where access to remote infrastructure may be impaired. Based on related work [HKR22; Reu+23; Reu+17; Sim+23; WH23], we selected seven crisis-specific use cases and asked participants which apps they would use for each: receiving warnings, getting safety tips, contacting emergency services, looking up information, communicating with others, navigating the

Table 6: App demand for crisis-specific use cases. The heat map shows the seven use cases and the ten most frequently selected app categories. Percentages are shown for the three most selected categories per use case, with bold text indicating the preferred category. M_c denotes the mean mention frequency of category c across use cases; M_u denotes the mean number of categories selected for use case u .

Use case \ Category ¹	N_u	Crisis	Messengers	News	Social Media	Web Browsers	Navigation	Weather	Reference	Music	Entertainment	M_u	None
Warnings (U1)	380	65 %	27 %	42 %								2.12	5 %
Safety tips (U2)	356	61 %		39 %		27 %						2.28	5 %
Emergency call (U3)	375	48 %	37 %			13 %						1.56	16 %
Information (U4)	348	53 %	32 %	60 %								2.42	2 %
Communication (U5)	364	22 %	72 %		35 %							1.88	6 %
Navigation (U6)	388	20 %	14 %				66 %					1.64	3 %
Distraction (U7)	360		37 %		27 %					27 %		2.42	9 %
M_c	367	41 %	35 %	29 %	19 %	15 %	14 %	9 %	7 %	4 %	3 %		

¹ The following app categories did not make the top 10 regarding M_c and were omitted in the heat map for better readability (in descending order according to M_c): Utilities, Games, Education, Books, Health & Fitness, Medical, Travel, Finance, Food & Drink, Creativity, Productivity, Shopping, Sports.

city, and finding distraction. To reduce survey fatigue, each participant was randomly assigned three of these use cases, resulting in 348 to 388 responses per use case.

Table 6 summarizes the results. Crisis applications were selected most often overall and ranked highest for receiving warnings, obtaining safety tips, and contacting emergency services. For other use cases, participants preferred general-purpose alternatives: news applications for finding information, messenger applications for communication and distraction, and navigation applications for traversing the city. These findings show that crisis support cannot rely on dedicated crisis apps alone; users also depend strongly on familiar general-purpose applications.

Messenger Apps. Under local-only connectivity, users overwhelmingly prefer messenger, news, and crisis applications. News and crisis tools are comparatively simpler to support in this setting because they are often read-only and can operate without user-to-service authentication, provided the region retains local news sources (e. g., city newsrooms or local broadcast stations) and local distribution infrastructure (e. g., local servers and content delivery networks (CDNs)).

Conversely, supporting messenger applications in a decentralized manner presents significant technical challenges. Mainstream messengers, including WhatsApp, Signal, and Telegram, rely on centralized communication architectures. Nevertheless, decentralizing these services is highly desirable because users rank messaging as the most important application category during regional outages. In a crisis, communication is urgent: affected individuals seek and share situational updates, coordinate informal response efforts, and

Table 7: Early reception and impact of paper [C].

Reception Type	Evidence
Community Recognition	CHI 2025 Honorable Mention Award (top 5%).
Scholarly Uptake	Cited in early work on local resilient 6G networks [Cho25; JH26].

reassure family members [PL07; Pal+09]. The main challenges are local message delivery, routing, and push mechanisms, but especially decentralized local authentication.

Decentralized Local Authentication. One specific challenge of local-only operation is preserving security and privacy without centralized trust infrastructure and remote identity providers. For messenger applications, this requires confidential, end-to-end encrypted (E2EE) communication and decentralized authentication.

Decentralized networks require authentication across two primary trust relationships: user-to-service and user-to-user. For user-to-service authentication, we explored usable protocols in Chapter 2, which in principle also support local-only operation. However, fully decentralized authentication requires distributing the service-side identity state.

For user-to-user authentication, most E2EE messaging apps already offer local trust-establishment mechanisms, such as QR codes or manual key-fingerprint comparisons. While these mechanisms theoretically allow users to establish ad hoc trust during a crisis, they are highly unusable in practice and fail to scale to group scenarios where multiple users must authenticate one another [FK23; D]. Unfortunately, the state-of-the-art proposal for group key exchange, SafeSlinger, does not support purely local use without external infrastructure [Far+13]. To resolve this limitation, we present a decentralized messenger authentication protocol in Chapter 4, designed to enable local trust establishment while addressing usability, scalability, and deployability concerns.

Decentralized Operation. Although the vision of local-only connectivity aligns well with the decentralized web [Ram+19] and local-first software [Kle+19], current cellular networks and mainstream smartphone apps are not yet island-ready. Putting local-only operation into practice involves multiple stakeholders: MNOs must enable seamless network transitions into local-only operation; service providers must adapt their backends and application logic to function during network partitions; and authorities need to define clear requirements for traffic prioritization and fairness. Finally, because users expect to rely on general-purpose applications during crises, developers must design these services to transition smoothly between global and local-only connectivity.

Early Reception. Our study [C], published in 2025, has received early community recognition and initial scholarly uptake in emerging work on resilient and local 6G networks (cf. Table 7). Subsequent work has already begun to address one technological challenge raised in the study: enabling island-ready cellular connectivity by decentralizing the 6G core network. Our work thus provides an early user-centered reference point for ongoing technical developments towards resilient 6G communication infrastructures.

Open Science. To support future research, we provide comprehensive artifacts for verifying, reproducing, and replicating our findings: a dedicated Zenodo repository [DS2] contains the complete dataset with quantitative and qualitative responses from all 857 participants. This repository also includes the Python and R Markdown evaluation scripts to reproduce our results. Additionally, we have released the original LimeSurvey structure to facilitate independent replication of our study.

3.2 Related Work

We use *local-only connectivity* to describe situations in which a geographic region (e. g., a city) remains covered by a public cellular access network but becomes temporarily disconnected from the global Internet [JH26]. In such scenarios, local communication and service access can remain feasible if the required application backends and authentication services are available within the affected region. However, many smartphone apps depend on infrastructure outside the local area, so local wireless coverage alone does not guarantee that users can continue to access the services they need.

We therefore review related work that is relevant to the *deployability* and *usability* of local-only operation. First, we discuss technical dependencies that can prevent smartphone services from operating inside a network island. We then review crisis communication systems and prior work on digital media use in crises. Together, these areas show that local-only operation is technically plausible, but previous research provides limited evidence on *how citizens prioritize* smartphone app categories under local-only connectivity, and thus on which user needs become most urgent in crises and other infrastructure disruptions. Our study contributes this user-centered understanding as a basis for reasoning about island-ready cellular networks and their authentication requirements.

3.2.1 Local-Only Connectivity as a Deployability Challenge

While the Internet’s routing and naming mechanisms were designed to scale across heterogeneous administrative domains, modern application ecosystems increasingly rely on *centralized service dependencies*. For smartphone apps, these commonly include:

- global naming and endpoint resolution (e. g., DNS),
- remote application backends (e. g., cloud APIs and data stores),
- online identity and authorization (e. g., OAuth, SSO, and token refresh), and
- centralized messaging infrastructures (e. g., push notifications).

When these dependencies become unreachable, an app may fail even if the user’s device, nearby peers, and local wireless networks remain available. These centralized dependencies are therefore a *deployability* risk for local-only connectivity.

Prior work offers several building blocks for reducing such dependencies. Multi-access edge computing (MEC) and CDNs move selected content and computation closer to

users, which can allow replicated resources to remain available during upstream outages [Zol+20]. Local-first and offline-first systems similarly reduce dependence on continuously reachable backends by allowing applications to operate on local state and synchronize later [KS92; Kle+19]. Peer-to-peer and decentralized web systems distribute storage, naming, or message routing across participating nodes and can, in principle, support communication within an isolated region [Doa+22; Kei+24]. However, these approaches often still depend on bootstrapping, rendezvous, or identity infrastructure, and they face practical smartphone constraints such as background execution limits and hard-to-maintain peer connectivity [Bal+23; Pet+19].

Cellular Infrastructure. Local-only operation also depends on the cellular network itself. Current architectures route many user flows through centralized core network functions before they reach the public Internet [3GP23], which can create path inflation and resilience bottlenecks for geographically local communication [Zar+14]. Prior work on distributed cellular cores and geo-redundant network functions reduces these bottlenecks by moving selected core functionality closer to users [Cor+24; MRR18]. Such architectures make it more plausible that an isolated region could retain basic mobile connectivity and host selected local application backends, complementing edge-computing approaches.

Authentication. Even if local routing, computation, and storage are available, many services still require users or devices to authenticate. Work on decentralized identity, self-sovereign identity, and web-of-trust mechanisms aims to reduce dependence on central identity providers, but key management, user-to-user authentication, and usability remain difficult on mobile devices [Bru+21; TCL23; Ung+15]. This challenge is especially relevant for messaging and coordination services, where users need to establish trust without access to globally reachable identity infrastructure.

Overall, prior work shows that local autonomy is technically plausible, but mainly addresses how services and networks *could* operate locally. It provides less guidance on which smartphone services to prioritize when local compute, storage, energy, and network resources are limited. Our study addresses this user-centered gap by examining which app categories citizens consider most important under local-only connectivity.

3.2.2 Crisis Communication under Degraded Connectivity

Crisis communication research has long recognized that failures of centralized information and communication infrastructure can impair warnings, coordination, and situational awareness [CH06]. Prior work therefore considers resilient and decentralized communication systems for crisis response, including self-organizing cellular networks, mobile ad hoc networks (MANETs), D2D communication, mesh networks, and portable “network-in-a-box” deployments [Höc+23; Poz+18; Tri+17]. These systems can restore local communication or provide temporary coverage, but they often require additional equipment, planning, physical deployment, or favorable operating conditions. This can limit their deployability for large urban populations that continue to rely on ordinary smartphones and public cellular infrastructure during a crisis.

Graceful Degradation. Other approaches aim to restore backhaul connectivity, for example through satellite or microwave links [Dee+19]. Such restoration is important, but it may be delayed, capacity-limited, or unevenly available across the affected region. Island-ready cellular networks take a complementary approach: rather than assuming immediate reconnection to the global Internet, they aim for *graceful degradation* by maintaining selected local services during the disruption.

Usability. From a usability perspective, this creates a distinct interaction problem. Users may move between full Internet connectivity, intermittent connectivity, and local-only connectivity, while applications expose different feature sets, availability states, or communication scopes. Prior work has considered how systems can remain usable across degraded or intermittent connectivity [H]. However, it remains unclear which applications citizens would actually want to use when connectivity is restricted to a local network island. Our study addresses this gap.

3.2.3 Smartphone App Priorities in Crises

The field of crisis informatics has produced a rich body of work on how citizens use social media during emergencies, including for warning reception, information sharing, coordination, and situational awareness [KR16; OVC15; Reu+23; SP11].

Crisis communication also extends beyond social media. Mobile crisis apps can support disaster-related information and communication [Tan+17b], while official warning and crisis apps provide alerts, behavioral guidance, and sometimes channels to authorities or emergency services [DM21]. Studies with German citizens further show increasing adoption and perceived usefulness of crisis apps [HKR22], and prior work has examined their usability [Tan+20], motivational factors [FBF22], and preparedness effects [HPR23]. These studies clarify which digital functions users value during crises, but they mainly study app use under the assumption that the required infrastructure remains available.

This assumption is problematic for local-only connectivity. Many crisis-relevant services depend on centralized platforms, app backends, identity providers, or push infrastructures. Events such as the 2021 Ahr Valley flooding illustrate that failures of energy and telecommunications infrastructure can severely disrupt access to digital services [Thi+23]. This shifts the question from general usefulness to service continuity under degraded infrastructure: which categories of applications should remain available when trust, backend, and network infrastructure are restricted to a local island?

Existing work provides only limited user-centric evidence on this question. Our study addresses this gap by comparing application priorities in everyday use and under a hypothetical crisis scenario with local-only connectivity.

Group Messenger Authentication

Our discussion in the previous chapter shows that a major source of fragility in digital communication lies in centralized dependencies. They create single points of trust and failure, especially in crisis scenarios or environments with degraded connectivity. This concern also applies to secure mobile messengers: although modern messengers advertise end-to-end encryption (E2EE), secure communication in practice often still depends on centrally mediated trust during contact authentication, as we illustrate below.

E2EE allows users to communicate securely without intermediary network nodes or service providers accessing the plaintext. Consequently, it provides an important mechanism for protecting the confidentiality and integrity of text messages, audio calls, and video calls in one-to-one and group settings. Since WhatsApp adopted the Signal protocol in 2016, E2EE has become a mainstream security feature. Today, most mobile messengers offer some form of E2EE, including WhatsApp, Telegram, iMessage, and Signal.

However, most users are unaware that E2EE alone cannot protect against compromised messaging providers or other active adversaries [Her+21]. In current mainstream messengers, users typically rely on their provider to distribute their contacts' public keys. Because these systems follow a trust-on-first-use model, users commonly accept these keys without independently authenticating them. This introduces a centralized trust dependency: a malicious or compromised provider can substitute keys to impersonate a contact in an active machine-in-the-middle (MitM) attack [AS23]. As illustrated in Figure 7, Alice cannot reliably detect such a substitution unless she performs an explicit *authentication ceremony* with Bob [Ung+15].

Active MitM attacks pose a practical threat to the security and privacy of users. Reports indicate that law enforcement agencies and other actors actively seek to compromise E2EE app providers to intercept supposedly secure communications [Goo18; LR18; Vas+20]. The consequences of attackers reading and manipulating private chats can be severe [AS23; Ung+15]. Exposed personal information can facilitate blackmail or reputational damage. Journalists and activists risk compromising their confidential sources. In corporate environments, leaked information or altered messages cause operational disruption and financial loss. Furthermore, such impersonation attacks can serve as an entry point for broader cyberattacks.

To mitigate the risk of active MitM attacks, messenger apps allow users to verify their contacts' cryptographic keys through an authentication ceremony [Ung+15]. In theory, such ceremonies provide genuine end-to-end authentication. In practice, however, current authentication ceremonies are highly unusable [FK23; Vaz+17]. They require considerable interaction, time, and effort, and many users are not even aware that these ceremonies

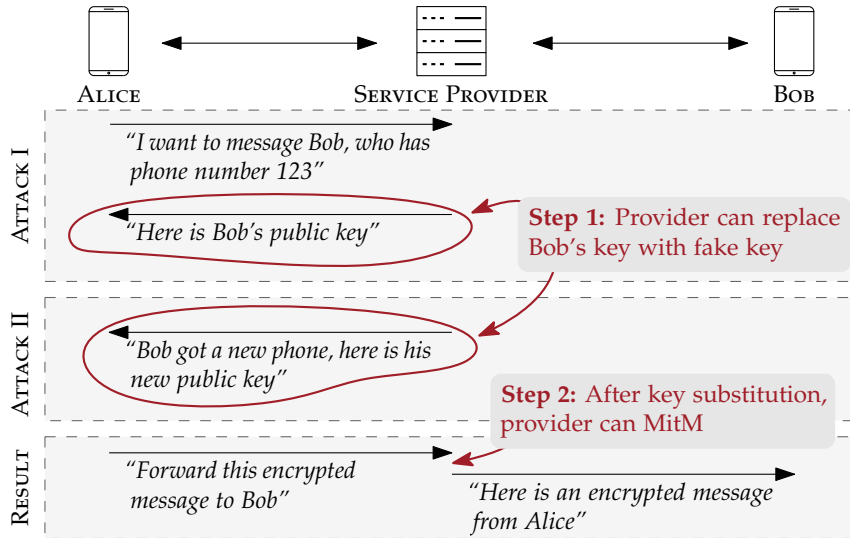


Figure 7: Two variants of active MitM attack against E2EE messaging. By substituting Bob’s public key with an attacker-controlled key using either variant, a malicious or compromised provider can impersonate Bob unless Alice independently authenticates the key.

exist or how to initiate them [FK23; Her+21]. Consequently, users rarely perform them, leaving their communications vulnerable [AS23].

This usability problem becomes even more severe in group settings. Current authentication ceremonies are designed for pairs of users and fail to scale to larger groups. Participants must perform bilateral ceremonies between every pair of group members, causing the number of manual verifications to grow quadratically with group size. For families, teams, activist groups, and other organizations that rely on group chats, this makes secure group authentication impractical.

The resulting challenge is to authenticate exchanged key material and contact information in a way that is usable, scalable, and independent of central trust infrastructures like public-key infrastructures or jointly trusted third parties. Addressing this challenge is important not only for trust, but also for resilience: decentralized authentication remains viable in scenarios with limited or disrupted connectivity, as discussed in Chapter 3.

Farb et al. [Far+13] made an important step towards usable group authentication with SafeSlinger, a protocol for securely exchanging contact information without trusted third parties. However, SafeSlinger still requires a central server for connectivity and a quadratic number of manual comparisons to achieve its security guarantees. Thus, it does not fully address the deployability and scalability requirements of decentralized group authentication.

In this chapter, we address the third research question:

RQ3: *How can authentication ceremonies for messengers be made usable and scalable for groups while remaining deployable without centralized infrastructure?*

Specifically, we present our contributions towards making messenger authentication more usable, decentralized, and practical for both pairs and larger groups. First, we

introduce *PairSonic*, a group authentication protocol for nearby mobile devices that uses acoustic communication as an out-of-band (OOB) channel (cf. §4.1.1). Second, motivated by limitations observed during our lab study, we systematically evaluate ad hoc acoustic communication between smartphones to identify robust technical solutions for direct device-to-device authentication (cf. §4.1.2). Finally, we position our approach within the broader design space by comparing it to relevant prior work (cf. §4.2).

4.1 Our Contributions

4.1.1 PairSonic: Group Authentication via Acoustic Data Transmission

To address the authentication challenges outlined above, we designed PairSonic, an open-source smartphone application and decentralized group authentication protocol for nearby users. PairSonic enables group members to verify contact information directly, without relying on central infrastructure during authentication.

This dissertation presents PairSonic through the following two publications, provided in Appendices D and E:

- [D] F. PUTZ, S. HAESLER, M. HOLLICK. “Sounds Good? Fast and Secure Contact Exchange in Groups.” In: *Proceedings of the ACM on Human-Computer Interaction* 8. CSCW ’24, Nov. 2024. DOI: [10.1145/3686964](https://doi.org/10.1145/3686964). CORE Rank A. Best Paper Award. Appendix D.
- [E] F. PUTZ, S. HAESLER, T. VÖKL, M. GEHRING, N. ROLLSHAUSEN, M. HOLLICK. “PairSonic: Helping Groups Securely Exchange Contact Information.” In: *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work & Social Computing* (Nov. 9–13, 2024). CSCW ’24. San José, Costa Rica: ACM, Nov. 2024, pp. 69–71. DOI: [10.1145/3678884.3681818](https://doi.org/10.1145/3678884.3681818). Appendix E.

To evaluate PairSonic, we conducted a user study with 45 participants comparing its usability against SafeSlinger, the current state of the art. Participants significantly preferred PairSonic over SafeSlinger. The following sections summarize the protocol design, implementation, and empirical findings.

Requirements. The primary objective of PairSonic is to enable groups of two or more nearby users to spontaneously verify their contact information, as depicted in Figure 8. Designing an effective group authentication protocol for this scenario requires clear definitions of usability, deployability, and security requirements (cf. Figure 1).

From a *usability* perspective, the protocol must require as little user interaction as possible so that it remains accessible to non-expert users. It must also scale well to larger groups and remain robust against user errors, particularly those caused by rushing users who might accept security prompts without proper verification [Kum+09].

From a *deployability* perspective, the system must operate in an ad hoc and decentralized manner, without depending on existing network or trust infrastructure. Specifically, it must function independently of an active Internet connection and run on standard smartphones without requiring new hardware or invasive system modifications.



Figure 8: Group contact verification using our acoustic group pairing protocol. Participants use their smartphones to authenticate each another in a shared authentication ceremony.

From a *security* perspective, the protocol must match SafeSlinger’s security properties [Far+13] by ensuring mutual authentication of the exchanged contact information. Upon successful protocol execution, each participant obtains the authentic contact details of all other group members, including their cryptographic public keys.

Design. PairSonic builds on the cryptographic foundation of SafeSlinger [Far+13] while improving both usability and deployability. As illustrated in Figure 9, the protocol operates in three phases:

First, the users designate one member as a temporary *coordinator*; all others act as *participants*. The coordinator’s device creates an ad hoc Wi-Fi network and broadcasts the connection credentials over an acoustic OOB channel, allowing the remaining devices to connect automatically.

Second, once all devices are connected, they perform a cryptographic exchange derived from SafeSlinger [Far+13] over the ad hoc Wi-Fi network. Participants send nested commitments containing their public keys to the coordinator, who aggregates them into a single hash value and broadcasts it acoustically. After verifying the hash successfully, the smartphones display a blue lock symbol. Users visually confirm that the symbol matches on all screens to complete the authentication ceremony.

Third, each participant reviews the received contacts and chooses which ones to import.

To minimize manual interaction, PairSonic uses the smartphones’ built-in speakers and microphones as an acoustic OOB channel. This design prioritizes deployability: sound-based communication requires no custom radio hardware and works on unmodified commercial off-the-shelf (COTS) smartphones. Compared with Wi-Fi or Bluetooth, the acoustic channel has a much shorter practical range, penetrates walls less easily, and lets

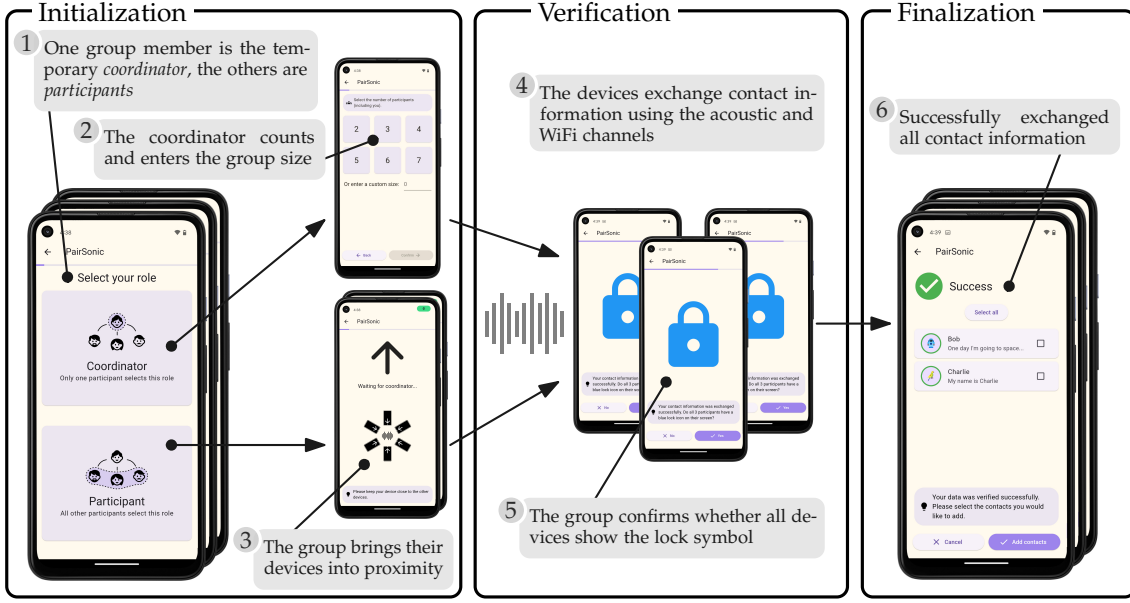


Figure 9: PairSonic user interface for a group of three users exchanging contact information. The annotations illustrate the main steps of the acoustic group authentication protocol.

users demonstratively identify participating devices by bringing them together [Bal+02]. These properties make it better suited to group authentication than pairwise methods such as NFC or QR codes. Finally, because the signal is software-defined at the application layer, acoustics can be deployed with advanced physical-layer security techniques [G].

Implementation. We implemented PairSonic as an Android app using the Flutter framework. The system relies on a novel communication stack that combines two channels. First, it creates a local ad hoc network via Wi-Fi Direct. Second, it establishes an acoustic OOB channel that transmits data via sound waves using the ggwave library [Ger26].

To verify hardware compatibility for this novel communication stack, we tested the prototype on a range of smartphones from Google, LG, Motorola, OnePlus, Oppo, Redmi, and Samsung. The application worked reliably on all tested devices that supported Wi-Fi Direct and ran Android 6.0 (API level 23, released in 2015) or newer.

Evaluation. To evaluate PairSonic’s usability, we conducted an IRB-approved within-subjects lab study with 45 participants, who tested both PairSonic and SafeSlinger. We assigned participants into groups of two to six, provided each participant with a smartphone, and tasked them with exchanging contact information using both systems.

Each group completed three trials with each system: First, in a warm-up phase, the group had to complete the authentication ceremony independently and without external guidance. Second, we simulated an active attack by a hypothetical adversary who could remotely manipulate smartphone radio communications, causing the protocol to fail on one device. Third, we disabled the attack simulation and measured the time required to complete the authentication ceremony after participants gained initial practice.

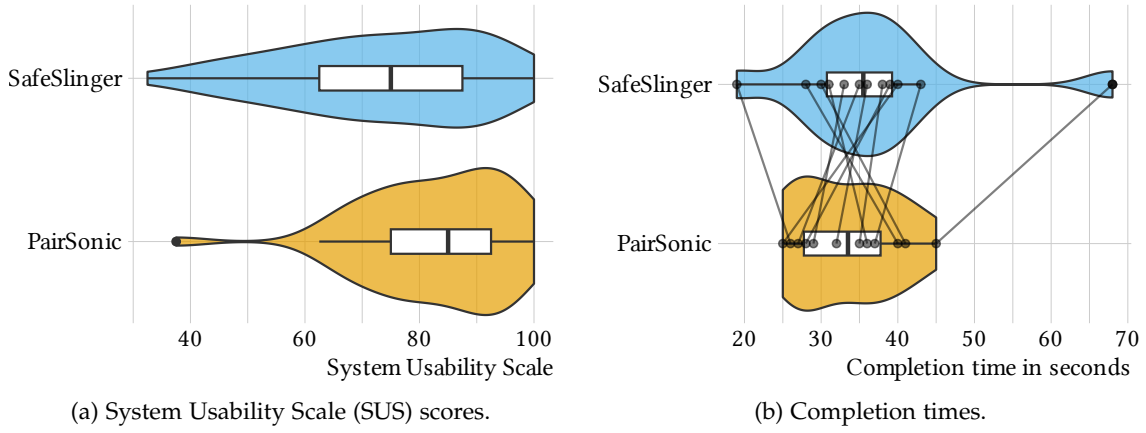


Figure 10: Participants’ SUS scores and completion times for SafeSlinger and PairSonic. Violin plots show density estimates of the distributions; boxplots show quartiles, medians, and outliers.

The sample comprised 45 participants aged 18 to 44. 22 identified as male, 18 as female, 2 as diverse, and 3 preferred not to state their gender. Most participants (37) were students, consistent with the sample’s overall high educational level, although 16 participants had no formal professional or university degree. Each participant received 15 EUR in cash as compensation, which was informed by the official minimum wage in Germany.

Our evaluation followed a mixed-methods approach combining lab-based usability testing, a post-test questionnaire, and open-ended oral interviews. The results showed that participants significantly preferred PairSonic (binomial test, $p = .016$, medium effect size $g = .19$) and gave it higher scores on the SUS (Wilcoxon signed-rank test, $V = 597.5$, $p = .004$, large effect size $r = -.43$, cf. Figure 10a). The security ratings of the two methods did not differ significantly. Most groups (67%) were faster with PairSonic, but not significantly (cf. Figure 10b). Interviews suggested that reducing user interaction improved usability but sometimes lowered perceived security, as many participants appeared to associate security with complexity based on prior experience. This highlights a challenging trade-off between actual usability and perceived security.

Overall, our study suggests that acoustic communication is a promising and user-friendly approach for decentralized ad hoc data exchange between nearby devices, reinforcing the intuitive act of bringing devices together to establish an association. However, the prototype’s acoustic channel was occasionally unreliable due to background noise and device handling, especially when users placed their devices on a table during transmission, causing impulsive noise that interfered with reception. We examine the acoustic physical layer, the causes of these failures, and potential solutions in more detail in §4.1.2.

Towards a More Robust OOB Channel. PairSonic deliberately prioritizes deployability by using an acoustic OOB channel available on unmodified smartphones. As the user study showed, however, this choice also introduced practical reliability limitations. A more robust alternative would be a custom RF-based OOB channel with physical-layer security, but such designs are typically not deployable on standard COTS devices.

Table 8: Early reception and impact of papers [D; E].

Reception Type	Evidence
Community Recognition	CSCW 2024 Best Paper Award (top 1%).
Artifact Visibility	Published as open-source software [SW3]; referenced by the gg-wave project as an example application for acoustic device pairing [Ger26].
Follow-up Research	Subsequent work evaluated more reliable acoustic transmission schemes [F] and extended the PairSonic codebase with a faster and more reliable RF-based OOB channel [N].

In follow-up work [N], we investigated whether this trade-off can be overcome on commodity smartphone hardware. Through extensive reverse engineering of a Broadcom Wi-Fi chip widely used in recent Samsung Galaxy and Google Pixel models, we modified its firmware to support arbitrary signal reception and transmission. This enabled us to implement an RF-based OOB channel with physical-layer security, which we integrated into PairSonic as a faster and more reliable alternative to the original acoustic channel.

Early Reception. Our study [D] and the corresponding demo [E] have received early recognition, artifact visibility, and follow-up research (cf. Table 8). PairSonic introduced an open-source prototype for fast and secure contact exchange in co-located groups. Building directly on our open-source contributions, subsequent research [F; N] has extended PairSonic and explored more robust channels for usable nearby-device authentication. Consequently, PairSonic serves as a reference implementation for exploring usable and secure device pairing between co-located smart devices.

Open Science. An overview of the PairSonic project is available online at <https://fputz.net/project/pairsonic>, with the open-source software hosted on GitHub [SW3]. We demonstrated the app at CSCW 2024 [E], where attendees could test it firsthand. To support replicability, we released our evaluation scripts and the pseudonymized study dataset [DS3], which includes usability, security, and preference scores; completion times; tool usage metrics; and demographic and control variables for all 45 participants.

4.1.2 Evaluating Acoustic Data Transmission Systems

Acoustic data transmission promises to be a practical alternative to Bluetooth and NFC by leveraging the speakers and microphones already present in smartphones and Internet of Things (IoT) devices. In the past two decades, acoustic communication has attracted growing research interest as a short-range, ad hoc wireless channel for mobile applications [MSS03; Mad+05; WZH21], particularly for *authentication* [Nan+13; Zha+19; D; G], where proximity and user-perceptible context can provide useful security properties. As a software-defined physical layer, acoustic communication is attractive from a *deployment* perspective because it can run entirely in software on COTS devices without requiring new hardware, making it broadly applicable to existing device ecosystems.

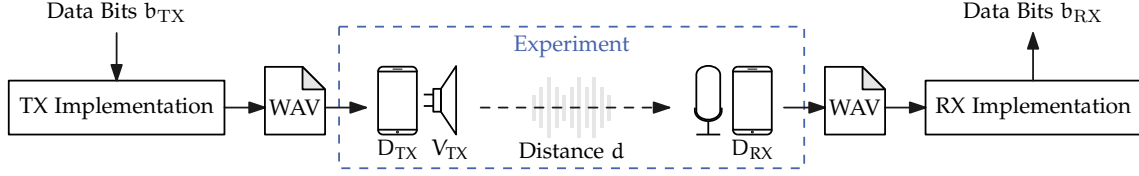


Figure 11: Acoustic data transmission model used in our experiments. Device D_{TX} plays WAV files generated by the TX implementation at volume V_{TX} , while device D_{RX} records the transmission. The resulting WAV file is decoded using the RX implementation.

Despite these advantages, acoustic channels can be sensitive to real-world operating conditions. In our lab study of PairSonic (cf. § 4.1.1), user handling and background noise frequently disrupted transmissions and caused decoding failures. These failures reduce the practicality of acoustic communication in everyday environments, especially when devices are used casually rather than under controlled conditions.

A large body of prior work proposes novel acoustic data transmission systems [Cai+22; Get+18; Gon+17; Lee+15; LA01; Zha+14]. Unfortunately, many have been evaluated only in simulation or under highly constrained real-world conditions. Testing on physical devices, however, is essential to capture hardware imperfections, realistic signal propagation, and the unpredictable effects of human handling and noise.

To address this gap, this dissertation contributes a rigorous practical evaluation of acoustic data transmission schemes in the following publication, included in Appendix F:

- [F] F. PUTZ, P. FORTMANN, J. FRANK, C. HAUGWITZ, M. KUPNIK, M. HOLLICK. “Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices.” In: *ACM Transactions on Internet of Things* 7.1, Feb. 2026. DOI: [10.1145/3779439](https://doi.org/10.1145/3779439). Appendix F.

As part of this work, we systematically reviewed 31 publications on acoustic communication for COTS devices and built a physical testbed with eight representative systems. Based on this testbed, we developed a repeatable methodology for evaluating their reliability and generalizability under realistic operating conditions. Our evaluation comprises more than 11 000 smartphone transmissions recorded both in everyday indoor environments and in a controlled anechoic chamber. Compared with prior work, we broaden the evaluation scope by testing these systems across multiple IoT use cases and explicitly quantifying the effects of distance, device model, user handling, and background noise on transmission errors. Our findings show that many existing schemes unfortunately fail to maintain reliability across practical use cases.

System Model. We evaluated *acoustic data transmission schemes*, i.e., software implementations that broadcast data to nearby devices using sound waves. We treat each scheme as a black box with two functions: a transmitter that converts data bits into a WAV file for playback, and a receiver that extracts data from a recorded WAV file (cf. Figure 11). Each scheme is evaluated as a complete communication system from a *link-layer* (Layer 2) perspective, including framing, synchronization, and error correction.

Table 9: Evaluated acoustic data transmission schemes, sorted by throughput. Implementation sources are encoded as follows: P – referenced in the paper; A – obtained by contacting the authors; R – re-implemented by us; N – non-academic solution included for comparison. The last three columns indicate each scheme’s a priori suitability for our three use case categories based on specified throughput and maximum distance.

Via	Name	Modulation	Frequencies	Inaudible	Throughput ¹	Near	Medium	Far
R	Lee et al. [Lee+15]	Chirp BOK	19.5–22 kHz	✓	15 bps	✗	✗	✓
A	Digital Voices [LA01]	MFSK	1–3 kHz	✗	80 bps	✗	✓	– ²
R	Nearby [Get+18]	DSSS + MFSK	18.5–20 kHz	✓	84 bps	✗	✓	✓
N	ggwave _a [Ger26]	MFSK	1.8–6.3 kHz	✗	268 bps	✗	✓	– ²
N	ggwave _i [Ger26]	MFSK	15–19.5 kHz	✓	268 bps	✗	✓	– ²
A	HRCSS [Cai+22]	OCSS	18–22 kHz	✓	500 bps	✓	✓	✓
A	Gonçalves et al. [Gon+17]	OFDM + QPSK	0.1–4 kHz	✗	600 bps	✓	✗	✗
R	PriWhisper [Zha+14]	MFSK	9–17 kHz	✗	729 bps	✓	✗	✗

¹ We list the net data rate, as this is the usable data rate after error correction from the user’s perspective.

² Maximum distance not specified, therefore no a priori assessment.

Scenarios. We focus on data exchange between nearby smartphones in settings where software can be modified (e.g., by installing smartphone apps), but hardware and firmware cannot. We group common use cases for acoustic communication into three categories based on communication distance, which is a discerning factor for usability:

1. **Near distance** ($d \leq 10$ cm): transmitting a public key or small file (≈ 4096 bit) over very short range, similar to NFC.
2. **Medium distance** (10 cm $< d \leq 1$ m): transmitting a public key fingerprint or hash value (≈ 128 bit) over moderate range.
3. **Far distance** ($d > 1$ m): transmitting a short ID (≈ 16 bit) over larger distances.

Acoustic Schemes. We surveyed the literature on acoustic communication schemes (cf. §4.2.5) and contacted all authors to request implementations. Unfortunately, we received only three working implementations from them. We therefore re-implemented three additional schemes from highly cited publications, selected to represent our three target use cases. For comparison, we also included two non-academic implementations from the popular open-source project ggwave.¹ Table 9 lists all eight evaluated schemes.

Evaluation. We conducted four experiments that varied distance, device model, background noise, and transmission environment (cf. Figure 12). We measured reliability using the total error rate (TER), defined as the final bit error rate after error correction, with synchronization failures counted as 100% TER.

In a typical office environment, most schemes were reliable only at short range (cf. Figure 13). At longer distances, only the inaudible ggwave variant and the scheme by Lee et al. [Lee+15] remained reliable. Unfortunately, the implementation of HRCSS [Cai+22]

¹ The ggwave project [Ger26] is the most-starred *data-over-sound* project on GitHub: <https://github.com/topics/data-over-sound> (as of May 20, 2026).

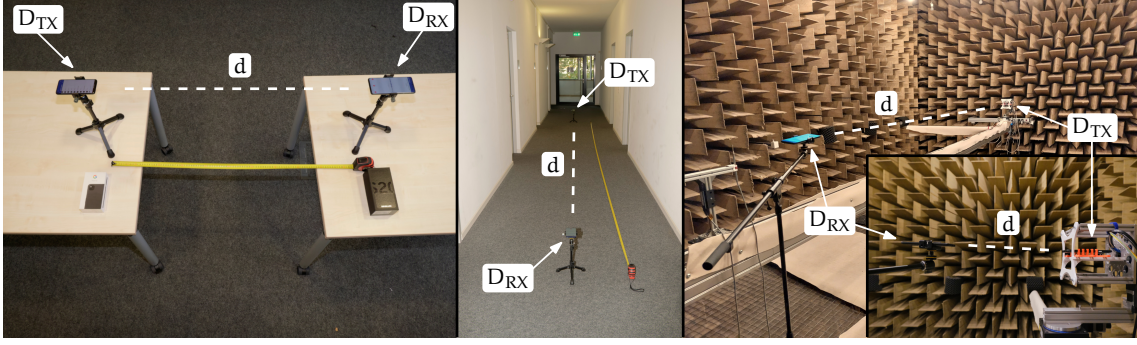


Figure 12: Experimental setups used to evaluate acoustic transmissions: office setup (*left*), office hallway (*middle*), and anechoic chamber (*right*).

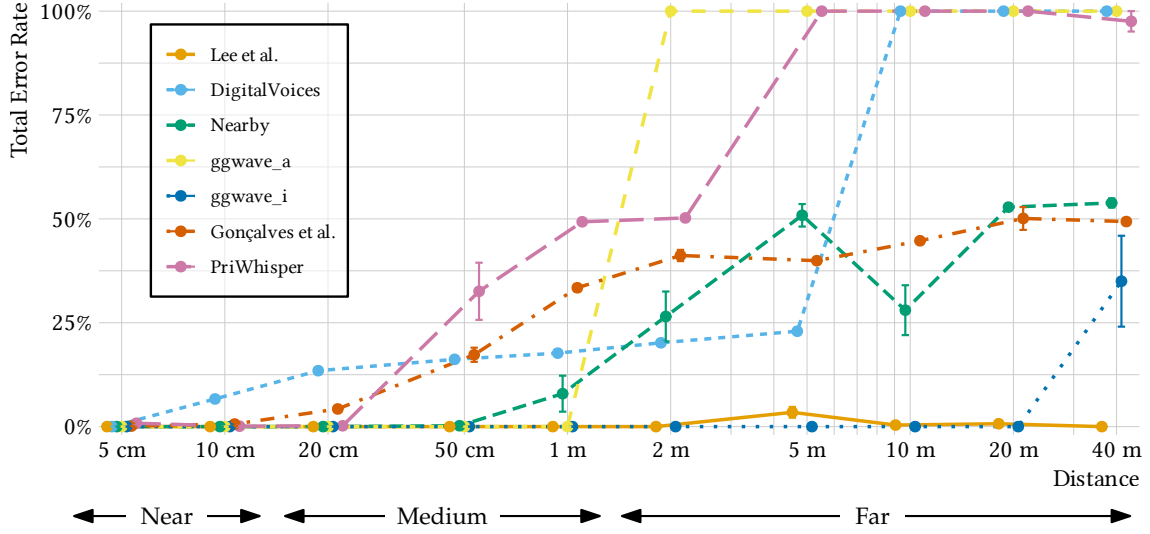


Figure 13: Reliability of acoustic transmissions at different distances, measured in a quiet indoor environment. Lower TER indicates higher reliability; error bars show standard errors. Distances are shown on a logarithmic axis.

provided by the authors was faulty and caused synchronization failures even when tested with studio hardware. We therefore excluded this scheme from further experiments.

Overall, our results show that many schemes are of limited practical use in realistic indoor settings. Among the medium-range schemes, both ggwave variants performed well, especially the inaudible version, while Nearby [Get+18] also performed strongly but did not generalize across all device combinations. Figure 14 summarizes how well each scheme performs for its intended use case in typical indoor environments.

Practical Challenges for Acoustic Communication. Our experiments identify four main challenges for acoustic communication on commodity devices that rarely appear in simulations or tightly controlled lab settings: multipath propagation, large variation in audio hardware across devices, limited usable bandwidth, and interference from everyday noise sources such as speech and device handling.

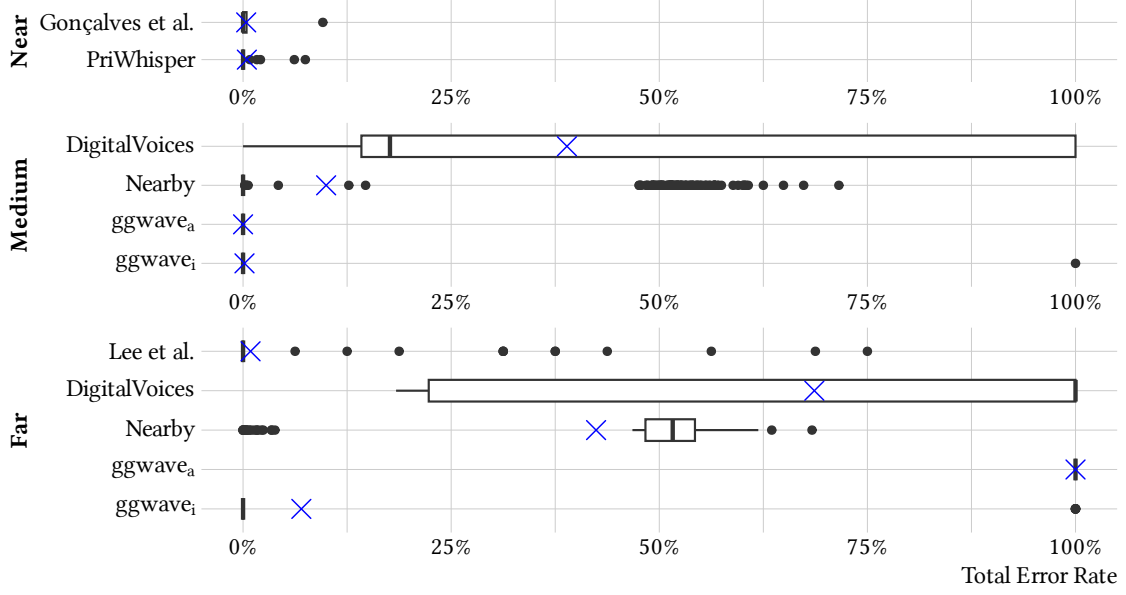


Figure 14: Summary of transmission results in a quiet indoor environment by use case category. Lower TER indicates higher reliability; schemes are sorted by throughput within each category. Box plots show medians (black bar), means (blue cross), and outliers.

These results suggest a practical design pattern: use the acoustic channel only to transmit a short and robust message, such as an identifier or hash, while transferring larger data over a higher-throughput channel such as Wi-Fi or cellular.

Open Science. Our evaluation demonstrates that many existing acoustic data transmission systems fail to remain reliable in practical scenarios. This gap highlights the need for rigorous and reproducible methodologies that assess how new protocols perform under real-world conditions. Currently, researchers face difficulties in replicating prior work because implementations are often not publicly available and testing methods are not transparent or inconsistent. Consequently, most studies evaluate new schemes in isolation, which limits meaningful comparative analysis.

To address this limitation, we contribute a black-box analysis framework. By evaluating acoustic data transmission systems solely through their observable inputs and outputs rather than their internal code, this framework can be applied to any unidirectional system and remains useful for future protocols. To support continued research, we document our experimental setup in Appendix F. In addition, we have released a public replication package [DS4]. This repository contains our full dataset of 11 900 recorded audio transmissions across eight systems, together with our software re-implementations of Nearby [Get+18], PriWhisper [Zha+14], and the scheme proposed by Lee et al. [Lee+15].

4.2 Related Work

This section situates our work on usable authentication via acoustic communication within prior research. We first examine authentication ceremonies in secure messengers,

user interactions for device association, and secure device pairing protocols, emphasizing their limitations in group settings. We then compare our approach to SafeSlinger [Far+13] and discuss acoustic data transmission on mobile devices.

4.2.1 Authentication Ceremonies in Secure Messengers

Modern E2EE messengers such as Signal and WhatsApp use opportunistic E2EE (trust on first use), which requires little user involvement but does not protect against active attacks unless users explicitly verify each other’s cryptographic keys through an *authentication ceremony* [AS23; Her+21; Ung+15]. In practice, these ceremonies typically ask users to scan a QR code or compare human-readable representations of key fingerprints, such as hexadecimal strings, word phrases, or visual encodings [Dec+16; Tan+17a].

Prior work shows that authentication ceremonies are difficult to use [Abu+17; Ass+15], error-prone [SS15; SSG17], and therefore rarely completed in practice [FK23; Sch+16; Vaz+18]. As a result, users who skip authentication often have weaker security than they expect, and many are unaware of this gap [HL16]. Furthermore, current authentication ceremonies scale poorly to groups because they require repeated pairwise verification, which does not generalize well to multiple participants [Far+13].

4.2.2 User Interactions for Device Association

Prior work has studied how users naturally associate nearby devices through physical actions and gestures [CG12; CMG14; Grø+20; Kra+10]. Although not security-focused, this work informs the design of usable authentication interactions. In particular, studies on group device association show that users prefer one shared interaction over repeated pairwise steps, typically by bringing devices close together or pointing them towards one another [CG11; CG13; Jok+15; JL14; JRV16]. PairSonic follows this design principle by using a single proximity-based group interaction for contact exchange.

4.2.3 Secure Device Pairing

Secure device pairing aims to establish an authentic channel between nearby devices [Fom+18; May08]. Most prior work focuses on two-device pairing [CG12; Kum+09; UKA07], whereas our setting involves two or more users. This case is more challenging because pairwise techniques do not securely generalize to groups [Far+13; KSP08].

Prior work on multi-device pairing can be grouped into four broad categories based on how they establish trust: human-in-the-loop approaches, zero-interaction schemes based on shared context, physical-layer security, and OOB channels.

Humans-in-the-loop. Human-in-the-loop schemes require users to perform explicit actions, such as entering a shared secret [AG00], comparing text [LP08] or visual patterns [Che+08; Lin+09], or moving devices together [MG09]. While these actions can help authenticate devices, they typically impose considerable user effort, scale poorly to groups, and often fail to protect against malicious insiders [KSP08].

Shared Homogeneous Context. Zero-interaction approaches derive shared keys from common environmental context [Fom+19; Xu+21]. Although attractive in principle, they are not well suited to ad hoc, user-controlled group association because they are designed for slower background pairing among devices of a single user [Far+23; Li+20].

Physical Layer Security. Other approaches exploit properties of the wireless channel or shielding techniques such as Faraday cages to support secure key establishment [GLL17; HLG13; Kuo+07; Law+11]. However, these methods often require specialized hardware, firmware, or additional equipment, which limits deployability on COTS smartphones.

Out-of-Band Channels. A further line of work uses OOB channels, i. e., low-bandwidth auxiliary channels whose physical properties help protect message integrity [SA00]. Prior work has explored infrared [Bal+02], light [SU09], NFC [CKG11], and audio channels [Goo+06; STU08; G]. PairSonic builds on this line of work by using an acoustic OOB channel for decentralized group authentication.

4.2.4 SafeSlinger

SafeSlinger is the closest prior work on group pairing and secure contact exchange [Far+13]. Building on earlier group pairing approaches [Che+08; Lin+09], it improved the usability and security of human-in-the-loop authentication for groups.

However, SafeSlinger still requires substantial manual coordination: users must agree on the group size, enter a shared value, compare verification phrases, and confirm the exchanged contact information. These steps become increasingly cumbersome in larger groups and remain prone to user error [KFR09; Tan+17a]. In addition, SafeSlinger depends on Internet connectivity and a central server.

PairSonic builds on SafeSlinger’s cryptographic foundations but simplifies the user interaction through an acoustic OOB channel and direct local communication. Instead of repeated group-wide coordination and phrase comparison, PairSonic reduces verification to a lightweight visual check, which scales better with group size (cf. Figure 15). At the same time, it removes the dependence on external infrastructure, making the protocol more deployable in decentralized and low-connectivity settings.

4.2.5 Acoustic Data Transmission

In the following, we briefly introduce acoustic data transmission as a form of wireless communication, relate academic proposals to commercial systems, and summarize application scenarios explored in prior work.

Background. Although the audio hardware of smartphones and IoT devices is primarily designed for media playback and voice communication, it can also be repurposed for aerial acoustic data transmission. Acoustic modems encode data in sound by modulating frequency, amplitude, or related signal properties, enabling communication among heterogeneous COTS devices without dedicated RF hardware [GB00; LA03]. Depending

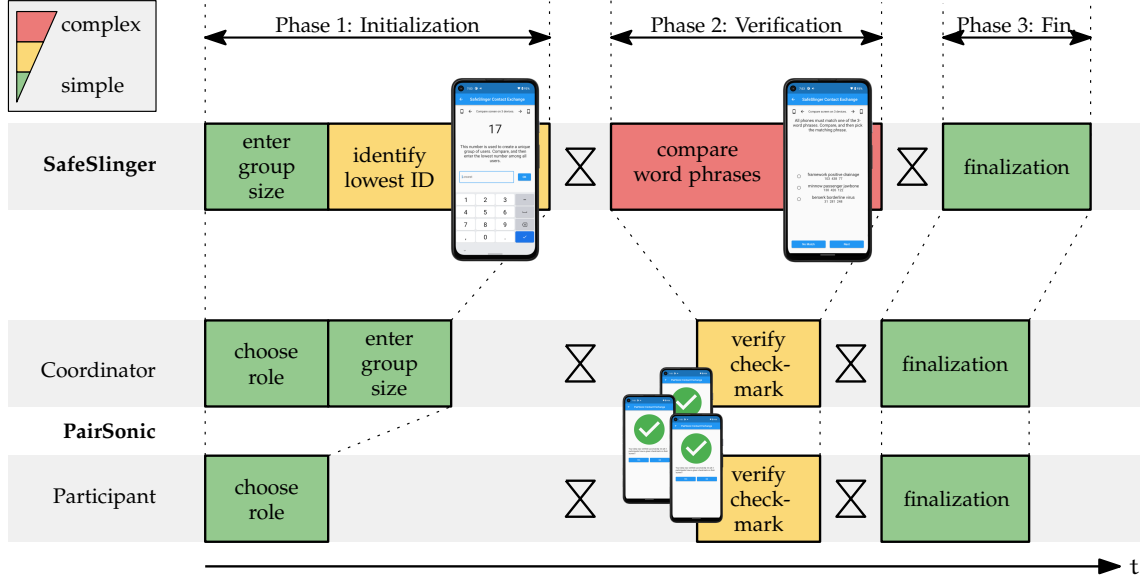


Figure 15: Qualitative timelines of the contact exchange process for a typical user group. The waiting symbol $\times$ indicates idle time. Color-coded actions represent increasing complexity and duration, illustrating PairSonic’s better scalability for larger groups.

on the application, transmissions may use audible or near-ultrasonic frequencies, with software-defined parameters allowing adaptation to different requirements.

Commercial Practice. A clear gap exists between academic and commercial acoustic communication systems. Commercial products for smart devices typically operate at only 10 bps to 100 bps [CUE; Get+18; Jon22], prioritizing robustness in noisy real-world environments and across heterogeneous commodity hardware. In contrast, academic systems often report substantially higher throughputs, ranging from roughly 500 bps to 10 000 bps [Cai+22; Gon+17; Zha+14] and sometimes beyond [YK22].

This discrepancy raises an important question: whether the higher rates reported in the literature can be sustained under realistic deployment conditions, or whether practical constraints such as hardware imperfections, user handling, and environmental noise ultimately force reliable systems to operate at much lower data rates. Addressing this question was a key motivation for the evaluation presented in §4.1.2.

Academic Approaches. To identify prior work on aerial acoustic software-defined communication systems for COTS devices such as smartphones, we conducted a systematic literature review. Our method was guided by PRISMA [Pag+21b], a structured framework for surveying and filtering large bodies of research. Starting from 631 papers, we excluded 422 as out of scope based on their abstracts. We then classified the remaining 209 papers according to their applicability to ubiquitous computing use cases and whether they presented a testable implementation for COTS devices. This process yielded 31 relevant papers published between 2000 and 2022 that describe testable acoustic communication schemes, summarized in Table 10.

Table 10: Surveyed acoustic data transmission schemes, sorted by throughput. The table continues from left to right; unavailable data is marked with “?”.

Name	Modulation	Inaudible	Throughput [bps]	Name	Modulation	Inaudible	Throughput [bps]
Moriyama et al. [MTT17]	ASK	✓	5	Gonçalves et al. [Gon+17]	OFDM + PSK	✗	600
Kim et al. [KML19]	FSK	✓	5	PriWhisper [Zha+14; Zha+19]	FSK	✗	1000
Ka et al. [Ka+16]	Chirp	✓	15	Digital Voices [LA01]	ASK, FSK, SS	✗	1280
Lee et al. [Lee+15]	Chirp	✓	16	A-NFC [LHC15]	QPSK	✓	2200
Lee et al. [Lee+20]	Chirp	✓	16	Dhwani [Nan+13]	OFDM + PSK	✗	2400
Hornych et al. [Hor+20]	ASK, FSK	✓	33	Hush [NTL19]	OFDM + QAM	✓	4900
TagScreen [LYL17]	Chirp	✓	36	He et al. [HB18]	AM, FSK	✗	6000
Haus et al. [Hau+19]	OFDM + ASK	✓	64	Yamamoto et al. [YK22]	QAM	?	32000
Miegel et al. [MBB18]	OOK, FSK	✓	84	Arentz et al. [AB11]	ASK	✓	?
Rustle [Jin+22]	custom ¹	✗	90	Bits-over-Sound [Kan18]	DBPSK	✗	?
Nearby [Get+18]	DSSS + FSK	✓	95	Jeon et al. [JKL16]	PSK	✓	?
Bang et al. [BCK16]	FSK	✓	170	No Entry [Kwa+19]	Chirp	?	?
Sadeq et al. [SH22]	FSK	✓	440	Sasano et al. [Sas+14b]	OFDM + PSK	✗	?
HRCSS [Cai+22]	Chirp	✓	500	SilentInformer [SCC20]	FSK	✓	?
Dolphin [Wan+16; Zho+19]	OFDM + ASK	✗	500				

¹ Rustle’s custom approach is called “furtive modulation”.

Usability. Beyond technical feasibility, researchers have also examined the user experience of acoustic communication [LA01; LA03; Meh+19]. This strand of research is particularly important because, unlike conventional wireless technologies, acoustic transmissions may be audible to users and bystanders, making them not only a communication mechanism but also an interaction cue.

Use Cases. Acoustic data transmission has been explored in prior work as a short-range communication channel for a variety of ubiquitous and mobile computing scenarios. Early research considered its use for context-aware computing [MSS03; Mad+05], while later work investigated communication between nearby devices when direct, localized exchange is desirable [AB11]. Because acoustic signals typically propagate only over limited distances and are constrained by the surrounding environment, they have also been studied for secure device interactions. Examples include device pairing [Goo+09; Goo+06; Zha+14; G] and access control for implantable medical devices [Hal+08].

The practical relevance of these use cases is reflected in commercial systems, which employ acoustic communication for mobile payments [Cop25; Ste14], context-aware applications [LIS], location-based services [CUE], and device pairing [Kum13; Tri20]. However, despite these deployments, prior research has not systematically evaluated acoustic communication systems on real devices under realistic conditions.

Conclusions

In this final chapter, we summarize the findings of this thesis (cf. § 5.1) and outline important perspectives for future work (cf. § 5.2).

5.1 Findings

This thesis investigated how mobile authentication protocols can be made more practical by improving their usability and deployability. Below, we briefly summarize our findings on web authentication, decentralized authentication, and messenger authentication.

Web Authentication. We were the first to show that both forms of FIDO2 authentication are highly usable on smartphones, with users expressing a slight preference for on-device biometrics (platform authentication) over external security tokens (roaming authentication). Considering that both forms of web authentication are now supported by most smartphones and web browsers, our findings indicate that FIDO2 provides users and service providers with practical and strong authentication mechanisms that are arguably more usable and deployable than previous attempts to replace passwords. We also identified remaining adoption barriers, particularly account recovery, revocation, and multi-device use. These issues have since started to be addressed by the FIDO Alliance through synced passkeys, indicating that our study captured practically significant barriers to passwordless authentication adoption at an important transitional moment.

At the same time, we found that the practical extensibility of the FIDO2 stack is limited, as browsers do not pass through third-party FIDO2 extensions. Consequently, custom extensions cannot realistically be deployed outside controlled or internal environments.

Decentralized Authentication. Our results show that, in crisis scenarios with severely limited Internet connectivity, users prefer to rely on familiar general-purpose apps for essential tasks rather than switching to dedicated crisis apps. Current mobile ecosystems, however, are not designed for this setting. Mainstream applications and cellular infrastructures commonly depend on centralized backend services and therefore often cease to function when these services become unreachable.

This thesis identifies centralized authentication as a critical point of failure in such scenarios. If essential applications are to remain usable during network outages, authentication must no longer depend exclusively on globally reachable infrastructure. This creates a clear need for decentralized or offline-capable authentication mechanisms that preserve security under local-only or intermittent connectivity. This need is particularly acute for messengers, which our study identified as the most important application category in both crises and everyday use.

Messenger Authentication. To address the problem of decentralized authentication in the messenger domain, we developed the open-source Android app PairSonic as an authentication ceremony for groups of users. PairSonic enables users to authenticate each other simultaneously, substantially reducing user effort compared to current authentication ceremonies in mainstream messengers such as Signal. Instead of relying on manual fingerprint comparisons, PairSonic uses an acoustic out-of-band channel, which our study participants found significantly more usable than the state of the art.

We further showed that PairSonic is compatible with all tested Android smartphones and requires only a speaker and microphone, making it highly deployable in practice. Our experiments identified four central challenges for acoustic communication on commodity devices that are rarely captured in simulations or tightly controlled lab settings and therefore require extensive real-world evaluation: multipath propagation, large variation in audio hardware across devices, limited usable bandwidth, and interference from everyday noise sources such as speech and device handling. These findings suggest a practical design pattern for mobile authentication: use the acoustic channel only to transmit a short, robust message, such as an identifier or hash, while transferring larger data over a higher-throughput channel such as Wi-Fi or cellular.

5.2 Future Work

In the following, we outline several directions for future work across the three investigated applications: web authentication (cf. § 5.2.1), decentralized authentication (cf. § 5.2.2), and messenger authentication (cf. § 5.2.3).

5.2.1 Web Authentication

While passkeys improve the security of web accounts, several open challenges remain in making them recoverable, understandable, and extensible in practice.

Account Recovery. Our lab study showed that users fear losing their authenticator and thus access to their accounts. While passkeys mitigate phishing at sign-in, the weakest link often shifts to account recovery procedures: SMS fallbacks remain phishable, and human-assisted recovery is vulnerable to social engineering, potentially reintroducing risks that passkeys are meant to avoid. Unfortunately, the FIDO Alliance’s recommendation to register a backup authenticator for recovery adds registration effort and can feel cumbersome [Mat+25].

Synced passkeys from providers such as Apple and Google are convenient for users, but this central dependency introduces a single point of failure. NIST explicitly identifies cloud-based recovery as a potential weakness, since a synced passkey is only as strong as the recovery process of its cloud account (US NIST SP 800-63B-4, Appendix B).

Prior work proposes cryptographic backup protocols for FIDO2 credentials [Aro+22; BSU25; Fry+20], but the socio-technical challenges of smartphone recovery remain unresolved: how users recover without a backup authenticator, how recovery interacts with synced passkey providers and their recovery mechanisms, and how services can design

understandable, secure lifecycle and escalation workflows that balance registration and recovery effort without falling back to phishable methods.

User Education. Website operators and the FIDO Alliance should improve communication around recovery, revocation, delegation, and multi-device use. Future research should examine mental models, onboarding messages, and risk communication (e.g., building upon the work of Amer et al. [Ame+25]), address common misconceptions and mistrust, and measure long-term user understanding.

Third-Party FIDO2 Extensions. Current FIDO2 client implementations restrict extensibility due to missing features and complex software architectures. Web browsers should consider (1) offering a configuration option for FIDO2 extension pass-through and (2) allowing web browser extensions to support custom FIDO2 extensions. Both steps would empower researchers and developers to build their own extensions, as described in [B].

5.2.2 Decentralized Authentication

Many messaging and service apps still rely on centralized trust infrastructure or remote identity providers, which makes fully local operation challenging.

Distributing Identity State. Decentralized authentication is required for both user-to-user and user-to-service trust. While mechanisms like our proposed PairSonic address user-to-user authentication by allowing the local, offline exchange of identity material, user-to-service authentication remains challenging to support in distributed deployments. Specifically, maintaining consistent server-side identity state, key updates, and revocation across intermittently disconnected islands requires further research.

User-Friendly Transitions for Island Operation. Our findings show that users prefer continuing to use general-purpose apps during crises rather than switching to dedicated crisis apps. Service providers and network operators therefore share a responsibility to support seamless transitions into and out of island operation as connectivity changes. Technically, this likely requires adopting local-first communication patterns, multi-access edge computing, and disruption-tolerant networking. Further usability research is needed to determine how to best communicate these network transitions, explain temporary functionality limitations, and establish local trust.

5.2.3 Messenger Authentication

Future work on messenger authentication spans two complementary areas. On the user side, verification mechanisms must become more understandable, discoverable, and compatible with everyday communication practices. On the technical side, the acoustic out-of-band channel requires more reliable, device-independent designs and more replicable evaluation practices.

Usable MitM Prevention. Mainstream messenger authentication ceremonies suffer from poor usability and limited adoption [Her+21]. Future work should therefore explore risk-aware authentication support: when perceived risk is low, users may rationally choose not to verify keys, so interfaces should help them understand the relevant threats, consequences, and appropriate responses rather than merely maximizing verification rates [Wu+19]. A promising direction is to automatically identify conversations or moments that warrant authentication (e.g., using local machine learning) and provide context-sensitive nudges [RGP26], though the ethical tradeoffs of such nudges must be carefully weighed. Users also need support for planning and completing the ceremony [FK23], especially in high-risk situations [Her+21].

Improving the user experience further requires studying ceremony discoverability and interoperable key management across end-to-end encryption (E2EE) messengers, so that users can authenticate contacts more easily across different apps and devices. These improvements should be evaluated through long-term field studies that capture real-world behavior, including repeated interactions, changing threat perceptions, and group dynamics, ideally in cooperation with messaging service providers [FK23].

As an alternative to explicit ceremonies, recent work explores automatic authentication approaches based on social authentication [LB24; Vaz+19] or key transparency [Fis+25; Mal+23]. For these emerging systems, future research should focus on user education, establishing accurate mental models, and designing effective, actionable warnings that help users respond appropriately when attacks are detected.

Reliable Acoustic Communication. PairSonic’s acoustic OOB channel was sensitive to environmental noise, reducing usability in everyday settings. Our follow-up evaluation revealed that existing acoustic transmission schemes generalize poorly across smartphone models due to hardware variances (e.g., differing frequency responses, amplifier nonlinearities, and on-chip post-processing). In our tests, no scheme reliably exceeded 270 bps at a distance of 1 m under realistic conditions. Consequently, a critical open question is whether it is possible to design high-data-rate acoustic communication schemes that remain robust to realistic noise and operate reliably across diverse off-the-shelf devices.

Improving Replicability. Finally, the field of acoustic communication research should urgently improve its adherence to open science practices to ensure scientific replicability. Currently, the widespread use of disparate, undocumented experimental setups makes comparative analyses between studies nearly impossible. To foster rigorous science, the community should prioritize transparency by sharing code, publishing datasets, and establishing standardized evaluation benchmarks. While standardizing tests outside of controlled settings (like anechoic chambers) is difficult, our unified experimental framework (detailed in Appendix F) offers a foundational step towards testing acoustic communication schemes in a more transparent, standardized, and reproducible manner.

Bibliography

This dissertation was written independently and was linguistically revised with the help of ChatGPT and Gemini.

- [3GP23] 3GPP. *System Architecture for the 5G System (5GS) 23.501, Version 18.4.0*. Technical Specification. 3rd Generation Partnership Project (3GPP), Dec. 2023. URL: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144> (visited on May 20, 2026).
- [Abu+17] R. ABU-SALMA, M. A. SASSE, J. BONNEAU, A. DANILOVA, A. NAIKSHINA, M. SMITH. "Obstacles to the Adoption of Secure Communication Tools." In: *2017 IEEE Symposium on Security and Privacy (S&P '17)*. May 2017, pp. 137–153. DOI: [10.1109/SP.2017.65](https://doi.org/10.1109/SP.2017.65).
- [Aeb+20] S. AEBISCHER, C. DETTONI, G. JENKINSON, K. KROL, D. LLEWELLYN-JONES, T. MASUI, F. STAJANO. "Deploying Authentication in the Wild: Towards Greater Ecological Validity in Security Usability Studies." In: *Journal of Cybersecurity* 6.1, Nov. 2020. DOI: [10.1093/cybsec/tyaa010](https://doi.org/10.1093/cybsec/tyaa010).
- [AE22] M. A. AL KABIR, W. ELMEDANY. "An Overview of the Present and Future of User Authentication." In: *4th IEEE Middle East and North Africa COMMunications Conference (MENACOMM '22)*. Dec. 2022, pp. 10–17. DOI: [10.1109/MENACOMM57252.2022.9998304](https://doi.org/10.1109/MENACOMM57252.2022.9998304).
- [AS23] M. ALATAWI, N. SAXENA. "SoK: An Analysis of End-to-End Encryption and Authentication Ceremonies in Secure Messaging Systems." In: *Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '23)*. ACM, June 2023, pp. 187–201. DOI: [10.1145/3558482.3581773](https://doi.org/10.1145/3558482.3581773).
- [Alq+20] F. ALQUBAISI, A. S. WAZAN, L. AHMAD, D. W. CHADWICK. "Should We Rush to Implement Password-Less Single Factor FIDO2 based Authentication?" In: *12th Annual Undergraduate Research Conference on Applied Computing (URC '20)*. 2020. DOI: [10.1109/URC49805.2020.9099190](https://doi.org/10.1109/URC49805.2020.9099190).
- [Ame+25] Y. AMER, E. U. HAQUE, Z. RONG, M. M. HASAN KHAN. "Understandability of the Technology and Benefit May Not Be Enough to Nudge Users: An Exploratory Study in the Context of FIDO2 Adoption Behavior." In: *2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC '25)*. July 2025, pp. 607–618. DOI: [10.1109/COMPSAC65507.2025.00083](https://doi.org/10.1109/COMPSAC65507.2025.00083).
- [Amf24] S. M. AMFT. "On the Usability of Authentication Security Communication." PhD Thesis. Leibniz University Hannover, Sept. 2024. DOI: [10.15488/17959](https://doi.org/10.15488/17959).

- [APX24] A. ANGELOGIANNI, I. POLITIS, C. XENAKIS. “How Many FIDO Protocols Are Needed? Analysing the Technology, Security and Compliance.” In: *ACM Computing Surveys* 56.8, Apr. 2024. DOI: [10.1145/3654661](https://doi.org/10.1145/3654661).
- [App20a] APPLE INC. *Meet Face ID and Touch ID for the Web*. WWDC20, session 10670. 2020. URL: <https://developer.apple.com/videos/play/wwdc2020/10670/> (visited on May 20, 2026).
- [App22b] APPLE INC. *Apple, Google, and Microsoft Commit to Expanded Support for FIDO Standard to Accelerate Availability of Passwordless Sign-Ins*. Apple Newsroom. May 5, 2022. URL: <https://www.apple.com/newsroom/2022/05/apple-google-and-microsoft-commit-to-expanded-support-for-fido-standard/> (visited on May 20, 2026).
- [App24a] APPLE INC. *About iOS 13 Updates*. Section: iOS 13.3. Feb. 26, 2024. URL: <https://support.apple.com/en-us/118392#133> (visited on May 20, 2026).
- [App26] APPLE INC. *Make Your Passwords and Passkeys Available Across Devices with iPhone and iCloud Keychain*. iPhone User Guide, iOS 26. 2026. URL: <https://support.apple.com/guide/iphone/passwords-devices-iph82d6721b2/26/ios/26> (visited on May 20, 2026).
- [AB11] W. A. ARENTZ, U. BANDARA. “Near Ultrasonic Directional Data Transfer for Modern Smartphones.” In: *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp ’11)*. ACM, Sept. 2011, pp. 481–482. DOI: [10.1145/2030112.2030181](https://doi.org/10.1145/2030112.2030181).
- [Aro+22] S. S. ARORA, S. BADRINARAYANAN, S. RAGHURAMAN, M. SHIRVANIAN, K. WAGNER, G. WATSON. *Avoiding Lock Outs: Proactive FIDO Account Recovery Using Managerless Group Signatures*. Technical Report. 2022. URL: <https://eprint.iacr.org/2022/1555>.
- [AG00] N. ASOKAN, P. GINZBOORG. “Key Agreement in Ad Hoc Networks.” In: *Computer Communications* 23.17, Nov. 2000, pp. 1627–1637. DOI: [10.1016/S0140-3664\(00\)00249-8](https://doi.org/10.1016/S0140-3664(00)00249-8).
- [Ass+15] H. ASSAL, S. HURTADO, A. IMRAN, S. CHIASSON. “What’s the Deal with Privacy Apps? A Comprehensive Exploration of User Perception and Usability.” In: *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM ’15)*. ACM, Nov. 2015, pp. 25–36. DOI: [10.1145/2836041.2836044](https://doi.org/10.1145/2836041.2836044).
- [Bal+23] L. BALDUF, M. KORCZYŃSKI, O. ASCIGIL, N. V. KEIZER, G. PAVLOU, B. SCHEUERMANN, M. KRÓL. “The Cloud Strikes Back: Investigating the Decentralization of IPFS.” In: *Proceedings of the 2023 ACM on Internet Measurement Conference (IMC ’23)*. ACM, Oct. 2023, pp. 391–405. DOI: [10.1145/3618257.3624797](https://doi.org/10.1145/3618257.3624797).
- [Bal+02] D. BALFANZ, D. K. SMETTERS, P. STEWART, H. C. WONG. “Talking to Strangers: Authentication in Ad-Hoc Wireless Networks.” In: *Proceedings of the Network and Distributed System Security Symposium (NDSS ’02)*. 2002. URL: <https://www.ndss-symposium.org/ndss2002/talking-strangers-authentication-ad-hoc-wireless-networks/> (visited on May 20, 2026).

- [BCK16] G. BANG, M. CHUNG, I. KO. "Data Communication Method Based on Inaudible Sound at Near Field." In: *2016 18th International Conference on Advanced Communication Technology (ICACT '16)*. 2016, pp. 639–642. DOI: [10.1109/ICACT.2016.7423502](https://doi.org/10.1109/ICACT.2016.7423502).
- [Bav25] S. BAVENDIEK. "End-to-End Encrypted Cloud Services with FIDO2." In: *Crypto Day Matters 38*. Gesellschaft für Informatik e.V. / FG KRYPTO, 2025. DOI: [10.18420/cdm-2025-38-11](https://doi.org/10.18420/cdm-2025-38-11).
- [BSU25] K. BICAKCI, M. SEKMEN, Y. UZUNAY. "Balancing Security and Usability in Enterprise FIDO Passkey Recovery With Peer Assistance." In: *IEEE Access* 13, 2025, pp. 150415–150425. DOI: [10.1109/ACCESS.2025.3602362](https://doi.org/10.1109/ACCESS.2025.3602362).
- [BU22] K. BICAKCI, Y. UZUNAY. "Is FIDO2 Passwordless Authentication a Hype or for Real?: A Position Paper." In: *2022 15th International Conference on Information Security and Cryptography (ISCTURKEY '22)*. 2022, pp. 68–73. DOI: [10.1109/ISCTURKEY56345.2022.9931832](https://doi.org/10.1109/ISCTURKEY56345.2022.9931832).
- [Bir+26] J. BIRKMANN, E.-M. KLOPRIES, H. SCHÜTTRUMPF, A. TRÜDINGER, S. STENGER-WOLF, eds. *Die Flutkatastrophe 2021: Vom Wiederaufbau zur Klimaresilienz*. Sustainable Development Goals (SDG) – Umsetzung in Praxis, Lehre und Entscheidungsprozessen. Wiesbaden: Springer Fachmedien, 2026. DOI: [10.1007/978-3-658-48688-4](https://doi.org/10.1007/978-3-658-48688-4).
- [Ble+25] J. BLESSING, D. HUGENROTH, R. ANDERSON, A. BERESFORD. "SoK: Web Authentication and Recovery in the Age of End-to-End Encryption." In: *Proceedings on Privacy Enhancing Technologies (PoPETs '25)* 3, 2025, pp. 560–589. DOI: [10.56553/popets-2025-0113](https://doi.org/10.56553/popets-2025-0113).
- [Ble25] V. BLESSING. "Securing End-to-End Encrypted Systems." PhD Thesis. University of Cambridge, May 2025. DOI: [10.17863/CAM.119942](https://doi.org/10.17863/CAM.119942).
- [BDW26] N. BOCK, C. DITTMER, P. WINDSHEIMER-KOLLA. *Civil Society Action and Operational Disaster Management: The Berlin Power Outage in January 2026*. Disaster Research Blog. Apr. 10, 2026. DOI: [10.17169/refubium-52004](https://doi.org/10.17169/refubium-52004).
- [Boe+08] S. BOEYEN, S. SANTESSON, T. POLK, R. HOUSLEY, S. FARRELL, D. COOPER. *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*. RFC 5280. May 2008. DOI: [10.17487/RFC5280](https://doi.org/10.17487/RFC5280).
- [Bon+12] J. BONNEAU, C. HERLEY, P. C. VAN OORSCHOT, F. STAJANO. "The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes." In: *2012 IEEE Symposium on Security and Privacy (S&P '12)*. May 2012, pp. 553–567. DOI: [10.1109/SP.2012.44](https://doi.org/10.1109/SP.2012.44).
- [Bon+15] J. BONNEAU, C. HERLEY, P. C. VAN OORSCHOT, F. STAJANO. "Passwords and the Evolution of Imperfect Authentication." In: *Communications of the ACM* 58.7, June 2015, pp. 78–87. DOI: [10.1145/2699390](https://doi.org/10.1145/2699390).
- [Boo22] L. VAN DEN BOOGAARD. "User Understanding and User Acceptance of Biometric Authentication on Mobile Phones." Bachelor Thesis. Radboud University, Mar. 2022. URL: https://www.cs.ru.nl/bachelors-theses/2022/Leon_van_den_Boogaard___4590058___User_understanding_and_user_acceptance_of_biometric_authentication_on_mobile_phones.pdf (visited on May 20, 2026).

- [Bri26] V. BRIEGLEB. *Berlin Power Outage: How Mobile Communications and Internet Come Back Online*. heise online. Jan. 7, 2026. URL: <https://www.heise.de/en/news/Berlin-power-outage-How-mobile-communications-and-internet-come-back-online-11132151.html> (visited on May 20, 2026).
- [Bro96] J. BROOKE. "SUS: A "Quick and Dirty" Usability Scale." In: *Usability Evaluation in Industry*. London: CRC Press, 1996. DOI: [10.1201/9781498710411-35](https://doi.org/10.1201/9781498710411-35).
- [Bru+21] C. BRUNNER, U. GALLERSDÖRFER, F. KNIRSCH, D. ENGEL, F. MATTHES. "DID and VC: Untangling Decentralized Identifiers and Verifiable Credentials for the Web of Trust." In: *Proceedings of the 2020 3rd International Conference on Blockchain Technology and Applications (ICBTA '20)*. ACM, Mar. 2021, pp. 61–66. DOI: [10.1145/3446983.3446992](https://doi.org/10.1145/3446983.3446992).
- [BSI25] BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK. *Digitaler Verbraucherschutz: BSI-Jahresrückblick 2024*. Jahresrückblick BSI-DVS25/001. Bonn, 2025. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/DVS-Berichte/dvs-jahresrueckblick_2024.pdf?__blob=publicationFile&v=6 (visited on May 20, 2026).
- [BNe22] BUNDESNETZAGENTUR. *Resilience of Telecommunications Networks*. Strategy Paper. Federal Network Agency for Electricity, Gas, Telecommunications, Post and Railway, Aug. 2022. URL: https://www.bundesnetzagentur.de/DE/Fachthemen/Telekommunikation/Resilienz/Strategiepapier_Resilienz_eng.pdf?__blob=publicationFile&v=2 (visited on May 20, 2026).
- [BG25] A. BÜTTNER, N. GRUSCHKA. *Device-Bound vs. Synced Credentials: A Comparative Evaluation of Passkey Authentication*. Technical Report. 2025. DOI: [10.48550/arXiv.2501.07380](https://doi.org/10.48550/arXiv.2501.07380).
- [Cai+22] C. CAI, Z. CHEN, J. LUO, H. PU, M. HU, R. ZHENG. "Boosting Chirp Signal Based Aerial Acoustic Communication Under Dynamic Channel Conditions." In: *IEEE Transactions on Mobile Computing* 21.9, Sept. 2022, pp. 3110–3121. DOI: [10.1109/TMC.2021.3051665](https://doi.org/10.1109/TMC.2021.3051665).
- [Che+08] C.-H. O. CHEN, C.-W. CHEN, C. KUO, Y.-H. LAI, J. M. McCUNE, A. STUDER, A. PERRIG, B.-Y. YANG, T.-C. WU. "GAnGS: Gather, Authenticate 'n Group Securely." In: *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom '08)*. ACM, Sept. 14, 2008, pp. 92–103. DOI: [10.1145/1409944.1409957](https://doi.org/10.1145/1409944.1409957).
- [CG11] M. K. CHONG, H. GELLERSEN. "How Users Associate Wireless Devices." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, May 2011, pp. 1909–1918. DOI: [10.1145/1978942.1979219](https://doi.org/10.1145/1978942.1979219).
- [CG12] M. K. CHONG, H. GELLERSEN. "Usability Classification for Spontaneous Device Association." In: *Personal and Ubiquitous Computing* 16.1, Jan. 2012, pp. 77–89. DOI: [10.1007/s00779-011-0421-1](https://doi.org/10.1007/s00779-011-0421-1).
- [CG13] M. K. CHONG, H. W. GELLERSEN. "How Groups of Users Associate Wireless Devices." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, Apr. 2013, pp. 1559–1568. DOI: [10.1145/2470654.2466207](https://doi.org/10.1145/2470654.2466207).

- [CKG11] M. K. CHONG, F. KAWSAR, H. GELLERSEN. “Spatial Co-Location for Device Association: The Connected Object Way.” In: *Proceedings of the 2011 International Workshop on Networking and Object Memories for the Internet of Things (NoME-IoT ’11)*. ACM, Sept. 2011, pp. 21–26. DOI: [10.1145/2029932.2029941](https://doi.org/10.1145/2029932.2029941).
- [CMG14] M. K. CHONG, R. MAYRHOFER, H. GELLERSEN. “A Survey of User Interaction for Spontaneous Device Association.” In: *ACM Computing Surveys* 47.1, May 2014. DOI: [10.1145/2597768](https://doi.org/10.1145/2597768).
- [Cho25] T. CHOWDHURY. “Technological Enablers and Radio Network Planning for Sustainability Transition in Local 6G Networks.” Master’s Thesis. University of Oulu, June 2025. 64 pp. URL: <https://urn.fi/URN:NBN:fi:oulu-202506164542> (visited on May 20, 2026).
- [CPD19] S. CIOLINO, S. PARKIN, P. DUNPHY. “Of Two Minds About Two-Factor: Understanding Everyday FIDO U2F Usability Through Device Comparison and Experience Sampling.” In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS ’19)*. USENIX Association, Aug. 2019, pp. 339–356. URL: <https://www.usenix.org/conference/soups2019/presentation/ciolino> (visited on May 20, 2026).
- [Col+18] J. COLNAGO, S. DEVLIN, M. OATES, C. SWOOPES, L. BAUER, L. CRANOR, N. CHRISTIN. “‘It’s not actually that horrible’: Exploring Adoption of Two-Factor Authentication at a University.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI ’18)*. ACM, 2018. DOI: [10.1145/3173574.3174030](https://doi.org/10.1145/3173574.3174030).
- [CH06] L. K. COMFORT, T. W. HAASE. “Communication, Coherence, and Collective Action: The Impact of Hurricane Katrina on Communications Infrastructure.” In: *Public Works Management & Policy* 10.4, Apr. 2006, pp. 328–343. DOI: [10.1177/1087724X06289052](https://doi.org/10.1177/1087724X06289052).
- [Cop25] COPSONIC TECHNOLOGIES. *CopSonic: Ultrasonic Authentication*. 2025. URL: <https://www.copsonic.com/> (visited on May 20, 2026).
- [Cor+24] M. CORICI, H. ZOPE, H. BUHR, C. SCHEICH, T. MAGEDANZ. “Shortcuts: A Simple Mechanism for Reducing the Data Path Delay in Beyond 5G and 6G Networks.” In: *27th Conference on Innovation in Clouds, Internet and Networks (ICIN ’24)*. IEEE, 2024, pp. 47–51. DOI: [10.1109/ICIN60470.2024.10494443](https://doi.org/10.1109/ICIN60470.2024.10494443).
- [CUE] CUE AUDIO, INC. *CUE: About CUE*. URL: <https://www.connectwithcue.com/about> (visited on May 20, 2026).
- [Daf+25] A. DAFFALLA, A. BHATTACHARYA, J. WILDER, R. CHATTERJEE, N. DELL, R. BELLINI, T. RISTENPART. “A Framework for Abusability Analysis: The Case of Passkeys in Interpersonal Threat Models.” In: *34th USENIX Security Symposium (USENIX Security ’25)*. USENIX Association, 2025, pp. 7819–7838. URL: <https://www.usenix.org/conference/usenixsecurity25/presentation/daffalla> (visited on May 20, 2026).
- [DM21] I. DALLO, M. MARTI. “Why Should I Use a Multi-Hazard App? Assessing the Public’s Information Needs and App Feature Preferences in a Participatory Process.” In: *International Journal of Disaster Risk Reduction* 57, Apr. 2021. DOI: [10.1016/j.ijdr.2021.102197](https://doi.org/10.1016/j.ijdr.2021.102197).

- [Dan+24] F. DANG, X. SUN, K. LIU, X. DING, X. WANG, Y. LIU. “An Anonymous Authentication Scheme Based on Blind Signatures for the FIDO Protocol.” In: *2024 International Scientific and Technical Conference Modern Computer Network Technologies (MoNeTeC '24)*. Oct. 2024. DOI: [10.1109/MoNeTec60984.2024.10768120](https://doi.org/10.1109/MoNeTec60984.2024.10768120).
- [Das+14] A. DAS, J. BONNEAU, M. CAESAR, N. BORISOV, X. WANG. “The Tangled Web of Password Reuse.” In: *Proceedings of the Network and Distributed System Security Symposium (NDSS '14)*. Internet Society, 2014. DOI: [10.14722/ndss.2014.23357](https://doi.org/10.14722/ndss.2014.23357).
- [DDC18] S. DAS, A. DINGMAN, L. J. CAMP. “Why Johnny Doesn’t Use Two Factor: A Two-Phase Usability Study of the FIDO U2F Security Key.” In: *Financial Cryptography and Data Security (FC '18)*. 2018, pp. 160–179. DOI: [10.1007/978-3-662-58387-6_9](https://doi.org/10.1007/978-3-662-58387-6_9).
- [Das+18] S. DAS, G. RUSSO, A. C. DINGMAN, J. DEV, O. KENNY, L. J. CAMP. “A Qualitative Study on Usability and Acceptability of Yubico Security Key.” In: *Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust (STAST '17)*. ACM, 2018, pp. 28–39. DOI: [10.1145/3167996.3167997](https://doi.org/10.1145/3167996.3167997).
- [Dec+16] S. DECHAND, D. SCHÜRMANN, K. BUSSE, Y. ACAR, S. FAHL, M. SMITH. “An Empirical Study of Textual Key-Fingerprint Representations.” In: *25th USENIX Security Symposium (USENIX Security '16)*. USENIX Association, Aug. 2016, pp. 193–208. URL: <https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/dechand> (visited on May 20, 2026).
- [Dee+19] G. DEEPAK, A. LADAS, Y. A. SAMBO, H. PERVAIZ, C. POLITIS, M. A. IMRAN. “An Overview of Post-Disaster Emergency Communication Systems in the Future Networks.” In: *IEEE Wireless Communications 26.6*, 2019, pp. 132–139. DOI: [10.1109/MWC.2019.1800467](https://doi.org/10.1109/MWC.2019.1800467).
- [Dev26] A. DEVERIA. *Can I use WebAuthn?* 2026. URL: <https://caniuse.com/webauthn> (visited on May 20, 2026).
- [DTH06] R. DHAMIJA, J. D. TYGAR, M. HEARST. “Why Phishing Works.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, 2006, pp. 581–590. DOI: [10.1145/1124772.1124861](https://doi.org/10.1145/1124772.1124861).
- [DTO23] F. DI NOCERA, G. TEMPESTINI, M. ORSINI. “Usable Security: A Systematic Literature Review.” In: *Information 14.12*, Dec. 2023. DOI: [10.3390/info14120641](https://doi.org/10.3390/info14120641).
- [Doa+22] T. V. DOAN, Y. PSARAS, J. OTT, V. BAJPAI. “Toward Decentralized Cloud Storage With IPFS: Opportunities, Challenges, and Future Considerations.” In: *IEEE Internet Computing 26.6*, Nov. 2022, pp. 7–15. DOI: [10.1109/MIC.2022.3209804](https://doi.org/10.1109/MIC.2022.3209804).
- [Eas11] D. E. EASTLAKE 3RD. *Transport Layer Security (TLS) Extensions: Extension Definitions*. RFC 6066. Jan. 2011. DOI: [10.17487/RFC6066](https://doi.org/10.17487/RFC6066).
- [Ell07] C. ELLISON. *Ceremony Design and Analysis*. Technical Report. 2007. URL: <https://eprint.iacr.org/2007/399> (visited on May 20, 2026).

- [Far+13] M. FARB, Y.-H. LIN, T. H.-J. KIM, J. McCUNE, A. PERRIG. “SafeSlinger: Easy-to-Use and Secure Public-Key Exchange.” In: *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom '13)*. ACM, Sept. 2013, pp. 417–428. DOI: [10.1145/2500423.2500428](https://doi.org/10.1145/2500423.2500428).
- [Far+22] F. M. FARKE, L. LASSAK, J. PINTER, M. DÜRMUTH. “Exploring User Authentication with Windows Hello in a Small Business Environment.” In: *Eighteenth Symposium on Usable Privacy and Security (SOUPS '22)*. 2022, pp. 523–540. URL: <https://www.usenix.org/conference/soups2022/presentation/farke> (visited on May 20, 2026).
- [Far+20] F. M. FARKE, L. LORENZ, T. SCHNITZLER, P. MARKERT, M. DÜRMUTH. “‘You still use the password after all’ – Exploring FIDO2 Security Keys in a Small Company.” In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS '20)*. USENIX Association, Aug. 2020, pp. 19–35. URL: <https://www.usenix.org/conference/soups2020/presentation/farke> (visited on May 20, 2026).
- [Far+23] H. FARRUKH, M. OZMEN, F. ORS, B. CELIK. “One Key to Rule Them All: Secure Group Pairing for Heterogeneous IoT Devices.” In: *2023 IEEE Symposium on Security and Privacy (S&P '23)*. 2023, pp. 3026–3042. DOI: [10.1109/SP46215.2023.10179369](https://doi.org/10.1109/SP46215.2023.10179369).
- [FK23] M. FASSL, K. KROMBHOLZ. “Why I Can’t Authenticate — Understanding the Low Adoption of Authentication Ceremonies with Autoethnography.” In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, Apr. 2023. DOI: [10.1145/3544548.3581508](https://doi.org/10.1145/3544548.3581508).
- [Fed25] FEDERAL BUREAU OF INVESTIGATION. *Internet Crime Report 2024*. Annual Report. Apr. 23, 2025. URL: https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf (visited on May 20, 2026).
- [Fek21] A. FEKETE. “Motivation, Satisfaction, and Risks of Operational Forces and Helpers Regarding the 2021 and 2013 Flood Operations in Germany.” In: *Sustainability* 13.22, Nov. 2021. DOI: [10.3390/su132212587](https://doi.org/10.3390/su132212587).
- [FSK10] N. FERGUSON, B. SCHNEIER, T. KOHNO. *Cryptography Engineering: Design Principles and Practical Applications*. Wiley, 2010. ISBN: 978-0-470-47424-2.
- [FID17] FIDO ALLIANCE. *Universal 2nd Factor (U2F) Overview*. Ed. by S. SRINIVAS, D. BALFANZ, E. TIFFANY, A. CZESKIS. FIDO Alliance Proposed Standard. Apr. 11, 2017. URL: <https://fidoalliance.org/specs/fido-u2f-v1.2-ps-20170411/fido-u2f-overview-v1.2-ps-20170411.html>.
- [FID19] FIDO ALLIANCE. *Android Now FIDO2 Certified, Accelerating Global Migration Beyond Passwords*. Feb. 25, 2019. URL: <https://fidoalliance.org/android-now-fido2-certified-accelerating-global-migration-beyond-passwords/> (visited on May 20, 2026).
- [FID21] FIDO ALLIANCE. *Client to Authenticator Protocol (CTAP)*. Version 2.1. Ed. by J. BRADLEY, J. HODGES, M. B. JONES, A. KUMAR, R. LINDEMANN, J. VERREPT. FIDO Alliance Proposed Standard. June 15, 2021. URL: <https://fidoalliance.org/specs/fido-v2.1-ps-20210615/fido-client-to-authenticator-protocol-v2.1-ps-20210615.html>.

- [FID24a] FIDO ALLIANCE. *Synched Passkey Deployment: Emerging Practices for Consumer Use Cases*. White Paper. FIDO Alliance, Jan. 2024. URL: https://fidoalliance.org/wp-content/uploads/2024/05/Synched-Passkey-Deployment_-Emerging-Practices-for-Consumer-Use-Cases_2024-Final.pdf (visited on May 20, 2026).
- [FID25] FIDO ALLIANCE. *FIDO Alliance Champions Widespread Passkey Adoption and a Passwordless Future on World Passkey Day 2025*. May 1, 2025. URL: <https://fidoalliance.org/fido-alliance-champions-widespread-passkey-adoption-and-a-passwordless-future-on-world-passkey-day-2025/> (visited on May 20, 2026).
- [Fis+25] K. FISCHER, M. KEIL, A. BUCKMANN, M. A. SASSE. “If You Want to Encrypt It Really, Really Hardcore...”: User Perceptions of Key Transparency in Whats-App.” In: *Proceedings on Privacy Enhancing Technologies (PoPETs ’25)* 4, 2025, pp. 1039–1054. DOI: [10.56553/popets-2025-0170](https://doi.org/10.56553/popets-2025-0170).
- [FBF22] D. FISCHER-PRESSLER, D. BONARETTI, K. FISCHBACH. “A Protection-Motivation Perspective to Explain Intention to Use and Continue to Use Mobile Warning Systems.” In: *Business & Information Systems Engineering* 64, 2022, pp. 167–182. DOI: [10.1007/s12599-021-00704-0](https://doi.org/10.1007/s12599-021-00704-0).
- [FH07] D. FLORENCIO, C. HERLEY. “A Large-Scale Study of Web Password Habits.” In: *Proceedings of the 16th International Conference on World Wide Web (WWW ’07)*. ACM, May 2007, pp. 657–666. DOI: [10.1145/1242572.1242661](https://doi.org/10.1145/1242572.1242661).
- [Fom+18] M. FOMICHEV, F. ALVAREZ, D. STEINMETZER, P. GARDNER-STEPHEN, M. HOLLICK. “Survey and Systematization of Secure Device Pairing.” In: *IEEE Communications Surveys & Tutorials* 20.1, 2018, pp. 517–550. DOI: [10.1109/COMST.2017.2748278](https://doi.org/10.1109/COMST.2017.2748278).
- [Fom+19] M. FOMICHEV, M. MAASS, L. ALMON, A. MOLINA, M. HOLLICK. “Perils of Zero-Interaction Security in the Internet of Things.” In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3.1, Mar. 2019. DOI: [10.1145/3314397](https://doi.org/10.1145/3314397).
- [FAW19] T. FRANKE, C. ATTIG, D. WESSEL. “A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale.” In: *International Journal of Human–Computer Interaction* 35.6, 2019, pp. 456–467. DOI: [10.1080/10447318.2018.1456150](https://doi.org/10.1080/10447318.2018.1456150).
- [Fry+20] N. FRYMANN, D. GARDHAM, F. KIEFER, E. LUNDBERG, M. MANULIS, D. NILSSON. “Asynchronous Remote Key Generation: An Analysis of Yubico’s Proposal for W3C WebAuthn.” In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS ’20)*. ACM, Nov. 2, 2020, pp. 939–954. DOI: [10.1145/3372297.3417292](https://doi.org/10.1145/3372297.3417292).
- [FGM22] N. FRYMANN, D. GARDHAM, M. MANULIS. “Unlinkable Delegation of WebAuthn Credentials.” In: *Computer Security (ESORICS ’22)*. Springer, Sept. 2022, pp. 125–144. DOI: [10.1007/978-3-031-17143-7_7](https://doi.org/10.1007/978-3-031-17143-7_7).

- [Gau+25] A. GAUTAM, T. YADAV, G. SMITH, K. SEAMONS, S. RUOTI. “Passwords and FIDO2 Are Meant To Be Secret: A Practical Secure Authentication Channel for Web Browsers.” In: *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS '25)*. ACM, Nov. 2025, pp. 1859–1873. doi: [10.1145/3719027.3765195](https://doi.org/10.1145/3719027.3765195).
- [GB00] V. GERASIMOV, W. BENDER. “Things That Talk: Using Sound for Device-to-Device and Device-to-Human Communication.” In: *IBM Systems Journal* 39.3.4, 2000, pp. 530–546–530–546. doi: [10.1147/sj.393.0530](https://doi.org/10.1147/sj.393.0530).
- [Ger26] G. GERGANOV. *ggwave. Tiny Data-Over-Sound Library*. GitHub Repository, Mar. 21, 2026. url: <https://github.com/ggerganov/ggwave> (visited on May 20, 2026).
- [Ger+23] E. GERLITZ, M. HÄRING, C. T. MÄDLER, M. SMITH, C. TIEFENAU. “Adventures in Recovery Land: Testing the Account Recovery of Popular Websites When the Second Factor Is Lost.” In: *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. 2023, pp. 227–243. isbn: 978-1-939133-36-6. url: <https://www.usenix.org/conference/soups2023/presentation/gerlitz-adventures> (visited on Mar. 14, 2026).
- [Get+18] P. GETREUER, C. GNEGY, R. F. LYON, R. A. SAUROUS. “Ultrasonic Communication Using Consumer Hardware.” In: *IEEE Transactions on Multimedia* 20.6, 2018, pp. 1277–1290. doi: [10.1109/TMM.2017.2766049](https://doi.org/10.1109/TMM.2017.2766049).
- [GLL17] N. GHOSE, L. LAZOS, M. LI. “HELP: Helper-Enabled in-Band Device Pairing Resistant against Signal Cancellation.” In: *26th USENIX Security Symposium (USENIX Security '17)*. USENIX Association, Aug. 2017, pp. 433–450. url: <https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/ghose> (visited on May 20, 2026).
- [Gil25] C. GILSENAN. “Exploring the Security and Privacy Impacts of Using 2FA Apps.” PhD Thesis. EECS Department, University of California, Berkeley, May 2025. url: <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2025/EECS-2025-49.html> (visited on May 20, 2026).
- [Gon+17] M. A. GONÇALVES, R. V. LOUREDO, M. P. TCHEOU, L. LOVISOLO. “Acoustic Modem Using OFDM and Least-Square Channel Estimation.” In: *2017 IEEE 9th Latin-American Conference on Communications (LATINCOM '17)*. Nov. 2017, pp. 1–5. doi: [10.1109/LATINCOM.2017.8240178](https://doi.org/10.1109/LATINCOM.2017.8240178).
- [Goo18] D. GOODIN. *Police Decrypt 258,000 Messages after Breaking Pricey IronChat Crypto App*. Ars Technica. Nov. 7, 2018. url: <https://arstechnica.com/information-technology/2018/11/police-decrypt-258000-messages-after-breaking-pricey-ironchat-crypto-app/> (visited on May 20, 2026).
- [Goo+09] M. T. GOODRICH, M. SIRIVIANOS, J. SOLIS, C. SORIENTE, G. TSUDIK, E. UZUN. “Using Audio in Secure Device Pairing.” In: *International Journal of Security and Networks* 4.1–2, Feb. 24, 2009, pp. 57–68. doi: [10.1504/IJSN.2009.023426](https://doi.org/10.1504/IJSN.2009.023426).
- [Goo+06] M. T. GOODRICH, M. SIRIVIANOS, J. SOLIS, G. TSUDIK, E. UZUN. “Loud and Clear: Human-Verifiable Authentication Based on Audio.” In: *26th IEEE International Conference on Distributed Computing Systems (ICDCS '06)*. IEEE, 2006. doi: [10.1109/icdcs.2006.52](https://doi.org/10.1109/icdcs.2006.52).

- [Goo26] GOOGLE. *About Credential Manager*. Android Developers. May 13, 2026. URL: <https://developer.android.com/identity/credential-manager> (visited on May 20, 2026).
- [Gooa] GOOGLE. *OpenSK. Open-Source Implementation for Security Keys Written in Rust*. GitHub Repository. URL: <https://github.com/google/OpenSK> (visited on May 20, 2026).
- [Goob] GOOGLE. *Use Passwords & Passkeys Across Your Devices*. Google Account Help. URL: <https://support.google.com/accounts/answer/6197437> (visited on May 20, 2026).
- [Grø+20] J. E. GRØNBÆK, M. S. KNUDSEN, K. O'HARA, P. G. KROGH, J. VERMEULEN, M. G. PETERSEN. "Proxemics Beyond Proximity: Designing for Flexible Social Interaction Through Cross-Device Interaction." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, Apr. 2020, pp. 1–14. DOI: [10.1145/3313831.3376379](https://doi.org/10.1145/3313831.3376379).
- [Gut+22] M. GUTFLEISCH, J. H. KLEMMER, N. BUSCH, Y. ACAR, M. A. SASSE, S. FAHL. "How Does Usable Security (Not) End Up in Software Products? Results From a Qualitative Interview Study." In: *2022 IEEE Symposium on Security and Privacy (S&P '22)*. May 2022, pp. 893–910. DOI: [10.1109/SP46214.2022.9833756](https://doi.org/10.1109/SP46214.2022.9833756).
- [Hac+22] T. HACKENJOS, B. WAGNER, J. HERR, J. RILL, M. WEHMER, N. GOERKE, I. BAUMGART. *FIDO2 With Two Displays-Or How to Protect Security-Critical Web Transactions Against Malware Attacks*. Technical Report. July 2022. DOI: [10.48550/arXiv.2206.13358](https://doi.org/10.48550/arXiv.2206.13358).
- [Hal+08] D. HALPERIN, T. S. HEYDT-BENJAMIN, B. RANSFORD, S. S. CLARK, B. DEFEND, W. MORGAN, K. FU, T. KOHNO, W. H. MAISEL. "Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses." In: *2008 IEEE Symposium on Security and Privacy S&P '08*. May 2008, pp. 129–142. DOI: [10.1109/SP.2008.31](https://doi.org/10.1109/SP.2008.31).
- [HKR22] J. HAUNSCHILD, M.-A. KAUFHOLD, C. REUTER. "Perceptions and Use of Warning Apps – Did Recent Crises Lead to Changes in Germany?" In: *Proceedings of Mensch und Computer 2022 (MuC '22)*. ACM, 2022, pp. 25–40. DOI: [10.1145/3543758.3543770](https://doi.org/10.1145/3543758.3543770).
- [HPR23] J. HAUNSCHILD, S. PAULI, C. REUTER. "Preparedness Nudging for Warning Apps? A Mixed-Method Study Investigating Popularity and Effects of Preparedness Alerts in Warning Apps." In: *International Journal on Human-Computer Studies (IJHCS)* 172, 2023. DOI: [10.1016/j.ijhcs.2023.102995](https://doi.org/10.1016/j.ijhcs.2023.102995).
- [Hau+19] M. HAUS, A. Y. DING, Q. WANG, J. TOIVONEN, L. TONETTO, S. TARKOMA, J. OTT. "Enhancing Indoor IoT Communication with Visible Light and Ultrasound." In: *2019 IEEE International Conference on Communications (ICC '19)*. May 2019, pp. 1–6. DOI: [10.1109/ICC.2019.8762001](https://doi.org/10.1109/ICC.2019.8762001).
- [HB18] Y. HE, J. BIAN. "Novel Ultrasonic Broadcast Communication System." In: *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP '18)*. 2018, pp. 1–7. DOI: [10.1109/WCSP.2018.8555902](https://doi.org/10.1109/WCSP.2018.8555902).

- [HL16] A. HERZBERG, H. LEIBOWITZ. “Can Johnny Finally Encrypt? Evaluating E2E-encryption in Popular IM Applications.” In: *Proceedings of the 6th Workshop on Socio-Technical Aspects in Security and Trust (STAST ’16)*. ACM, Dec. 5, 2016, pp. 17–28. DOI: [10.1145/3046055.3046059](https://doi.org/10.1145/3046055.3046059).
- [Her+21] A. HERZBERG, H. LEIBOWITZ, K. SEAMONS, E. VAZIRIPOUR, J. WU, D. ZAPPALA. “Secure Messaging Authentication Ceremonies Are Broken.” In: *IEEE Security & Privacy* 19.2, Mar. 2021, pp. 29–37. DOI: [10.1109/MSEC.2020.3039727](https://doi.org/10.1109/MSEC.2020.3039727).
- [Höc+23] J. HÖCHST, L. BAUMGÄRTNER, F. KUNTKE, A. PENNING, A. STERZ, M. SOMMER, B. FREISLEBEN. “Mobile Device-to-Device Communication for Crisis Scenarios Using Low-Cost LoRa Modems.” In: *Disaster Management and Information Technology: Professional Response and Recovery Management in the Age of Disasters*. Cham: Springer International Publishing, Apr. 2023, pp. 235–268. DOI: [10.1007/978-3-031-20939-0_12](https://doi.org/10.1007/978-3-031-20939-0_12).
- [HMJ20] J. HODGES, G. MANDYAM, M. B. JONES. *Registries for Web Authentication (WebAuthn)*. RFC 8809. Aug. 2020. DOI: [10.17487/RFC8809](https://doi.org/10.17487/RFC8809).
- [Hol24] K. HOLDSWORTH. *Passkeys on Windows: Authenticate Seamlessly with Passkey Providers*. Windows Developer Blog. Oct. 8, 2024. URL: <https://blogs.windows.com/windowsdeveloper/2024/10/08/passkeys-on-windows-authenticate-seamlessly-with-passkey-providers/> (visited on May 20, 2026).
- [Hol25] K. HOLDSWORTH. *Windows 11 Expands Passkey Manager Support*. Windows IT Pro Blog. Nov. 11, 2025. URL: <https://techcommunity.microsoft.com/blog/windows-itpro-blog/windows-11-expands-passkey-manager-support/4467572> (visited on May 20, 2026).
- [Höl+24] S. HÖLTERVENNHOF, N. WÖHLER, A. MÖHLE, M. OLTROGGE, Y. ACAR, O. WIESE, S. FAHL. “A Mixed-Methods Study on User Experiences and Challenges of Recovery Codes for an End-to-End Encrypted Service.” In: *33rd USENIX Security Symposium (USENIX Security ’24)*. USENIX Association, 2024, pp. 7267–7284. URL: <https://www.usenix.org/conference/usenixsecurity24/presentation/h%C3%B6ltervennhoff> (visited on May 20, 2026).
- [Hol+25] J.-U. HOLTGRAVE, S. KLIVAN, K. MARKY, S. FAHL. “A Qualitative Study of Adoption Barriers and Challenges for Passwordless Authentication in German Public Administrations.” In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI ’25)*. ACM, Apr. 2025. DOI: [10.1145/3706598.3713252](https://doi.org/10.1145/3706598.3713252).
- [Hor+20] Y. HORNYCH, J. C. TOLEDO, B. WANG, W.-J. YI, J. SANIE. “Near-Ultrasonic Communications for IoT Applications Using Android Smartphone.” In: *2020 IEEE International Conference on Electro Information Technology (EIT ’20)*. July 2020, pp. 407–410. DOI: [10.1109/EIT48999.2020.9208265](https://doi.org/10.1109/EIT48999.2020.9208265).
- [HLG13] Y. HOU, M. LI, J. D. GUTTMAN. “Chorus: Scalable In-Band Trust Establishment for Multiple Constrained Devices over the Insecure Wireless Channel.” In: *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec ’13)*. ACM, 2013, pp. 167–178. DOI: [10.1145/2462096.2462124](https://doi.org/10.1145/2462096.2462124).

- [Hun26] T. HUNT. *Have I Been Pwned: Who's Been Pwned*. 2026. URL: <https://haveibeenpwned.com/PwnedWebsites> (visited on May 20, 2026).
- [Hyl+24] C. D. HYLENDER, P. LANGLOIS, A. PINTO, S. WIDUP. *2024 Data Breach Investigations Report*. Report. Verizon, 2024. URL: <https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf> (visited on May 20, 2026).
- [IAN] IANA. *Web Authentication (WebAuthn) Registries*. URL: <https://www.iana.org/assignments/webauthn/webauthn.xhtml> (visited on May 20, 2026).
- [Isl25] M. ISLAM. "Safeguarding Online Authentication Systems From Attacks." PhD Thesis. University of Wisconsin-Madison, 2025. URL: <https://asset.library.wisc.edu/1711.dl/F56XAC6IRAKUX8B/R/file-5ebad.pdf> (visited on May 20, 2026).
- [Isl+25] M. ISLAM, S. S. ARORA, R. CHATTERJEE, K. C. WANG. "Detecting Compromise of Passkey Storage on the Cloud." In: *34th USENIX Security Symposium (USENIX Security '25)*. USENIX Association, 2025, pp. 7743–7762. URL: <https://www.usenix.org/conference/usenixsecurity25/presentation/islam> (visited on May 20, 2026).
- [IWS04] B. IVES, K. R. WALSH, H. SCHNEIDER. "The Domino Effect of Password Reuse." In: *Communications of the ACM* 47.4, Apr. 2004, pp. 75–78. DOI: [10.1145/975817.975820](https://doi.org/10.1145/975817.975820).
- [JH26] L. JANZEN, M. HOLICK. "Making Cellular Networks Crisis-Proof: Towards Island-Ready, Resilient-by-Design 6G Communication Networks." In: *Computer Communications* 247, Feb. 2026. DOI: [10.1016/j.comcom.2025.108407](https://doi.org/10.1016/j.comcom.2025.108407).
- [JKL16] K. M. JEON, H. K. KIM, M. J. LEE. "Noncoherent Low-Frequency Ultrasonic Communication System with Optimum Symbol Length." In: *International Journal of Distributed Sensor Networks* 12.5, May 2016. DOI: [10.1155/2016/9713180](https://doi.org/10.1155/2016/9713180).
- [Jin+22] M. JIN, Y. HE, Y. LIU, X. WANG. "Furtively Connecting IoT Devices with Acoustic Noise." In: *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN '22)*. May 2022, pp. 195–207. DOI: [10.1109/IPSN54338.2022.00023](https://doi.org/10.1109/IPSN54338.2022.00023).
- [Jok+15] T. JOKELA, M. K. CHONG, A. LUCERO, H. GELLERSEN. "Connecting Devices for Collaborative Interactions." In: *Interactions* 22.4, June 2015, pp. 39–43. DOI: [10.1145/2776887](https://doi.org/10.1145/2776887).
- [JL14] T. JOKELA, A. LUCERO. "FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch." In: *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, Sept. 2014, pp. 369–378. DOI: [10.1145/2628363.2628376](https://doi.org/10.1145/2628363.2628376).
- [JRV16] T. JOKELA, P. P. REZAEI, K. VÄÄNÄNEN. "Natural Group Binding and Cross-Display Object Movement Methods for Wearable Devices." In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, Sept. 2016, pp. 206–216. DOI: [10.1145/2935334.2935346](https://doi.org/10.1145/2935334.2935346).

- [Jon22] D. JONES. *How Near-Ultrasonic Audio Adds Spatial Awareness to the Sonos System*. Sonos Tech Blog. Jan. 27, 2022. URL: <https://tech-blog.sonos.com/posts/how-near-ultrasonic-audio-adds-spatial-awareness-to-the-sonos-system/> (visited on May 20, 2026).
- [JSS25] M. JUBUR, P. SHRESTHA, N. SAXENA. "An In-Depth Analysis of Password Managers and Two-Factor Authentication Tools." In: *ACM Computing Surveys* 57.5, Jan. 2025. DOI: [10.1145/3711117](https://doi.org/10.1145/3711117).
- [Ka+16] S. KA, T. H. KIM, J. Y. HA, S. H. LIM, S. C. SHIN, J. W. CHOI, C. KWAK, S. CHOI. "Near-Ultrasound Communication for TV's 2Nd Screen Services." In: *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, May 2016, pp. 42–54. DOI: [10.1145/2973750.2973774](https://doi.org/10.1145/2973750.2973774).
- [KFR09] R. KAINDA, I. FLECHAS, A. W. ROSCOE. "Usability and Security of Out-of-Band Channels in Secure Device Pairing Protocols." In: *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, July 2009. DOI: [10.1145/1572532.1572547](https://doi.org/10.1145/1572532.1572547).
- [Kan18] V. A. KANADE. "Audio-Internet [A-I]: A Novel Technique for Streaming [Bits-over-Sound]: SoundData." In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI '18)*. Sept. 2018, pp. 1690–1694. DOI: [10.1109/ICACCI.2018.8554953](https://doi.org/10.1109/ICACCI.2018.8554953).
- [KR16] M.-A. KAUFHOLD, C. REUTER. "The Self-Organization of Digital Volunteers across Social Media: The Case of the 2013 European Floods in Germany." In: *Journal of Homeland Security and Emergency Management* 13.1, Apr. 2016, pp. 137–166. DOI: [10.1515/jhsem-2015-0063](https://doi.org/10.1515/jhsem-2015-0063).
- [Kei+24] N. KEIZER, O. ASCIGIL, M. KRÓL, D. KUTSCHER, G. PAVLOU. "A Survey on Content Retrieval on the Decentralised Web." In: *ACM Computing Surveys* 56.8, Apr. 2024. DOI: [10.1145/3649132](https://doi.org/10.1145/3649132).
- [Kep+23] M. KEPKOWSKI, M. MACHULAK, I. WOOD, D. KAAFAR. "Challenges with Passwordless FIDO2 in an Enterprise Setting: A Usability Study." In: *2023 IEEE Secure Development Conference (SecDev)*. Oct. 2023, pp. 37–48. DOI: [10.1109/SecDev56634.2023.00017](https://doi.org/10.1109/SecDev56634.2023.00017).
- [KML19] S. KIM, H. MUN, Y. LEE. "A Data-Over-Sound Application: Attendance Book." In: *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS '19)*. Sept. 2019. DOI: [10.23919/APNOMS.2019.8892996](https://doi.org/10.23919/APNOMS.2019.8892996).
- [KS92] J. J. KISTLER, M. SATYANARAYANAN. "Disconnected Operation in the Coda File System." In: *ACM Transactions on Computer Systems (TOCS)* 10.1, Feb. 1992, pp. 3–25. DOI: [10.1145/146941.146942](https://doi.org/10.1145/146941.146942).
- [Kle+19] M. KLEPPMANN, A. WIGGINS, P. VAN HARDENBERG, M. McGRANAGHAN. "Local-First Software: You Own Your Data, in Spite of the Cloud." In: *Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (Onward! 2019)*. ACM, Oct. 2019, pp. 154–178. DOI: [10.1145/3359591.3359737](https://doi.org/10.1145/3359591.3359737).

- [Kli+20] E. KLIEME, J. WILKE, N. VAN DORNICK, C. MEINEL. “FIDOnuous: A FIDO2/Web-Authn Extension to Support Continuous Web Authentication.” In: *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '20)*. Dec. 2020, pp. 1857–1867. DOI: [10.1109/TrustCom50675.2020.00254](https://doi.org/10.1109/TrustCom50675.2020.00254).
- [Kli+23] S. KLIVAN, S. HÖLTERVENNHOF, N. HUAMAN, A. KRAUSE, L. SIMKO, Y. ACAR, S. FAHL. ““We’ve Disabled MFA for You”: An Evaluation of the Security and Usability of Multi-Factor Authentication Recovery Deployments.” In: *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*. ACM, Nov. 2023, pp. 3138–3152. DOI: [10.1145/3576915.3623180](https://doi.org/10.1145/3576915.3623180).
- [Kli+24] S. KLIVAN, S. HÖLTERVENNHOF, R. PANSKUS, K. MARKY, S. FAHL. “Everyone for Themselves? A Qualitative Study about Individual Security Setups of Open Source Software Contributors.” In: *2024 IEEE Symposium on Security and Privacy (S&P '24)*. May 2024, pp. 1065–1082. DOI: [10.1109/SP54263.2024.00214](https://doi.org/10.1109/SP54263.2024.00214).
- [Kok+22] E. E. KOKS, K. C. H. VAN GINKEL, M. J. E. VAN MARLE, A. LEMNITZER. “Brief Communication: Critical Infrastructure Impacts of the 2021 Mid-July Western European Flood Event.” In: *Natural Hazards and Earth System Sciences* 22.12, Nov. 2022, pp. 3831–3838. ISSN: 1561-8633. DOI: [10.5194/nhess-22-3831-2022](https://doi.org/10.5194/nhess-22-3831-2022).
- [Kra+10] C. KRAY, D. NESBITT, J. DAWSON, M. ROHS. “User-Defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops.” In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, Sept. 2010, pp. 239–248. DOI: [10.1145/1851600.1851640](https://doi.org/10.1145/1851600.1851640).
- [Kre21] M. KREICHGAUER. *Intent to Deprecate and Remove: U2F API (Cryptotoken)*. blink-dev Google Group. June 11, 2021. URL: https://groups.google.com/a/chromium.org/g/blink-dev/c/xHC3AtU_65A (visited on May 20, 2026).
- [Kuc23] D. KUCHHAL. “Building Trust in the Online Ecosystem Through Empirical Evaluations of Web Security and Privacy Concerns.” PhD Thesis. Georgia Institute of Technology, July 2023. URL: <https://hdl.handle.net/1853/72724> (visited on May 20, 2026).
- [Kuc+23] D. KUCHHAL, M. SAAD, A. OEST, F. LI. “Evaluating the Security Posture of Real-World FIDO2 Deployments.” In: *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*. ACM, Nov. 2023, pp. 2381–2395. DOI: [10.1145/3576915.3623063](https://doi.org/10.1145/3576915.3623063).
- [Kum+09] A. KUMAR, N. SAXENA, G. TSUDIK, E. UZUN. “A Comparative Study of Secure Device Pairing Methods.” In: *Pervasive and Mobile Computing (PerCom '09)* 5.6, Dec. 2009, pp. 734–749. DOI: [10.1016/j.pmcj.2009.07.008](https://doi.org/10.1016/j.pmcj.2009.07.008).
- [Kum13] G. KUMPARAK. *SlickLogin Aims To Kill The Password By Singing A Silent Song To Your Smartphone*. TechCrunch. Sept. 9, 2013. URL: <https://techcrunch.com/2013/09/09/slicklogin-wants-to-kill-the-password-by-singing-a-silent-song-to-your-smartphone/> (visited on May 22, 2026).

- [Kun+21] J. KUNKE, S. WIEFLING, M. ULLMANN, L. LO IACONO. "Evaluation of Account Recovery Strategies with FIDO2-based Passwordless Authentication." In: *Open Identity Summit 2021*. Gesellschaft für Informatik e.V., 2021, pp. 59–70. URL: <https://dl.gi.de/handle/20.500.12116/36502> (visited on May 22, 2026).
- [Kuo+07] C. KUO, M. LUK, R. NEGI, A. PERRIG. "Message-in-a-Bottle: User-Friendly and Secure Key Deployment for Sensor Nodes." In: *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys '07)*. ACM, Nov. 2007, pp. 233–246. DOI: [10.1145/1322263.1322286](https://doi.org/10.1145/1322263.1322286).
- [KSP08] C. KUO, A. STUDER, A. PERRIG. "Mind Your Manners: Socially Appropriate Wireless Key Establishment for Groups." In: *Proceedings of the First ACM Conference on Wireless Network Security (WiSec '08)*. ACM, Mar. 2008, pp. 125–130. DOI: [10.1145/1352533.1352553](https://doi.org/10.1145/1352533.1352553).
- [Kwa+19] C. KWAK, S. KIM, S. KA, J. LEE, S. CHOI. "No Entry: Anti-Noise Energy Detector for Chirp-Based Acoustic Communication." In: *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON '19)*. IEEE, 2019. DOI: [10.1109/SAHCN.2019.8824986](https://doi.org/10.1109/SAHCN.2019.8824986).
- [Lan+17] J. LANG, A. CZESKIS, D. BALFANZ, M. SCHILDER, S. SRINIVAS. "Security Keys: Practical Cryptographic Second Factors for the Modern Web." In: *International Conference on Financial Cryptography and Data Security (FC '16)*. Springer, May 2017. DOI: [10.1007/978-3-662-54970-4_25](https://doi.org/10.1007/978-3-662-54970-4_25).
- [Lan20] A. LANGLEY. *Comment 4 on Chromium Issue 40701687: Support WebAuthn uvi & uvm Extension*. Chromium Issue Tracker, May 22, 2020. URL: <https://issues.chromium.org/issues/40701687#comment4> (visited on May 20, 2026).
- [Las+21] L. LASSAK, A. HILDEBRANDT, M. GOLLA, B. UR. "'It's Stored, Hopefully, on an Encrypted Server': Mitigating Users' Misconceptions About FIDO2 Biometric WebAuthn." In: *30th USENIX Security Symposium (USENIX Security '21)*. USENIX Association, Aug. 2021, pp. 91–108. URL: <https://www.usenix.org/conference/usenixsecurity21/presentation/lassak>.
- [Las+24] L. LASSAK, E. PAN, B. UR, M. GOLLA. "Why Aren't We Using Passkeys? Obstacles Companies Face Deploying FIDO2 Passwordless Authentication." In: *33rd USENIX Security Symposium (USENIX Security '24)*. USENIX Association, Aug. 2024, pp. 7231–7248. URL: <https://www.usenix.org/conference/usenixsecurity24/presentation/lassak> (visited on May 22, 2026).
- [LP08] S. LAUR, S. PASINI. "SAS-Based Group Authentication and Key Agreement Protocols." In: *Public Key Cryptography (PKC '08)*. Springer, 2008, pp. 197–213. DOI: [10.1007/978-3-540-78440-1_12](https://doi.org/10.1007/978-3-540-78440-1_12).
- [Law+11] Y. W. LAW, G. MONIAVA, Z. GONG, P. HARTEL, M. PALANISWAMI. "KALWEN: A New Practical and Interoperable Key Management Scheme for Body Sensor Networks." In: *Security and Communication Networks* 4.11, Nov. 2011, pp. 1309–1329. DOI: [10.1002/sec.256](https://doi.org/10.1002/sec.256).

- [Lee+15] H. LEE, T. H. KIM, J. W. CHOI, S. CHOI. "Chirp Signal-Based Aerial Acoustic Communication for Smart Devices." In: *2015 IEEE Conference on Computer Communications (INFOCOM '15)*. IEEE, 2015, pp. 2407–2415. doi: [10.1109/INFOCOM.2015.7218629](https://doi.org/10.1109/INFOCOM.2015.7218629).
- [Lee+25] J. LEE, M. KANG, S. LEE, H.-K. CHOI. "False Promises of Passwordless: Defeating Windows Hello through TPM Misuses." In: *2025 IEEE Annual Computer Security Applications Conference (ACSAC '25)*. Dec. 2025, pp. 308–321. doi: [10.1109/ACSAC67867.2025.00037](https://doi.org/10.1109/ACSAC67867.2025.00037).
- [Lee+20] J. LEE, C. KWAK, S. KIM, S. BAHK. "Reliable and Low-Complexity Chirp Spread Spectrum-Based Aerial Acoustic Communication." In: *IEEE Access* 8, 2020, pp. 151589–151601. doi: [10.1109/ACCESS.2020.3017097](https://doi.org/10.1109/ACCESS.2020.3017097).
- [LR18] I. LEVY, C. ROBINSON. *Principles for a More Informed Exceptional Access Debate*. The Lawfare Institute. Nov. 29, 2018. URL: <https://www.lawfaremedia.org/article/principles-more-informed-exceptional-access-debate> (visited on May 20, 2026).
- [Li+20] X. LI, Q. ZENG, L. LUO, T. LUO. "T2Pair: Secure and Usable Pairing for Heterogeneous IoT Devices." In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. ACM, Nov. 2020, pp. 309–323. doi: [10.1145/3372297.3417286](https://doi.org/10.1145/3372297.3417286).
- [LHC15] M.-C. LIN, F.-Y. HUANG, T.-D. CHUEH. "A-NFC: Two-way Near-Field Communications (NFC) via Inaudible Acoustics." In: *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA '15)*. July 2015. doi: [10.1109/IISA.2015.7388085](https://doi.org/10.1109/IISA.2015.7388085).
- [LYL17] Q. LIN, L. YANG, Y. LIU. "TagScreen: Synchronizing Social Televisions through Hidden Sound Markers." In: *IEEE Conference on Computer Communications (INFOCOM '17)*. May 2017. doi: [10.1109/INFOCOM.2017.8057170](https://doi.org/10.1109/INFOCOM.2017.8057170).
- [Lin+09] Y.-H. LIN, A. STUDER, H.-C. HSIAO, J. M. McCUNE, K.-H. WANG, M. KROHN, P.-L. LIN, A. PERRIG, H.-M. SUN, B.-Y. YANG. "SPATE: Small-Group PKI-less Authenticated Trust Establishment." In: *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*. ACM, June 2009. doi: [10.1145/1555816.1555818](https://doi.org/10.1145/1555816.1555818).
- [LB24] F. LINKER, D. BASIN. "SOAP: A Social Authentication Protocol." In: *33rd USENIX Security Symposium (USENIX Security '24)*. 2024, pp. 3223–3240. URL: <https://www.usenix.org/conference/usenixsecurity24/presentation/linker> (visited on May 20, 2026).
- [LIS] LISNR, INC. *Ultrasonic Proximity SDK*. URL: <https://lisnr.com/radius-ultrasonic-sdk-3/> (visited on May 22, 2026).
- [LA01] C. V. LOPES, P. M. Q. AGUIAR. "Aerial Acoustic Communications." In: *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*. Oct. 2001, pp. 219–222. doi: [10.1109/ASPAA.2001.969582](https://doi.org/10.1109/ASPAA.2001.969582).
- [LA03] C. V. LOPES, P. M. Q. AGUIAR. "Acoustic Modems for Ubiquitous Computing." In: *IEEE Pervasive Computing* 2.3, 2003, pp. 62–71. doi: [10.1109/MPRV.2003.1228528](https://doi.org/10.1109/MPRV.2003.1228528).

- [LN23] E. LUNDBERG, D. NILSSON. *webauthn-recovery-extension: Asynchronous Delegated Key Generation Without Shared Secrets*. GitHub Repository, Nov. 24, 2023. URL: <https://github.com/Yubico/webauthn-recovery-extension> (visited on May 20, 2026).
- [LBB23] S. G. LYASTANI, S. BUGIEL, M. BACKES. “A Systematic Study of the Consistency of Two-Factor Authentication User Journeys on Top-Ranked Websites.” In: *Proceedings of the Network and Distributed System Security Symposium (NDSS ’23)*. Internet Society, Feb. 2023. DOI: [10.14722/ndss.2023.23362](https://doi.org/10.14722/ndss.2023.23362).
- [Lya+20] S. G. LYASTANI, M. SCHILLING, M. NEUMAYR, M. BACKES, S. BUGIEL. “Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication.” In: *2020 IEEE Symposium on Security and Privacy (S&P ’20)*. IEEE, May 2020, pp. 842–859. DOI: [10.1109/SP40000.2020.00047](https://doi.org/10.1109/SP40000.2020.00047).
- [MSS03] A. MADHAVAPEDDY, D. SCOTT, R. SHARP. “Context-Aware Computing with Sound.” In: *Ubiquitous Computing (UbiComp 2003)*. Springer, 2003, pp. 315–332. DOI: [10.1007/978-3-540-39653-6_25](https://doi.org/10.1007/978-3-540-39653-6_25).
- [Mad+05] A. MADHAVAPEDDY, R. SHARP, D. SCOTT, A. TSE. “Audio Networking: The Forgotten Wireless Technology.” In: *IEEE Pervasive Computing* 4.3, 2005, pp. 55–60. DOI: [10.1109/MPRV.2005.50](https://doi.org/10.1109/MPRV.2005.50).
- [MAI25] A. MAHFOUZ, R. ABDULGHAFOR, A. A. K. ISMAEEL. “Passkeys in Practice: An Empirical Evaluation of FIDO2/WebAuthn Compliance and Interoperability.” In: *2025 IEEE 24th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom ’25)*. IEEE, Nov. 2025, pp. 2781–2788. DOI: [10.1109/Trustcom66490.2025.00329](https://doi.org/10.1109/Trustcom66490.2025.00329).
- [Mal+23] H. MALVAI, L. KOKORIS-KOGIAS, A. SONNINO, E. GHOSH, E. OZTÜRK, K. LEWI, S. LAWLOR. “Parakeet: Practical Key Transparency for End-to-End Encrypted Messaging.” In: *Proceedings of the Network and Distributed System Security Symposium (NDSS ’23)*. Internet Society, 2023. DOI: [10.14722/ndss.2023.24545](https://doi.org/10.14722/ndss.2023.24545).
- [Mat+25] A. MATZEN, A. RÜFFER, M. BYLLEMOS, O. HEINE, M. PAPAIOANNOU, G. CHOUDHARY, N. DRAGONI. “Challenges and Potential Improvements for Passkey Adoption—A Literature Review with a User-Centric Perspective.” In: *Applied Sciences* 15.8, Apr. 2025. DOI: [10.3390/app15084414](https://doi.org/10.3390/app15084414).
- [May08] R. MAYRHOFFER. “Ubiquitous Computing Security: Authenticating Spontaneous Interactions.” Habilitation Thesis. University of Vienna, Sept. 15, 2008. URL: <https://www.mayrhofer.eu.org/downloads/publications/Habil-Authenticating-Spontaneous-Interactions.pdf> (visited on May 20, 2026).
- [MG09] R. MAYRHOFFER, H. GELLERSEN. “Shake Well Before Use: Intuitive and Secure Pairing of Mobile Devices.” In: *IEEE Transactions on Mobile Computing* 8.6, June 2009, pp. 792–806. DOI: [10.1109/TMC.2009.51](https://doi.org/10.1109/TMC.2009.51).
- [Mec+24] L. MECKE, A. SAAD, S. PRANGE, U. GRUENEFELD, S. SCHNEEGASS, F. ALT. *Do They Understand What They Are Using? – Assessing Perception and Usage of Biometrics*. Technical Report. Oct. 2024. DOI: [10.48550/arXiv.2410.12661](https://doi.org/10.48550/arXiv.2410.12661).

- [Meh+19] A. MEHRABI, A. MAZZONI, D. JONES, A. STEED. "Evaluating the User Experience of Acoustic Data Transmission." In: *Personal and Ubiquitous Computing* 24, Dec. 2019, pp. 655–668. doi: [10.1007/s00779-019-01345-7](https://doi.org/10.1007/s00779-019-01345-7).
- [Mica] MICROSOFT. *Passkeys Overview*. Microsoft Support. URL: <https://support.microsoft.com/en-us/windows/passkeys-overview-301c8944-5ea2-452b-9886-97e4d2ef4422> (visited on May 20, 2026).
- [Micb] MICROSOFT. *Win32 Headers for WebAuthn*. GitHub Repository. URL: <https://github.com/microsoft/webauthn> (visited on May 20, 2026).
- [MBB18] J. MIEGEL, P. BRANCH, P. BLAMEY. "Wireless Communication Between Personal Electronic Devices and Hearing Aids Using High Frequency Audio and Ultrasound." In: *The Journal of the Acoustical Society of America* 144.4, Oct. 2018, pp. 2598–2604. doi: [10.1121/1.5063813](https://doi.org/10.1121/1.5063813).
- [Moh+23] S. MOHR et al. "A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe – Part 1: Event Description and Analysis." In: *Natural Hazards and Earth System Sciences* 23.2, Feb. 2023, pp. 525–551. doi: [10.5194/nhess-23-525-2023](https://doi.org/10.5194/nhess-23-525-2023).
- [Moo20] N. MOONEY. *Addition of a Network Transport*. GitHub Issue. Feb. 29, 2020. URL: <https://github.com/w3c/webauthn/issues/1381> (visited on May 20, 2026).
- [Mor+07] N. MORAVEJI, J. LI, J. DING, P. O'KELLEY, S. WOOLF. "Comicboarding: Using Comics as Proxies for Participatory Design with Children." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, 2007, pp. 1371–1374. doi: [10.1145/1240624.1240832](https://doi.org/10.1145/1240624.1240832).
- [MTT17] S. MORIYAMA, J. TAKAHASHI, Y. TOBE. "Creation of Communication Groups Using Sound of Smartphones." In: *Intelligent Environments 2017*. IOS Press, 2017, pp. 8–14. doi: [10.3233/978-1-61499-796-2-8](https://doi.org/10.3233/978-1-61499-796-2-8).
- [MRR18] S. MUKHERJEE, R. RAVINDRAN, D. RAYCHAUDHURI. "A Distributed Core Network Architecture for 5G Systems and Beyond." In: *Proceedings of the 2018 Workshop on Networking for Emerging Applications and Technologies (NEAT '18)*. ACM, 2018, pp. 33–38. doi: [10.1145/3229574.3229583](https://doi.org/10.1145/3229574.3229583).
- [Nan+13] R. NANDAKUMAR, K. K. CHINTALAPUDI, V. PADMANABHAN, R. VENKATESAN. "Dhwani: Secure Peer-to-Peer Acoustic NFC." In: *ACM SIGCOMM Computer Communication Review* 43.4, Aug. 2013, pp. 63–74. doi: [10.1145/2534169.2486037](https://doi.org/10.1145/2534169.2486037).
- [Naq20] B. NAQVI. "Towards Aligning Security and Usability During the System Development Lifecycle." PhD Thesis. Lappeenranta-Lahti University of Technology LUT, Nov. 17, 2020. URL: <https://urn.fi/URN:ISBN:978-952-335-586-6> (visited on May 20, 2026).
- [NTL19] E. NOVAK, Z. TANG, Q. LI. "Ultrasound Proximity Networking on Smart Mobile Devices for IoT Applications." In: *IEEE Internet of Things Journal* 6.1, Feb. 2019, pp. 399–409. doi: [10.1109/JIOT.2018.2848099](https://doi.org/10.1109/JIOT.2018.2848099).

- [OR20] S. OESCH, S. RUOTI. "That Was Then, This Is Now: A Security Evaluation of Password Generation, Storage, and Autofill in Browser-Based Password Managers." In: *29th USENIX Security Symposium (USENIX Security '20)*. USENIX Association, Aug. 2020, pp. 2165–2182. URL: <https://www.usenix.org/conference/usenixsecurity20/presentation/oesch> (visited on May 20, 2026).
- [OVC15] A. OLTEANU, S. VIEWEG, C. CASTILLO. "What to Expect When the Unexpected Happens: Social Media Communications Across Crises." In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, Feb. 2015, pp. 994–1009. DOI: [10.1145/2675133.2675242](https://doi.org/10.1145/2675133.2675242).
- [Oog+20] W. OOGAMI, H. GOMI, S. YAMAGUCHI, S. YAMANAKA, T. HIGURASHI. *Observation Study on Usability Challenges for Fingerprint Authentication Using WebAuthn-Enabled Android Smartphones*. Poster. Aug. 2020. URL: <https://www.usenix.org/conference/soups2020/presentation/oogami> (visited on May 20, 2026).
- [Ort26] A.-M. ORTLOFF. "Investigating Research Practice and Methods in Usable Security and Privacy." PhD Thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, Apr. 2026. DOI: [10.48565/bonndoc-852](https://doi.org/10.48565/bonndoc-852).
- [Ort+25] A.-M. ORTLOFF, J. TANG, A. ARUMUGAM, D. HUSCHINA, L. GEIERHAAS, F. MARTIUS, L. JANSEN, K. VON DER TWER, L. JUNGBLUTH, M. SMITH. "Replication: "No One Can Hack My Mind" - 10 Years Later: An Update and Outlook on Experts' and Non-Experts' Security Practices and Advice." In: *Twenty-First Symposium on Usable Privacy and Security (SOUPS '25)*. USENIX Association, 2025, pp. 435–454. URL: <https://www.usenix.org/conference/soups2025/presentation/ortloff> (visited on May 20, 2026).
- [Owe+21] K. OWENS, O. ANISE, A. KRAUSS, B. UR. "User Perceptions of the Usability and Security of Smartphones as FIDO2 Roaming Authenticators." In: *Seventeenth Symposium on Usable Privacy and Security (SOUPS '21)*. USENIX Association, Aug. 2021, pp. 57–76. URL: <https://www.usenix.org/conference/soups2021/presentation/owens> (visited on May 20, 2026).
- [OUA20] K. OWENS, B. UR, O. ANISE. "A Framework for Evaluating the Usability and Security of Smartphones as FIDO2 Roaming Authenticators." In: *Who Are You?! Adventures in Authentication Workshop (WAY '20)*. Aug. 2020. URL: <https://wayworkshop.org/2020/papers/way2020-owens.html> (visited on May 20, 2026).
- [Pag+21b] M. J. PAGE et al. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." In: *BMJ* 372, Mar. 2021. DOI: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71).
- [PL07] L. PALEN, S. B. LIU. "Citizen Communications in Crisis: Anticipating a Future of ICT-Supported Public Participation." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, Apr. 2007, pp. 727–736. DOI: [10.1145/1240624.1240736](https://doi.org/10.1145/1240624.1240736).

- [Pal+09] L. PALEN, S. VIEWEG, S. B. LIU, A. L. HUGHES. "Crisis in a Networked World: Features of Computer-Mediated Communication in the April 16, 2007, Virginia Tech Event." In: *Social Science Computer Review* 27.4, Nov. 2009, pp. 467–480. DOI: [10.1177/0894439309332302](https://doi.org/10.1177/0894439309332302).
- [Pee+22] C. PEETERS, C. PATTON, I. N. S. MUNYAKA, D. OLSZEWSKI, T. SHRIMPION, P. TRAYNOR. "SMS OTP Security (SOS): Hardening SMS-Based Two Factor Authentication." In: *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security (ASIA CCS '22)*. ACM, May 2022, pp. 2–16. DOI: [10.1145/3488932.3497756](https://doi.org/10.1145/3488932.3497756).
- [Pet+19] O. PETTER, M. HIRSCH, E. MUSHTAQ, P. HEVESI, P. LUKOWICZ. "Crowdsensing Under Recent Mobile Platform Background Service Restrictions: A Practical Approach." In: *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*. ACM, Sept. 2019, pp. 793–797. DOI: [10.1145/3341162.3344867](https://doi.org/10.1145/3341162.3344867).
- [Poz+18] M. POZZA, A. RAO, H. FLINCK, S. TARKOMA. "Network-in-a-Box: A Survey About On-Demand Flexible Networks." In: *IEEE Communications Surveys & Tutorials* 20.3, 2018, pp. 2407–2428. DOI: [10.1109/COMST.2018.2807125](https://doi.org/10.1109/COMST.2018.2807125).
- [QHD18] N. QUERMANN, M. HARBACH, M. DÜRMUTH. "The State of User Authentication in the Wild." In: *Who Are You?! Adventures in Authentication Workshop (WAY '18)*. Aug. 2018. URL: <https://wayworkshop.org/2018/papers/way2018-quermann.pdf> (visited on May 20, 2026).
- [Ram+19] A. RAMAN, S. JOGLEKAR, E. D. CRISTOFARO, N. SASTRY, G. TYSON. "Challenges in the Decentralised Web: The Mastodon Case." In: *Proceedings of the Internet Measurement Conference (IMC '19)*. ACM, 2019, pp. 217–229. DOI: [10.1145/3355369.3355572](https://doi.org/10.1145/3355369.3355572).
- [Ras21] B. RASMUSSEN. "A Usability Study of FIDO2 Roaming Software Tokens as a Password Replacement." Master's Thesis. Brigham Young University, Aug. 6, 2021. URL: <https://scholarsarchive.byu.edu/etd/9227> (visited on May 20, 2026).
- [Ree+19] K. REESE, T. SMITH, J. DUTSON, J. ARMKNECHT, J. CAMERON, K. SEAMONS. "A Usability Study of Five Two-Factor Authentication Methods." In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS '19)*. USENIX Association, Aug. 2019, pp. 357–370. URL: <https://www.usenix.org/conference/soups2019/presentation/reese> (visited on May 20, 2026).
- [RGP26] T. REITTINGER, M. GLAS, G. PERNUL. "Moving Beyond Passwords: Investigating the Effect of Digital Nudges on Passkey Adoption." In: *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. ACM, Apr. 2026. DOI: [10.1145/3772318.3791425](https://doi.org/10.1145/3772318.3791425).
- [Res18] E. RESCORLA. *The Transport Layer Security (TLS) Protocol Version 1.3*. RFC 8446. Aug. 2018. DOI: [10.17487/RFC8446](https://doi.org/10.17487/RFC8446).

- [RIB22] C. REUTER, L. L. IACONO, A. BENLIAN. "A Quarter Century of Usable Security and Privacy Research: Transparency, Tailorability, and the Road Ahead." In: *Behaviour & Information Technology* 41.10, June 2022, pp. 2035–2048. doi: [10.1080/0144929X.2022.2080908](https://doi.org/10.1080/0144929X.2022.2080908).
- [Reu+23] C. REUTER, M.-A. KAUFHOLD, T. BISELLI, H. PLEIL. "Increasing Adoption Despite Perceived Limitations of Social Media in Emergencies: Representative Insights on German Citizens' Perception and Trends from 2017 to 2021." In: *International Journal of Disaster Risk Reduction* 96, Oct. 2023. doi: [10.1016/j.ijdr.2023.103880](https://doi.org/10.1016/j.ijdr.2023.103880).
- [Reu+17] C. REUTER, M.-A. KAUFHOLD, T. SPIELHOFER, A. S. HAHNE. "Social Media in Emergencies: A Representative Study on Citizens' Perception in Germany." In: *Proceedings of the ACM on Human-Computer Interaction*. CSCW 1, Dec. 2017. doi: [10.1145/3134725](https://doi.org/10.1145/3134725).
- [Rey+18] J. REYNOLDS, T. SMITH, K. REESE, L. DICKINSON, S. RUOTI, K. SEAMONS. "A Tale of Two Studies: The Best and Worst of YubiKey Usability." In: *2018 IEEE Symposium on Security and Privacy (S&P '18)*. 2018, pp. 872–888. doi: [10.1109/SP.2018.00067](https://doi.org/10.1109/SP.2018.00067).
- [Riv+26] M. RIVERA-DOURADO, R. PÉREZ-JOVE, A. PAZOS, J. VÁZQUEZ-NAYA. *Usability of Passwordless Authentication in Wi-Fi Networks: A Comparative Study of Passkeys and Passwords in Captive Portals*. Technical Report. Mar. 2026. doi: [10.48550/arXiv.2603.25290](https://doi.org/10.48550/arXiv.2603.25290).
- [Riv+25] M. RIVERA-DOURADO, C. XENAKIS, A. PAZOS, J. VÁZQUEZ-NAYA. "EAP-FIDO: A Novel EAP Method for Using FIDO2 Credentials for Network Authentication." In: *Computer Networks* 266, July 2025. doi: [10.1016/j.comnet.2025.111348](https://doi.org/10.1016/j.comnet.2025.111348).
- [Ruo+19] S. RUOTI, J. ANDERSEN, L. DICKINSON, S. HEIDBRINK, T. MONSON, M. O'NEILL, K. REESE, B. SPENDLOVE, E. VAZIRIPOUR, J. WU, D. ZAPPALA, K. SEAMONS. "A Usability Study of Four Secure Email Tools Using Paired Participants." In: *ACM Transactions on Privacy and Security (TOPS)* 22.2, Apr. 2019. doi: [10.1145/3313761](https://doi.org/10.1145/3313761).
- [RS19] S. RUOTI, K. SEAMONS. "Johnny's Journey Toward Usable Secure Email." In: *IEEE Security & Privacy* 17.6, Oct. 2019, pp. 72–76. doi: [10.1109/MSEC.2019.2933683](https://doi.org/10.1109/MSEC.2019.2933683).
- [SH22] N. SADEQ, M. S. HOSSAIN. "Privacy Breach in Android Smartphone Through Inaudible Sound." In: *2022 IEEE Wireless Communications and Networking Conference (WCNC '22)*. Apr. 2022, pp. 2441–2446. doi: [10.1109/WCNC51071.2022.9771721](https://doi.org/10.1109/WCNC51071.2022.9771721).
- [Sas+14b] S. SASANO, F. SASAMORI, O. TAKYU, S. HANDA. "Performance Evaluation of Acoustic OFDM with Repetition Coding." In: *2014 International Conference on Information and Communication Technology Convergence (ICTC '14)*. Oct. 2014, pp. 798–799. doi: [10.1109/ICTC.2014.6983292](https://doi.org/10.1109/ICTC.2014.6983292).

- [SCC20] U. SATAPATHY, S. CHATTERJEE, S. CHAKRABORTY. "SilentInformer: A Framework for Information Dispersion Using Inaudible Acoustic Signals." In: *2020 International Conference on COMMunication Systems & NETworkS (COMSNETS '20)*. Jan. 2020, pp. 543–546. DOI: [10.1109/COMSNETS48256.2020.9027493](https://doi.org/10.1109/COMSNETS48256.2020.9027493).
- [SU09] N. SAXENA, M. B. UDDIN. "Blink 'Em All: Scalable, User-Friendly and Secure Initialization of Wireless Sensor Nodes." In: *Cryptology and Network Security (CANS '09)*. Springer, 2009, pp. 154–173. DOI: [10.1007/978-3-642-10433-6_11](https://doi.org/10.1007/978-3-642-10433-6_11).
- [Sch+16] S. SCHRÖDER, M. HUBER, D. WIND, C. ROTTERMANNER. "When SIGNAL Hits the Fan: On the Usability and Security of State-of-the-Art Secure Mobile Messaging." In: *EuroUSEC '16*. Internet Society, July 2016. URL: <http://dx.doi.org/10.14722/eurosec.2016.23012> (visited on May 25, 2026).
- [Set+24] S. C. SETHURAMAN, A. MITRA, A. GHOSH, R. T. S. BABU. "TUSH-Key: Transferable User Secrets on Hardware Key." In: *2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS '24)*. Oct. 2024, pp. 176–185. DOI: [10.1109/ICPADS63350.2024.00032](https://doi.org/10.1109/ICPADS63350.2024.00032).
- [SS25] M. SHIRAISHI, T. SHINAGAWA. "Hardware Authenticator Binding: A Secure Alternative to Passkeys." In: *2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC '24)*. IEEE, July 2025, pp. 1095–1100. DOI: [10.1109/COMPSAC65507.2025.00141](https://doi.org/10.1109/COMPSAC65507.2025.00141).
- [SS15] M. SHIRVANIAN, N. SAXENA. "On the Security and Usability of Crypto Phones." In: *Proceedings of the 31st Annual Computer Security Applications Conference (ACSAC '15)*. ACM, Dec. 2015, pp. 21–30. DOI: [10.1145/2818000.2818007](https://doi.org/10.1145/2818000.2818007).
- [SSG17] M. SHIRVANIAN, N. SAXENA, J. J. GEORGE. "On the Pitfalls of End-to-End Encrypted Communications: A Study of Remote Key-Fingerprint Verification." In: *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC '17)*. ACM, Dec. 2017, pp. 499–511. DOI: [10.1145/3134600.3134610](https://doi.org/10.1145/3134600.3134610).
- [Sil+14] D. SILVER, S. JANA, D. BONEH, E. CHEN, C. JACKSON. "Password Managers: Attacks and Defenses." In: *23rd USENIX Security Symposium (USENIX Security '14)*. USENIX Association, Aug. 2014, pp. 449–464. URL: <https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/silver> (visited on May 20, 2026).
- [Sim+23] L. SIMKO, H. SRI RAMULU, T. KOHNO, Y. ACAR. "The Use and Non-Use of Technology During Hurricanes." In: *Proceedings of the ACM on Human-Computer Interaction*. CSCW2 7, Oct. 2023. DOI: [10.1145/3610215](https://doi.org/10.1145/3610215).
- [Sol18] SOLOKEYS. *Solo 1: Open Security Key Supporting FIDO2 & U2F over USB + NFC*. GitHub Repository, 2018. URL: <https://github.com/solokeys/solo> (visited on May 20, 2026).
- [STU08] C. SORIENTE, G. TSUDIK, E. UZUN. "HAPADEP: Human-Assisted Pure Audio Device Pairing." In: *Information Security (ISC '08)*. Springer, 2008, pp. 385–400. DOI: [10.1007/978-3-540-85886-7_27](https://doi.org/10.1007/978-3-540-85886-7_27).
- [SA00] F. STAJANO, R. ANDERSON. "The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks." In: *Security Protocols*. Springer, 2000, pp. 172–182. DOI: [10.1007/10720107_24](https://doi.org/10.1007/10720107_24).

- [SP11] K. STARBIRD, L. PALEN. “Voluntweeters’ Self-Organizing by Digital Volunteers in Times of Crisis.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’11)*. ACM, 2011, pp. 1071–1080. doi: [10.1145/1978942.1979102](https://doi.org/10.1145/1978942.1979102).
- [Sta26] STATCOUNTER. *Desktop vs Mobile Market Share Worldwide*. Apr. 2026. URL: <https://gs.statcounter.com/platform-market-share/desktop-mobile/worldwide/> (visited on May 20, 2026).
- [Ste14] STEVIE AWARDS. *VeriFone’s Way2ride*. The American Business Awards. 2014. URL: <https://stevieawards.com/aba/verifones-way2ride> (visited on May 20, 2026).
- [SJ14] B. STOCK, M. JOHNS. “Protecting Users against XSS-based Password Manager Abuse.” In: *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security (ASIA CCS ’14)*. ACM, June 2014, pp. 183–194. DOI: [10.1145/2590296.2590336](https://doi.org/10.1145/2590296.2590336).
- [Sur+24] J. SURAY, J. H. KLEMMER, J. SCHMÜSER, S. FAHL. *How the Future Works at SOUPS: Analyzing Future Work Statements and Their Impact on Usable Security and Privacy Research*. Technical Report. May 2024. DOI: [10.48550/arXiv.2405.20785](https://doi.org/10.48550/arXiv.2405.20785).
- [Tan+17a] J. TAN, L. BAUER, J. BONNEAU, L. F. CRANOR, J. THOMAS, B. UR. “Can Unicorns Help Users Compare Crypto Key Fingerprints?” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI ’17)*. ACM, May 2017, pp. 3787–3798. DOI: [10.1145/3025453.3025733](https://doi.org/10.1145/3025453.3025733).
- [TCL23] K. L. TAN, C.-H. CHI, K.-Y. LAM. “Survey on Digital Sovereignty and Identity: From Digitization to Digitalization.” In: *ACM Computing Surveys* 56.3, Oct. 2023. DOI: [10.1145/3616400](https://doi.org/10.1145/3616400).
- [Tan+20] M. L. TAN, R. PRASANNA, K. STOCK, E. E. DOYLE, G. LEONARD, D. JOHNSTON. “Usability Factors Influencing the Continuance Intention of Disaster Apps: A Mixed-Methods Study.” In: *International Journal of Disaster Risk Reduction* 50, Nov. 2020. DOI: [10.1016/j.ijdr.2020.101874](https://doi.org/10.1016/j.ijdr.2020.101874).
- [Tan+17b] M. L. TAN, R. PRASANNA, K. STOCK, E. HUDSON-DOYLE, G. LEONARD, D. JOHNSTON. “Mobile Applications in Crisis Informatics Literature: A Systematic Review.” In: *International Journal of Disaster Risk Reduction* 24, Sept. 2017, pp. 297–311. DOI: [10.1016/j.ijdr.2017.06.009](https://doi.org/10.1016/j.ijdr.2017.06.009).
- [Tay23] D. TAYLOR. “What’s the Password? Account Sharing in the Context of Passwordless Authentication.” Undergraduate Honors Thesis. Brigham Young University, Aug. 2023. URL: https://scholarsarchive.byu.edu/studentpub_uht/321 (visited on May 20, 2026).
- [Tem+25b] D. TEMOSHOK, D. PROUD-MADRUGA, Y.-Y. CHOONG, R. GALLUZZO, S. GUPTA, C. LASALLE, N. LEFKOVITZ, A. REGENSCHEID. *Digital Identity Guidelines*. NIST Special Publication (SP) 800-63-4. National Institute of Standards and Technology, July 2025. DOI: [10.6028/NIST.SP.800-63-4](https://doi.org/10.6028/NIST.SP.800-63-4).

- [Thi+23] A. H. THIEKEN, P. BUBECK, A. HEIDENREICH, J. VON KEYSERLINGK, L. DILLENARDT, A. OTTO. "Performance of the Flood Warning System in Germany in July 2021 — Insights from Affected Residents." In: *Natural Hazards and Earth System Sciences* 23.2, Mar. 2023, pp. 973–990. DOI: [10.5194/nhess-23-973-2023](https://doi.org/10.5194/nhess-23-973-2023).
- [Tie+25] E. TIEFENAU, J. A. GROHS, M. HÄRING, M. SMITH, C. TIEFENAU. "'They Are Responsible for Ensuring That I Can Continue to Use the Service. Investigating Users' Expectations Towards 2FA Recovery in Germany.'" In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, Apr. 2025. DOI: [10.1145/3706598.3714245](https://doi.org/10.1145/3706598.3714245).
- [Tri+17] S. TRIFUNOVIC, S. T. KOUYOUMDJIEVA, B. DISTL, L. PAJEVIC, G. KARLSSON, B. PLATTNER. "A Decade of Research in Opportunistic Networks: Challenges, Relevance, and Future Directions." In: *IEEE Communications Magazine* 55.1, Jan. 2017, pp. 168–173. DOI: [10.1109/MCOM.2017.1500527CM](https://doi.org/10.1109/MCOM.2017.1500527CM).
- [Tri20] TRILLBIT INC. *TrillBit: Contactless, Fast and Secure Authentication Protocol*. 2020. URL: <https://www.trillbit.com/technology-trillbit> (visited on May 20, 2026).
- [Ulq+21] E. ULQINAKU, H. ASSAL, A. ABDOU, S. CHIASSON, S. CAPKUN. "Is Real-time Phishing Eliminated with FIDO? Social Engineering Downgrade Attacks against FIDO Protocols." In: *30th USENIX Security Symposium (USENIX Security '21)*. 2021, pp. 3811–3828. ISBN: 978-1-939133-24-3. URL: <https://www.usenix.org/conference/usenixsecurity21/presentation/ulqinaku> (visited on Mar. 14, 2026).
- [ULC19] E. ULQINAKU, D. LAIN, S. CAPKUN. "2FA-PP: 2nd Factor Phishing Prevention." In: *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '19)*. ACM, May 2019, pp. 60–70. DOI: [10.1145/3317549.3323404](https://doi.org/10.1145/3317549.3323404).
- [Ung+15] N. UNGER, S. DECHAND, J. BONNEAU, S. FAHL, H. PERL, I. GOLDBERG, M. SMITH. "SoK: Secure Messaging." In: *2015 IEEE Symposium on Security and Privacy (S&P '15)*. IEEE, May 2015, pp. 232–249. DOI: [10.1109/SP.2015.22](https://doi.org/10.1109/SP.2015.22).
- [Ur+15] B. UR, F. NOMA, J. BEES, S. M. SEGRETI, R. SHAY, L. BAUER, N. CHRISTIN, L. F. CRANOR. "'I Added '!' at the End to Make It Secure': Observing Password Creation in the Lab." In: *Eleventh Symposium On Usable Privacy and Security (SOUPS '15)*. USENIX Association, 2015, pp. 123–140. URL: <https://www.usenix.org/conference/soups2015/proceedings/presentation/ur> (visited on May 20, 2026).
- [UKA07] E. UZUN, K. KARVONEN, N. ASOKAN. "Usability Analysis of Secure Pairing Methods." In: *Financial Cryptography and Data Security (FC '07)*. Springer, 2007, pp. 307–324. DOI: [10.1007/978-3-540-77366-5_29](https://doi.org/10.1007/978-3-540-77366-5_29).
- [VHD97] J. D. VAN DER LAAN, A. HEINO, D. DE WAARD. "A Simple Procedure for the Assessment of Acceptance of Advanced Transport Telematics." In: *Transportation Research Part C: Emerging Technologies*, Feb. 1997. DOI: [10.1016/S0968-090X\(96\)00025-3](https://doi.org/10.1016/S0968-090X(96)00025-3).

- [Vas+20] D. A. VASILE, M. KLEPPMANN, D. R. THOMAS, A. R. BERESFORD. "Ghost Trace on the Wire? Using Key Evidence for Informed Decisions." In: *Security Protocols XXVII*. Springer International Publishing, 2020, pp. 245–257. DOI: [10.1007/978-3-030-57043-9_23](https://doi.org/10.1007/978-3-030-57043-9_23).
- [Vaz+19] E. VAZIRIPOUR, D. HOWARD, J. TYLER, M. O'NEILL, J. WU, K. SEAMONS, D. ZAPPALA. "I Don't Even Have to Bother Them! Using Social Media to Automate the Authentication Ceremony in Secure Messaging." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, May 2019. DOI: [10.1145/3290605.3300323](https://doi.org/10.1145/3290605.3300323).
- [Vaz+17] E. VAZIRIPOUR, J. WU, M. O'NEILL, R. CLINTON, J. WHITEHEAD, S. HEIDBRINK, K. SEAMONS, D. ZAPPALA. "Is That You, Alice? A Usability Study of the Authentication Ceremony of Secure Messaging Applications." In: *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security (SOUPS '17)*. USENIX Association, July 2017, pp. 29–47. URL: <https://www.usenix.org/conference/soups2017/technical-sessions/presentation/vaziripour> (visited on May 20, 2026).
- [Vaz+18] E. VAZIRIPOUR, J. WU, M. O'NEILL, D. METRO, J. COCKRELL, T. MOFFETT, J. WHITEHEAD, N. BONNER, K. SEAMONS, D. ZAPPALA. "Action Needed! Helping Users Find and Complete the Authentication Ceremony in Signal." In: *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security (SOUPS '18)*. USENIX Association, Aug. 2018, pp. 47–62. URL: <https://www.usenix.org/conference/soups2018/presentation/vaziripour> (visited on May 20, 2026).
- [WAR25] K. C. WANG, S. S. ARORA, M. K. REITER. "The 2FA Illusion: Uncovering Weak Links of Web Account Access in the Wild." In: *2025 IEEE Annual Computer Security Applications Conference (ACSAC '25)*. IEEE, Dec. 2025, pp. 658–672. DOI: [10.1109/ACSAC67867.2025.00060](https://doi.org/10.1109/ACSAC67867.2025.00060).
- [Wan+16] Q. WANG, K. REN, M. ZHOU, T. LEI, D. KOUTSONIKOLAS, L. SU. "Messages Behind the Sound." In: *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, Oct. 2016. DOI: [10.1145/2973750.2973765](https://doi.org/10.1145/2973750.2973765).
- [WZH21] Y. WANG, X. ZHU, H. HAN. "ChirpMu: Chirp Based Imperceptible Information Broadcasting with Music." In: *2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS '21)*. IEEE, June 2021. DOI: [10.1109/IWQOS52092.2021.9521353](https://doi.org/10.1109/IWQOS52092.2021.9521353).
- [Was+16] R. WASH, E. RADER, R. BERMAN, Z. WELLMER. "Understanding Password Choices: How Frequently Entered Passwords Are Re-used across Websites." In: *Twelfth Symposium on Usable Privacy and Security (SOUPS '16)*. USENIX Association, 2016, pp. 175–188. URL: <https://www.usenix.org/conference/soups2016/technical-sessions/presentation/wash> (visited on May 20, 2026).
- [WT99] A. WHITTEN, J. D. TYGAR. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." In: *8th USENIX Security Symposium (USENIX Security '99)*, Aug. 1999. URL: <https://www.usenix.org/conference/8th-usenix-security->

- [symposium/why-johnny-cant-encrypt-usability-evaluation-pgp-50](#) (visited on May 20, 2026).
- [Wor19] WORLD WIDE WEB CONSORTIUM. *Web Authentication: An API for Accessing Public Key Credentials Level 1*. Ed. by D. BALFANZ, A. CZESKIS, J. HODGES, J. C. JONES, M. B. JONES, A. KUMAR, A. LIAO, R. LINDEMANN, E. LUNDBERG. W3C Recommendation. Mar. 4, 2019. URL: <https://www.w3.org/TR/2019/REC-webauthn-1-20190304/>.
- [Wor21] WORLD WIDE WEB CONSORTIUM. *Web Authentication: An API for Accessing Public Key Credentials Level 2*. Ed. by J. HODGES, J. C. JONES, M. B. JONES, A. KUMAR, E. LUNDBERG. W3C Recommendation. Apr. 8, 2021. URL: <https://www.w3.org/TR/2021/REC-webauthn-2-20210408/>.
- [Wu+19] J. WU, C. GATRELL, D. HOWARD, J. TYLER, E. VAZIRIPOUR, K. SEAMONS, D. ZAPPALA. “Something Isn’t Secure, but I’m Not Sure How That Translates into a Problem: Promoting Autonomy by Designing for Understanding in Signal.” In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS ’19)*. USENIX Association, Aug. 2019, pp. 137–153. URL: <https://www.usenix.org/conference/soups2019/presentation/wu> (visited on May 20, 2026).
- [WH23] L. WÜRSCHING, M. HOLLICK. “Maintaining App Services in Disrupted Cities: A Crisis and Resilience Evaluation Tool.” en. In: *NetMob ’23 Book of Abstracts*. IMDEA, Oct. 2023, pp. 141–142. URL: https://netmob.org/www23/book_of_abstract/book_of_abstract.pdf (visited on May 20, 2026).
- [Xu+21] W. XU, J. ZHANG, S. HUANG, C. LUO, W. LI. “Key Generation for Internet of Things: A Contemporary Survey.” In: *ACM Computing Surveys* 54.1, Jan. 2021. DOI: [10.1145/3429740](#).
- [YS24] T. K. YADAV, K. SEAMONS. “A Security and Usability Analysis of Local Attacks Against FIDO2.” In: *Proceedings of the Network and Distributed System Security Symposium (NDSS ’24)*. Internet Society, 2024. DOI: [10.14722/ndss.2024.24327](#).
- [YK22] K. YAMAMOTO, H. KUBO. “32kbps Terrestrial Acoustic Communications of Single Carrier Block Transmission Employing Transmit Precoding and Channel Interpolation.” In: *IEICE Communications Express* 11.3, 2022, pp. 141–147. DOI: [10.1587/comex.2021XBL0206](#).
- [Yan+04] J. YAN, A. BLACKWELL, R. ANDERSON, A. GRANT. “Password Memorability and Security: Empirical Results.” In: *IEEE Security & Privacy* 2.5, 2004, pp. 25–31. DOI: [10.1109/MSP.2004.81](#).
- [YHV24] W.-Z. YEOH, L. HANZLIK, O. VALTA. “Delegating FIDO Credentials Using Single-use ECDSA Signatures.” In: *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security (ASIA CCS ’24)*. ACM, July 2024, pp. 1801–1813. DOI: [10.1145/3634737.3657004](#).
- [Yeo+23] W.-Z. YEOH, M. KEPKOWSKI, G. HEIDE, D. KAAFAR, L. HANZLIK. “Fast IDentity Online with Anonymous Credentials (FIDO-AC).” In: *32nd USENIX Security Symposium (USENIX Security ’23)*. USENIX Association, Aug. 2023, pp. 3029–3046. URL: <https://www.usenix.org/conference/usenixsecurity23/presentation/yeoh> (visited on May 20, 2026).

- [Yub20] YUBICO. *YubiKey 5 Series: The Multi-Protocol, Highest Assurance Security Key That Enables Passwordless*. 2020. URL: <https://www.yubico.com/products/yubikey-5-overview/> (visited on May 20, 2026).
- [Yuba] YUBICO. *libfido2: Provides Library Functionality and Command-Line Tools to Communicate with a FIDO Device over USB or NFC*. GitHub Repository. URL: <https://github.com/Yubico/libfido2> (visited on May 20, 2026).
- [Yubb] YUBICO. *python-fido2: Provides Library Functionality for Communicating with a FIDO Device over USB*. GitHub Repository. URL: <https://github.com/Yubico/python-fido2> (visited on May 20, 2026).
- [Zar+14] K. ZARIFIS, T. FLACH, S. NORI, D. CHOFFNES, R. GOVINDAN, E. KATZ-BASSETT, Z. M. MAO, M. WELSH. "Diagnosing Path Inflation of Mobile Client Traffic." In: *Passive and Active Measurement (PAM '14)*. Springer, 2014, pp. 23–33. DOI: [10.1007/978-3-319-04918-2_3](https://doi.org/10.1007/978-3-319-04918-2_3).
- [Zha+14] B. ZHANG, Q. ZHAN, S. CHEN, M. LI, K. REN, C. WANG, D. MA. "PriWhisper: Enabling Keyless Secure Acoustic Communication for Smartphones." In: *IEEE Internet of Things Journal* 1.1, Feb. 2014, pp. 33–45. DOI: [10.1109/JIOT.2014.2297998](https://doi.org/10.1109/JIOT.2014.2297998).
- [Zha+19] X. ZHANG, J. LIU, S. CHEN, Y. KONG, K. REN. "PriWhisper+: An Enhanced Acoustic Short-Range Communication System for Smartphones." In: *IEEE Internet of Things Journal* 6.1, Feb. 2019, pp. 614–627. DOI: [10.1109/JIOT.2018.2850524](https://doi.org/10.1109/JIOT.2018.2850524).
- [Zho+19] M. ZHOU, Q. WANG, K. REN, D. KOUTSONIKOLAS, L. SU, Y. CHEN. "Dolphin: Real-Time Hidden Acoustic Signal Capture with Smartphones." In: *IEEE Transactions on Mobile Computing* 18.3, Mar. 2019, pp. 560–573. DOI: [10.1109/TMC.2018.2842771](https://doi.org/10.1109/TMC.2018.2842771).
- [Zim+18] V. ZIMMERMANN, N. GERBER, M. KLEBOTH, A. VON PREUSCHEN, K. SCHMIDT, P. MAYER. "The Quest to Replace Passwords Revisited - Rating Authentication Schemes." In: *Proceedings of the 12th International Symposium on Human Aspects of Information Security and Assurance (HAISA '18)*. 2018, pp. 38–48. DOI: [10.5445/IR/1000085729](https://doi.org/10.5445/IR/1000085729).
- [Zol+20] B. ZOLFAGHARI, G. SRIVASTAVA, S. ROY, H. R. NEMATİ, F. AFGHAH, T. KOSHIBA, A. RAZI, K. BIBAK, P. MITRA, B. K. RAI. "Content Delivery Networks: State of the Art, Trends, and Future Roadmap." In: *ACM Computing Surveys* 53.2, Apr. 2020. DOI: [10.1145/3380613](https://doi.org/10.1145/3380613).

List of Figures

Figure 1	Dissertation Goal	2
Figure 2	Authentication Scenarios	4
Figure 3	Smartphone Interactions in our FIDO2 Authentication Experiment	10
Figure 4	SUS and Acceptance Scores by Authenticator Type	12
Figure 5	FIDO2 Extension Protocol Flow	16
Figure 6	Comic Board Introducing the Crisis Scenario	26
Figure 7	Active MitM Attack	34
Figure 8	Contact Verification in a Group	36
Figure 9	PairSonic User Interface	37
Figure 10	SUS Scores and Completion Times for SafeSlinger and PairSonic	38
Figure 11	Acoustic Data Transmission Model	40
Figure 12	Experimental Setup	42
Figure 13	Acoustic Transmission Reliability by Distance	42
Figure 14	Acoustic Transmission Results by Use Case Category	43
Figure 15	Scalability Comparison of PairSonic and SafeSlinger	46

List of Tables

Table 1	Perceived Strengths and Weaknesses of FIDO2	13
Table 2	Early Reception and Impact of Paper [A]	15
Table 3	FIDO2 Extension Compatibility	18
Table 4	Early Reception and Impact of Paper [B]	20
Table 5	App Demand for Local-Only Connectivity	27
Table 6	App Demand for Crisis-Specific Use Cases	28
Table 7	Early Reception and Impact of Paper [C]	29
Table 8	Early Reception and Impact of Papers [D; E]	39
Table 9	Evaluated Acoustic Data Transmission Schemes	41
Table 10	Surveyed Acoustic Data Transmission Schemes	47

List of Acronyms

2FA	Two-Factor Authentication
6G	Sixth-Generation Technology Standard for Cellular Networks
API	Application Programming Interface
ATI	Affinity for Technology Interaction
BLE	Bluetooth Low Energy
BSI	German Federal Office for Information Security
CBOR	Concise Binary Object Representation
CDN	Content Delivery Network
COTS	Commercial Off-the-Shelf
CTAP	Client to Authenticator Protocol
D2D	Device-to-Device
DNS	Domain Name System
DOI	Digital Object Identifier
E2EE	End-to-End Encryption
EAP	Extensible Authentication Protocol
eID	Electronic Identification
FBI	Federal Bureau of Investigation
FIDO2	Fast IDentity Online 2
IETF	Internet Engineering Task Force
IoT	Internet of Things
IRB	Institutional Review Board
ISO	International Organization for Standardization
JSON	JavaScript Object Notation

MANET	Mobile Ad Hoc Network
MEC	Multi-Access Edge Computing
MitM	Machine-in-the-Middle
MNO	Mobile Network Operator
NFC	Near-Field Communication
NIST	National Institute of Standards and Technology
OOB	Out-of-Band
OTP	One-Time Password
PGP	Pretty Good Privacy
PIN	Personal Identification Number
PKI	Public-Key Infrastructure
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QR	Quick-Response
RF	Radio Frequency
RX	Receiver
SMS	Short Message Service
SSO	Single Sign-On
SUS	System Usability Scale
TER	Total Error Rate
TLS	Transport Layer Security
TOFU	Trust on First Use
TOTP	Time-Based One-Time Password
TX	Transmitter
U2F	Universal 2nd Factor
USB	Universal Serial Bus
UX	User Experience
W3C	World Wide Web Consortium
WAV	Waveform Audio File Format
WebAuthn	Web Authentication

FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones

Citation	L. Würsching, F. Putz, S. Haesler, and M. Hollick. “FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones.” In: <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (Apr. 23–28, 2023). CHI ’23. Hamburg, Germany: ACM, Apr. 2023. ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3580993
Contribution	This publication is discussed in Chapter 2. My contribution is described in the section “ <i>Collaborations and My Contribution</i> ” on page x.
Rights notice	© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author’s version of the work. It is included in this dissertation in accordance with ACM author rights. The definitive Version of Record was published in <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI ’23)</i> , DOI: https://doi.org/10.1145/3544548.3580993 .

FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones

Leon Würsching*

Technical University of Darmstadt
Darmstadt, Germany
lwuersching@seemoo.tu-darmstadt.de

Steffen Haesler

Technical University of Darmstadt
Darmstadt, Germany
haesler@peasec.tu-darmstadt.de

Florentin Putz*

Technical University of Darmstadt
Darmstadt, Germany
fputz@seemoo.tu-darmstadt.de

Matthias Hollick

Technical University of Darmstadt
Darmstadt, Germany
mhollick@seemoo.tu-darmstadt.de

ABSTRACT

Modern smartphones support FIDO2 passwordless authentication using either external security keys or internal biometric authentication, but it is unclear whether users appreciate and accept these new forms of web authentication for their own accounts. We present the first lab study (N=87) comparing platform and roaming authentication on smartphones, determining the practical strengths and weaknesses of FIDO2 as perceived by users in a mobile scenario. Most participants were willing to adopt passwordless authentication during our in-person user study, but closer analysis shows that participants prioritize usability, security, and availability differently depending on the account type. We identify remaining adoption barriers that prevent FIDO2 from succeeding password authentication, such as missing support for contemporary usage patterns, including account delegation and usage on multiple clients.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI; Laboratory experiments; Empirical studies in ubiquitous and mobile computing*; • **Security and privacy** → *Usability in security and privacy*.

KEYWORDS

Usability, Security, Passwordless, User Authentication, Biometrics, Accounts

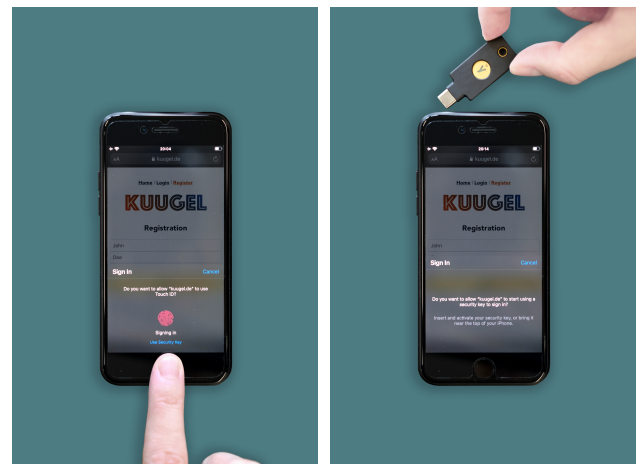
ACM Reference Format:

Leon Würsching, Florentin Putz, Steffen Haesler, and Matthias Hollick. 2023. FIDO2 the Rescue? Platform vs. Roaming Authentication on Smartphones. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3544548.3580993>

*Both authors contributed equally to this research.

CHI '23, April 23–28, 2023, Hamburg, Germany

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 23–28, 2023, Hamburg, Germany, <https://doi.org/10.1145/3544548.3580993>.



(a) Group P: Apple Touch ID

(b) Group R: Yubico Yubikey

Figure 1: The different user interactions with the smartphone during the experiment. Group P used Apple Touch ID as a platform authenticator and Group R used a Yubico YubiKey as a roaming authenticator.

1 INTRODUCTION

Authentication is one of the central building blocks for securing the Internet. Since the 1970s, web services have conventionally relied on text-based passwords as the de-facto standard for user authentication. Secure passwords, however, are hard to memorize [72], password reuse enables severe attacks [38], and password-based authentication is prone to phishing [22]. Despite stronger security guarantees, all previous approaches in the “quest to replace passwords” failed to get widespread adoption, mainly because of their inferior deployability and usability [11].

To address the first entry barrier – *deployability* – more than 250 technology companies and browser vendors [29] have founded the Fast Identity Online Alliance (FIDO), jointly designing the FIDO2 standards for passwordless authentication [30]. As a result, all major web browsers are now FIDO2-ready [46], as they support the corresponding W3C Web Authentication (WebAuthn) standard [69]. FIDO2 mandates public-key cryptography to provide user authentication based on authenticators containing the user's private

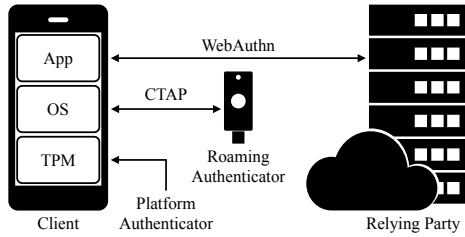


Figure 2: The parties and protocols involved in FIDO2 authentication scenarios. On the client side, “App” and “OS” are abbreviations for the application accessing the relying party and the client’s operating system, respectively.

key [28]. Unlike passwords, FIDO2 authentication is resistant to phishing, keylogging, replay attacks, and server breaches [41].

As more and more websites support FIDO2 authentication, it becomes essential to study the second entry barrier – *usability* – to find out how users react to this paradigm shift from knowledge-based factors to possession-based and biometrics-based factors. There are two variants of FIDO2 authentication with fundamentally different user interactions:

For *roaming authentication*, the private keys are stored on an external roaming authenticator, e.g., a YubiKey [76], which connects to the client device via Universal Serial Bus (USB), near-field communication (NFC), or Bluetooth Low Energy (BLE). A recent usability study within the desktop environment by Lyastani et al. [44] suggests that users accept and prefer roaming authentication as an alternative to passwords. Many participants, however, were concerned about carrying a security key with them physically and potentially losing it. Despite the prevalence of FIDO2 support in recent iPhones and Android devices [2, 27], **smartphones have not been studied as FIDO2 clients for roaming authentication yet**, but only as roaming authenticators themselves [51, 57].

The second variant of FIDO2, *platform authentication*, mitigates these availability concerns, as the smartphone’s integrated trusted platform module (TPM) stores the private keys, protected by an additional local authentication using the smartphone’s unlock mechanism, e.g., Apple Touch ID [7]. However, platform authentication raises new usability concerns about the fundamentally different mental model. Users need to grasp a more complex authentication method combining a private key stored on the smartphone’s TPM that is further protected by biometrics-based local authentication. Platform authentication on smartphones has received little attention so far. Oogami et al. [50] studied how to improve compatible websites’ user experience, and Lassak et al. [43] developed smartphone notifications addressing platform authentication misconceptions. However, **the research community lacks an understanding of whether users understand, accept, and trust passwordless platform authentication as an alternative to roaming authentication and passwords**.

In this work, we try to bridge this knowledge gap by reporting, to the best of our knowledge, the first large-scale lab study comparing FIDO2 platform authentication and roaming authentication on smartphones. We recruited 87 participants, randomly assigned

them to one of two groups, and had them perform a series of practical web authentication tasks on an Apple iPhone. The first group used platform authentication with Apple Touch ID for the web (Figure 1a), while the second group used roaming authentication with an NFC-based YubiKey (Figure 1b). The participants reflected on their experience in a survey, which featured a combination of quantitative questions using standardized metrics and qualitative open-ended text questions. This paper presents our work as follows:

- As our main contribution, we conduct the **first large-scale lab study that compares FIDO2 platform and roaming authentication on smartphones** (Section 4).
- We show that platform and roaming authentication have excellent usability on smartphones, but lay users generally prefer platform authentication (Section 5).
- From our questionnaire’s qualitative categories, we identify the strengths and weaknesses of FIDO2 on smartphones (Section 6). Based on our participants’ feedback, we investigate how to address the weaknesses of passwordless authentication and discuss account-specific adoption decisions and usage patterns (Section 7).
- We provide a replication package with our evaluation scripts and the pseudonymized dataset [71] collected in our study, consisting of 22 variables for each of our 87 participants. We also release our mockup website’s source code [70] to facilitate future work.

Our questionnaire encouraged participants to reflect on their everyday authentication use cases and whether they would be willing to use FIDO2 for their own accounts. While most of our participants were generally willing to adopt passwordless authentication, account-specific adoption barriers remain for both roaming and platform authentication, which we discuss in detail.

2 BACKGROUND AND RELATED WORK

This section explains how the FIDO2 standards enable passwordless authentication on smartphones. Section 2.1 describes relevant parts of the standards. Then, we explain roaming authentication (Section 2.2) and platform authentication (Section 2.3), including summaries of related studies of FIDO2. We also consider authentication-related studies outside the FIDO2 cosmos (Section 2.4).

2.1 FIDO2

The FIDO2 standards provide strong and passwordless authentication for the Internet. All major web browsers support them [46], and an increasing number of online services offer FIDO2 authentication both in one-factor authentication (1FA) and multi-factor authentication (MFA) scenarios [29]. FIDO2 consists of two standards, WebAuthn [69] and Client to Authenticator Protocol (CTAP) [28], which specify the communication between relying party, the client, and the authenticator (Figure 2). Together they form a challenge-response protocol, confirming the user’s identity to the relying party by verifying the user’s possession of the authenticator that manages their public key credentials.

In our study, WebAuthn handles the communication between a smartphone (client device) and a website (relying party). However, FIDO2 is flexible and also supports local authentication scenarios, e.g., a user unlocking a Windows computer via Windows Hello

[47]. On the client-side, WebAuthn expects an authenticator with access to the secret key corresponding to the public key stored at the relying party. FIDO2 supports two types of authenticators:

2.2 Roaming Authentication

A roaming authenticator is a CTAP-conforming external device that manages all public key credentials. During authentication, the relying party's challenge is forwarded to the roaming authenticator, solved locally with the secret key, and returned to the relying party. There is a large ecosystem of suitable roaming authenticators which are compatible with smartphones via NFC [75], BLE [36], or Apple Lightning/USB Type C [25]. The smartphone itself must be equipped with a CTAP-conforming interface and a WebAuthn-compatible web browser. Since 2019, FIDO2 roaming authentication has been supported on Android smartphones running Android 7 or later [27] and on Apple iPhones running iOS 13.3 or later via Apple Lightning [73] and NFC [2].

Lyastani et al. [44] conducted the first large-scale user study of FIDO2 roaming authentication on computers. In their between-groups study, participants either authenticated with passwords or used a YubiKey, concluding that roaming authentication is more usable and accepted than passwords. Farke et al. [23] studied FIDO2 roaming authentication as an unlocking mechanism for computers, using YubiKeys with enabled personal identification number (PIN) protection. They reported that participants stopped using roaming authentication because it was slower than their password manager.

Previous studies investigated whether smartphones, themselves, could be used as external FIDO2 roaming authenticators to address deployment and availability issues of traditional security keys. Owens et al. [51, 52] reported a between-groups observation study, comparing passwords to smartphones as roaming authenticators. They concluded that users understand the security benefits of FIDO2 but still find password-based authentication more usable. Rasmussen [57] conducted a between-groups user study showing that smartphones as roaming authenticators have similar usability and acceptance as YubiKeys. Both studies [51, 57] reported availability concerns regarding empty smartphone batteries.

We continue with a brief overview of user studies on the Universal 2nd Factor (U2F) [26], which is CTAP's predecessor and has similar user interaction. Lang et al. [41] reported on a two-year enterprise deployment of security keys within Google, laying the foundation for the U2F standard that is CTAP's predecessor. Das et al. [18, 20] studied U2F in non-enterprise environments and found that a two-factor authentication (2FA) method's acceptance did not correlate with its usability. Reynolds et al. [60] and Reese et al. [58] conducted longitudinal studies of U2F, reporting that the initial setup is cumbersome compared to the fast and easy authentication afterward. Ciolino et al. and Das et al. investigated lay users' perception of U2F, finding that, while sentiment towards U2F is lower than for SMS-based 2FA [13], some lay users do not feel the need to protect their accounts with MFA at all [19]. Colnago et al. [14] conducted a large-scale longitudinal study to explore 2FA adoption at a university, finding that $< 1\%$ use U2F as a 2FA method. More recent works on U2F by Das et al. and Reynolds et al. found that users are more likely to adopt MFA for essential accounts [21] and

that the users' initial negative perception of MFA methods fades over time [59].

2.3 Platform Authentication

Client devices qualified to manage cryptographic key pairs can be used as platform authenticators. During authentication, the relying party's challenge is solved directly on the client's platform authenticator. Most platform authenticators require the user to authenticate locally with a PIN or biometrics, augmenting FIDO2 to an MFA method. Any Android phone running Android 7 or later provides a FIDO2 platform authenticator, as Android's FIDO2 certification includes the built-in biometrics sensors [27]. Since iOS 14, Apple has equipped iPhones with a platform authenticator, namely Touch ID for the web [3], which we also refer to as *Touch ID* for simplicity. For Touch ID platform authentication, the iPhone's Secure Enclave manages FIDO2 credentials and requires the user to authenticate locally with Touch ID.

Oogami et al. [50] studied the usability of FIDO2 platform authentication on smartphones, conducting interviews to improve the website user experience for first-time FIDO2 users. However, they did not collect FIDO2 weaknesses or adoption barriers of platform authentication. Lassak et al. [43] studied misconceptions about FIDO2 platform authentication on Android phones and developed smartphone notifications to address them.

2.4 Miscellaneous

Independently of FIDO2, van den Boogaard [66] conducted an online study of users' understanding of biometrics-based authentication on mobile phones, finding that most participants use biometric options when available. Connors et al. [15] conducted an online study of *Let's authenticate* [16], which is similar to FIDO2 but uses authenticators (roaming or platform) to issue certificates for user authentication. Alqubaisi et al. [1] studied elements hindering the adoption of FIDO2 passwordless authentication, basing their analysis on the FIDO developer mailing list.

3 RESEARCH QUESTIONS

We continue the work of recent FIDO2 user studies but focus on passwordless authentication on smartphone clients. Thus, the research questions addressed in this work are:

- **RQ1** *What is the usability and acceptance of FIDO2 passwordless authentication using platform and roaming authentication on smartphones?*
- **RQ2** *What benefits and concerns do users consider when using FIDO2 passwordless authentication on smartphones?*
- **RQ3** *Which account types do users want to secure using FIDO2 passwordless authentication on smartphones?*

In contrast to related work, to the best of our knowledge, we are the first to study FIDO2 roaming authentication on smartphone clients and the first to conduct a lab study on any device type that compares FIDO2 platform authentication and roaming authentication. Our main goal is to identify the strengths and weaknesses of passwordless authentication on smartphones. To this end, we determine account-specific adoption barriers by encouraging our participants to reflect on whether they want to use passwordless authentication for their own accounts.

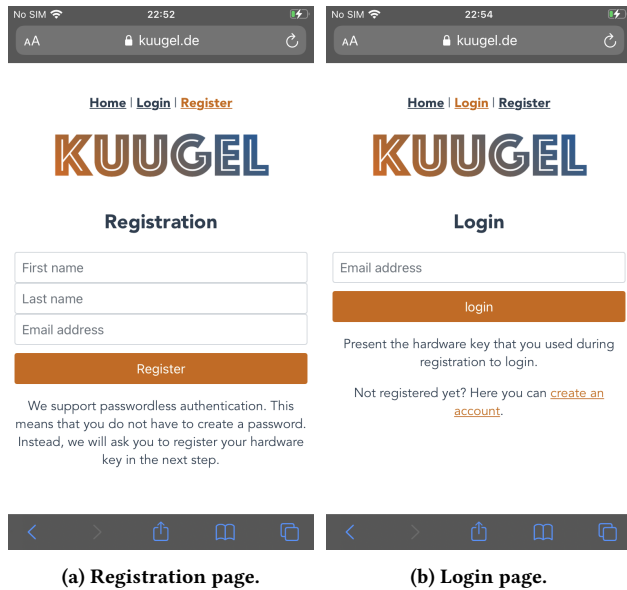


Figure 3: Screenshots of our mockup website's registration and login page. In our experiment, each participant experienced the complete authentication flow of FIDO2 passwordless authentication consisting of account registration and subsequent login.

4 METHODS

To answer our research questions on the usability and acceptance of platform and roaming authentication on smartphones, we conducted a lab study with practical web authentication tasks (Section 4.1) and a follow-up survey containing quantitative and qualitative questions (Section 4.2). We address our study sample (Section 4.3) and ethical considerations (Section 4.4), including the pilot study (Section 4.5), and analysis toolbox (Section 4.6).

4.1 Material

Our lab study's goal was to create a passwordless authentication scenario resembling our participants' everyday lives. In our study, the participants gained hands-on authentication experience on a smartphone, namely the Apple iPhone SE (2nd generation), referred to as *iPhone* for simplicity, running iOS 14.5.1 and Safari 14.1. During our study, this setup represented the most popular mobile operating system in North America [64]. We developed the mockup website KUUGEL as a relying party (Figure 3), which supports FIDO2 registration/login, e.g., using either (1) platform authentication with Touch ID [7] or (2) roaming authentication with a Yubico YubiKey 5C NFC [76], referred to as *YubiKey* for simplicity. The YubiKey supports FIDO2 via an NFC interface. Furthermore, it features optional PIN protection, adding a knowledge-based authentication factor. Farke et al. [23] reported negative user feedback on this feature. Thus, we opted against activating the YubiKey's PIN protection to avoid additional authentication overheads.

4.2 Study Design

We used a between-groups design to identify the differences between platform and roaming authentication on smartphones. Our between-groups design eliminates the influence of order and concentration level with an extended study duration. We decided for a lab study over alternative study designs to ensure consistent conditions for all participants, minimizing confounding factors. We randomly distributed the participants into two groups: The study group used platform authentication with Touch ID (Group P), while the control group used roaming authentication with a YubiKey (Group R). We conducted our lab study in neutral meeting rooms, where the participants met the study conductor in one-on-one sessions¹. Each room contained a chair in front of a desk with an iPhone and, for Group R, also a YubiKey. For all participants, we ensured the same neutral environment and hardware to help participants only evaluate the effect of the presented authentication method. The study conductor was always present during each trial to verify that the participant performed the tasks and watched the educational videos.

Our study consisted of eight stages, which are shown in Figure 4. We first asked participants to read and sign our consent form, which explained the study's goal, procedure, and privacy policy. We then instructed participants in Group P to register one of their fingerprints in the iPhone's Touch ID settings, which was required for the following experiment. We explicitly decided to do this step early before the introductory videos and practical tasks so that the rest of the study procedure was comparable for both groups.

4.2.1 Task Instruction. All participants watched group-specific educational videos (approx. 5 minutes total) before the experiment, introducing FIDO2 passwordless authentication and ensuring that all participants had identical conditions². The video continued with a brief explanation of our mockup website's registration and login process. We decided to use pre-recorded videos for instruction to increase the internal validity of our study by ensuring that all participants received the same information. Both groups watched identically structured videos that only differed in the platform and roaming authentication details, allowing for fair comparison in our between-groups design. Our goal was to avoid misunderstandings resulting in confounding effects, that are invisible in the worst case.

4.2.2 Experiment. After the task instruction, we handed our participants the smartphone (and for Group R, also the YubiKey). The participants used the iPhone to browse our mockup website and register an account using their assigned authenticator type. Our mockup website complies with Apple's recommendations for websites with Touch ID for the Web [7]. Besides the pages necessary for registration and login, the mockup website provides only minimal additional functionality to help participants assess only the usability of FIDO2.

After registration, participants in Group P logged into our mockup website with an email address and their fingerprint, as shown in Figure 1a. Participants in Group R used their email address and the YubiKey for authentication at our mockup website.

¹Protocols were in place for the in-person lab study ensuring the safety of participants and the study conductor, as we conducted our study during the COVID-19 pandemic.

²Our introduction video was partly based on the introductory videos from the study of Lyastani et al. [44].

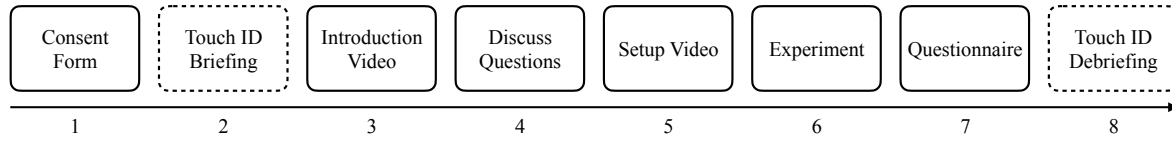


Figure 4: Our lab study’s eight stages. Dashed borders indicate stages that only participants of Group P participated in to register and delete their fingerprints from the iPhone’s Touch ID settings.

The user brings the YubiKey near the iPhone to register a new credential on the YubiKey, as shown in Figure 1b. After registration, the participants logged into their accounts using the same method.

4.2.3 Questionnaire. After collecting the hardware, we asked the participants to fill in our questionnaire to reflect on their experiences during the experiment. Appendix B contains the corresponding questionnaire. We included variables from related work on authentication factors [44, 50, 57, 60] for comparability with previous user studies:

- **DEPENDENT VARIABLES:** The dependent variables captured our participants’ experience and assessment of roaming or platform authentication. To answer RQ1, we used the System Usability Scale (SUS) [12] to determine the authentication methods’ usability, as well as the acceptance scale from van der Laan et al. [67] to measure how much our participants accepted their assigned authentication method. To answer RQ3, we used 11 five-level Likert items to measure how likely our participants would use the demonstrated authentication method on 11 different types of online accounts. As support for FIDO2 is currently limited in practice, we asked our participants to assume that each service supported the presented authentication method. Participants could also select “not available” if they did not use this type of account.
- **GENERAL IMPRESSION:** Our questionnaire included four open-ended text questions to answer RQ2. Participants reflected on their general impression of the studied authentication method, but we also specifically asked our participants to state any strengths and weaknesses that came to their minds. Our fourth text question, which directly follows the quantitative question on adoption for 11 different account types, was designed to learn why participants do (not) want to use the authentication methods for specific account types. Our goal was to recognize account characteristics increasing or decreasing participants’ likelihood of using passwordless authentication and identify adoption barriers for common account types.
- **CONTROL VARIABLES:** As control variables, we measured our participants’ technology affinity using the Affinity for Technology Interaction Scale (ATI) [32] as well as their level of privacy concerns using four Likert items from Langer et al. [42]. Furthermore, we asked participants which 2FA methods they had already used and whether they had prior experience with Apple iOS, which was the operating system running on our study’s smartphone.
- **DEMOGRAPHIC VARIABLES:** We collected the gender, age, education, and field of study/work of our participants as demographic variables.

After filling in the questionnaire, we ensured participants of Group P deleted their fingerprints from the iPhone’s Touch ID settings. Before concluding the study, we thanked the participants for their time. It usually took participants 15-25 minutes to participate in our study.

4.3 Participants

We conducted our lab study with 89 participants between July and September 2021. For recruitment, we used mailing lists, social media groups, word-of-mouth, and snowball sampling, both within and outside our university (participants had to be over 18 years old to be eligible). Interested participants self-registered with a registration form to choose a time slot for voluntary participation in our study. Participants did not receive compensation for their participation in our study, except for one student from our university’s psychology department who received a participation point as part of their study program. Within this study program, the student had sufficient studies to choose from.

Two participants did not fully complete the questionnaire, so we removed them from our final sample ($N = 87$). Of these, 55 (63%) identified as male and 32 (37%) as female, with no participants opting for the “other” or “no answer” options. The education level of our participants was high, with the most common highest educational degree in our sample being a Bachelor’s degree (40%; 35). Our diverse recruitment sources resulted in at least 39 (45%) of non-students. Most participants (59%; 51) were between 20 and 29 years old, but a substantial share of participants (34%; 30) were older than 30, of which 10 participants (11%) were over 50. According to Pearson’s chi-squared test, the demographic structure (gender, age, and education) did not differ significantly between Group P and Group R (Table 7).

4.4 Ethical Concerns

Our university’s institutional review board (IRB) reviewed and approved this study. We informed all participants about the study’s purpose and data collection while adhering to the General Data Protection Regulation (GDPR). We collected written consent prior to the lab study.

During the study, participants in Group P temporarily registered their fingerprints on the iPhone. We briefed the participants about this and ensured the deletion of their fingerprints after each trial. For participants in Group R, we reset the YubiKey to factory settings after each trial to ensure equal conditions for all participants. The names and email addresses for authentication at our mockup website were only temporarily stored on the iPhone and deleted after each trial. We did not collect any other sensitive information. We pseudonymously stored each participant’s answers without

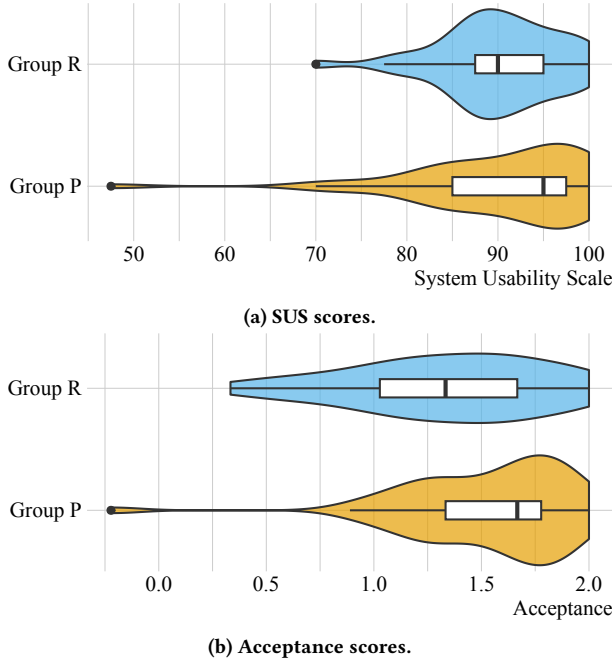


Figure 5: Comparison of our participants' SUS scores and acceptance scores for platform authentication (Group P) and roaming authentication (Group R). The boxplots show quartiles, median, and outliers.

identifying information by assigning sequential numbers. Participation in our study was voluntary and without compensation in accordance with common practice of our university's computer science department.

4.5 Pilot Study

We conducted a pilot study ($N=6$) to identify technical problems with the study setup and ambiguities in the instructional videos and the questionnaire. The pilot study showed that the setup worked reliably and that the participants understood the instructional videos and all categories of the questionnaire.

4.6 Data Analysis

We used the statistical software R 4.2.0 for all quantitative data analysis [56]. We calculated the central tendencies and correlations to answer our research questions. We compared Group P and Group R using the Wilcoxon-Mann-Whitney test [45] for ordinal variables and Pearson's chi-squared test [53] for nominal variables. For multiple tests, we controlled the false discovery rate (FDR) using the Benjamini-Hochberg method [10]. We used Kendall's rank correlation coefficient [39] to determine the relationship between our control and dependent variables. For all statistical tests, we used an alpha level of .05.

Furthermore, we conducted a qualitative analysis of the four open-ended text questions, performing the following qualitative coding steps: (1) Two researchers independently constructed an initial codebook for each question using inductive coding [34, 48].

(2) The researchers merged these codebooks through discussions and formed clusters to identify a suitable level of detail. (3) Two further independent researchers deductively coded all data according to the final codebook. (4) They discussed to resolve coding differences. After discussions, the researchers agreed on most coding decisions and reached satisfactory inter-coder reliability (mean Krippendorff's Alpha = .984, minimum .789 [40]). Finally, the researchers who created the initial codebook discussed and agreed on any remaining coding inconsistencies. Appendix C contains our final coding system.

5 QUANTITATIVE RESULTS

In this section, we report on the evaluation of our questionnaire, which represents the results of our study. The questionnaire (Section 4.2) contains quantitative scales, control variables, and open-ended qualitative questions. We used well-known metrics [12, 67] to study the participants' perceptions of FIDO2 regarding its usability (Section 5.1) and acceptance (Section 5.2). Furthermore, we determined the participants' likelihood to adopt FIDO2 for specific online account types (Section 5.3) and analyzed the control variables (Section 5.4).

5.1 Usability

After getting hands-on experience with passwordless authentication on smartphones, the participants completed our questionnaire, including the SUS [12]. We use the Shapiro-Wilk test [63] to determine whether this variable is normally distributed. The SUS scores for Group R ($W = 0.95$, $p = .058$) are approximately normal, but the scores for Group P ($W = 0.81$, $p < .001$) are significantly non-normal. Figure 5a depicts the statistics of the SUS scores per group. The distribution in Group P has visible negative skew, indicating that most participants selected values on the higher end of the scale. Additionally, there is a single outlier in both groups. Parametric methods such as the t-test assume normally distributed sampling distributions and can give inaccurate results in the presence of outliers [68]. We, therefore, use non-parametric statistical methods for data analysis (Section 4.6).

SUS scores in Group P (Mdn = 95) did not differ significantly from Group R (Mdn = 90) according to the Wilcoxon-Mann-Whitney test, $W = 835.5$, $p = .351$, $r = -.10$. **Both platform and roaming authentication on smartphones have a similarly high level of usability.**

5.2 Acceptance

We collected data on the acceptance of passwordless authentication using the scale of van der Laan et al. [67]. Figure 5b shows the acceptance scores for both groups, which are approximately normal for Group R ($W = 0.96$, $p = .196$), but significantly non-normal for Group P ($W = 0.83$, $p < .001$). As with the SUS scores, the acceptance scores in Group P have strong negative skew and an outlier, which reaffirms our choice of the non-parametric Wilcoxon-Mann-Whitney test. The acceptance scores in Group P (Mdn = 1.7) differed significantly from Group R (Mdn = 1.3), $W = 691$, $p = .031$, $r = -.23$, indicating that **platform authentication has higher acceptance on smartphones than roaming authentication.**

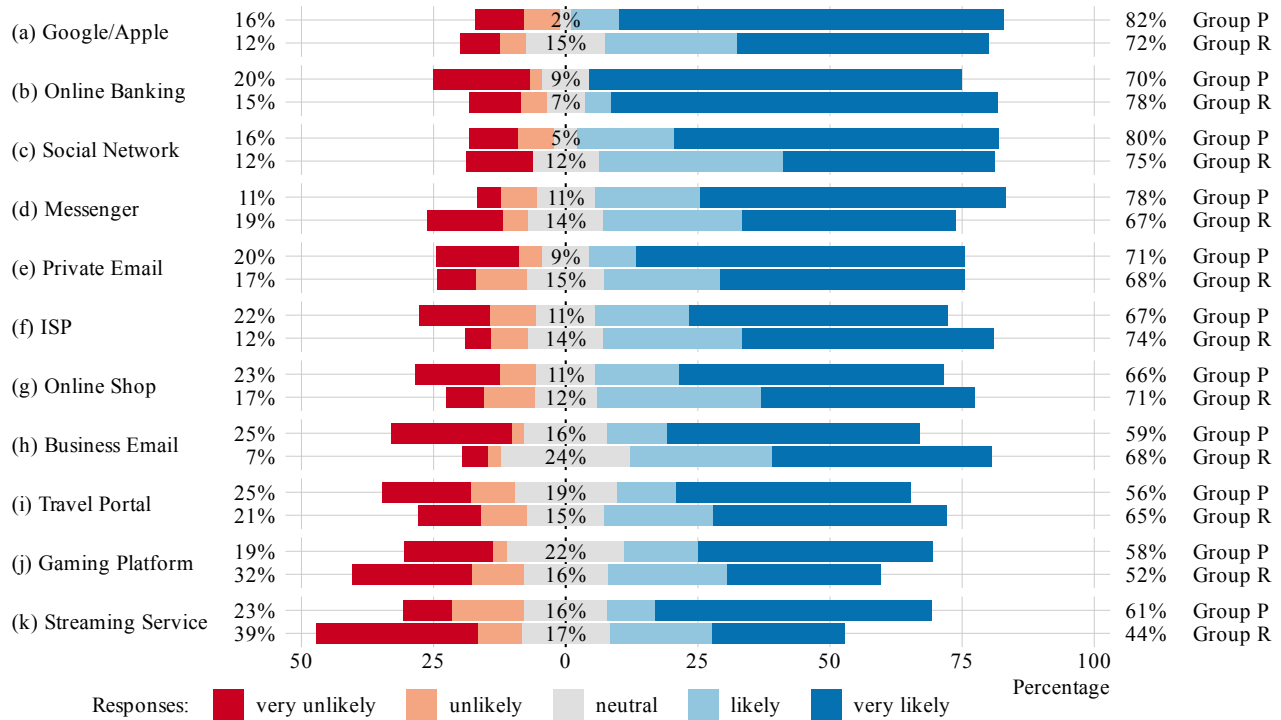


Figure 6: Comparison of our participants' adoption likelihood for different account types ordered by their average likelihood. The percentages (left, middle, right) represent the share of negative, neutral, and positive responses, respectively.

5.3 Adoption

The category dedicated to adoption listed 11 account types and asked participants how likely they were to use the studied passwordless authentication method, provided that the online service supported it. Figure 6 shows the responses for all accounts.

Most participants were optimistic about using passwordless authentication on their smartphones. The three account types with the highest share of positive responses were Google/Apple accounts (77%; 65), online banking accounts (74%; 63), and social network accounts (77%; 65), although the last had less “very likely” responses compared to the former two. Our participants were least likely to use passwordless authentication for streaming services (54%; 43 positive responses) and gaming platforms (55%; 37 positive responses). For every account type, however, the share of positive responses was higher than the share of negative responses, indicating that **overall, our participants want to use passwordless authentication on smartphones when possible.**

We compared the adoption scores of both groups using Pearson's chi-squared test. The adoption did not differ significantly between both groups for any account type after controlling the FDR using the Benjamini-Hochberg method and a target FDR of 10% [10]. We excluded participants without a certain account from the analysis of that account type. This occurred only for a few participants per account, with the following exceptions: Twenty participants did not have a gaming account (11 P + 9 R), 17 participants did not have an account at a travel portal (9 P + 8 R), and 7 participants did not have an account at a streaming service (1 P + 6 R).

5.4 Control Variables

Our control variables are part of the descriptive overview in Table 1. Additionally, Appendix A contains our participants' demographic information. Group P and Group R did not differ significantly for ATI, privacy concerns, and iOS familiarity. We calculate bivariate Kendall's rank correlation coefficients in Table 2 to determine the relationship between the control and dependent variables. As shown, the acceptance scores significantly correlated with the SUS scores ($\tau = .36, p < .001$) and the predictor variable representing the differences between the two groups ($\tau = .20, p = .030$). There also was a significant negative correlation between the ATI scores and iPhone familiarity ($\tau = -.23, p = .009$).

2FA FAMILIARITY. We asked our participants whether they were familiar with any 2FA methods. Most of our participants had prior experience with SMS-based 2FA (87%; 76), push-based smartphone apps (83%; 72), and transaction authentication number (TAN) lists (76%; 66). About half had used one-time password (OTP) generator apps before (51%; 44), and only a few had experience with security keys such as the YubiKey (8%; 7). Group P and Group R did not differ significantly for any of the five 2FA methods.

6 QUALITATIVE RESULTS

The questionnaire contained open-ended text questions to investigate the participants' general experience with passwordless authentication as well as its strengths and weaknesses. We also included an open-ended question to understand better why participants do

Table 1: Comparison of our participants' descriptive data, including the control and dependent variables.

Variable	Group P (N = 45)	Group R (N = 42)	Statistic	ES
iOS Familiarity			$\chi^2(1) = 0.05$	.03
Yes	67% (30)	64% (27)	$p = .815$	
No	33% (15)	36% (15)		
ATI	4.2 (2.2)	4.5 (1.5)	$W = 1072$ $p = .282$	-.12
Privacy Concerns	5.0 (1.8)	5.0 (1.8)	$W = 966$ $p = .861$	-.02
SUS	95 (12.5)	90 (7.5)	$W = 835.5$ $p = .351$	-.10
Acceptance	1.7 (0.4)	1.3 (0.6)	$W = 691$ $p = .031$	-.23

Note: For iOS familiarity, we report in-group percentages (and frequencies), Pearson's chi-square test, and Cramér's V [17] as the ES. For the other scores based on Likert scales, we report the median (and the interquartile range), the Wilcoxon-Mann-Whitney test, and the effect size estimate based on Rosenthal's method [61].

Table 2: Kendall's correlation between control and dependent variables.

	2	3	4	5	Accept.
1 Group (P)	-.10	-.02	.03	.09	.20*
2 ATI		.05	-.23**	.15	.12
3 Privacy Concerns			-.09	-.12	-.14
4 iOS Familiarity				.00	-.02
5 SUS					.36***

Note: N = 87 * $p < .05$ ** $p < .01$ *** $p < .001$

(not) want to adopt passwordless authentication for specific account types. In summary, our participants' general impression was "good" (87% P, 81% R) or "good, but" (13% P, 19% R), with only two participants in Group R not describing a positive general experience. Based on our participants' comments, Section 6.1 summarizes the strengths of FIDO2 (Table 3), and Section 6.2 reports its weaknesses (Table 4). We also quote from our participants' free-text responses to better present the details of their reasoning.

6.1 FIDO2 Strengths

Overall, our participants reacted positively to platform and roaming authentication on smartphones. The SUS scores for both groups were concentrated at the top of the scale (> 90), indicating *excellent* usability [9]. The acceptance scores in both groups were high but significantly higher in Group P, albeit with a small effect size.

In our lab study, participants gained hands-on experience with passwordless authentication on our mockup website. However, we also asked them if they wanted to use the authentication methods on their own accounts, encouraging the participants to reflect on their everyday authentication use cases and whether they would like to use passwordless authentication beyond this user study. There was a clear positive trend for each of the 11 account types, with the majority wanting to adopt passwordless authentication.

The median share of positive responses ("likely" and "very likely" to adopt) across all account types was 69.8%. When we asked users about their general experience with the studied authentication methods, most participants (86%) had a positive impression.

Why did users like the studied authentication methods? The coded responses to the open-ended text questions revealed four main strengths of passwordless authentication related to security, passwords, usability, and authentication times (Table 3).

SECURE. Most participants in both groups commented on the methods' security benefits, e.g., one participant wrote:

"From my point of view, it feels more secure than, for example, a password manager." (P32, Group R)

This is encouraging, as for any new authentication method to succeed, it must not only be secure, but users must also perceive it as secure and trust it [44].

PASSWORD REPLACEMENT. Users generally dislike passwords and the associated overhead of creating, memorizing, and updating them [65]. Many participants stated eliminating this effort as an advantage of passwordless methods.

USABLE. Participants in both groups praised the usability of the studied authentication methods and described them as "intuitive", "simple", and "efficient":

"Easy to use, even for non-technical people." (P44, Group P)

FAST. Participants, mainly in Group P but also in Group R, mentioned that the authentication method is fast.

6.2 FIDO2 Weaknesses

The median share of negative responses ("unlikely" or "very unlikely" adoption) across all account types was 17.6%. Why do some participants not want to adopt FIDO2 authentication on their smartphones, despite security benefits and excellent usability scores? To illustrate this, note that those participants still had a respectable median SUS score of 93.8, demonstrating that usability alone is insufficient for users to adopt FIDO2 in their everyday lives. Instead, there are further adoption barriers, which we now present in more detail. We start with problems common to both platform and roaming authentication and then turn to each of them individually. Table 4 summarizes all weaknesses mentioned by our participants within a smartphone environment.

6.2.1 Platform and Roaming Weaknesses. Overall, in both Group P and Group R, the usability of registration and login on the smartphone was fine, but some participants raised concerns about account recovery and delegation, authenticator revocation, and availability. We identified the following weaknesses in both groups, indicating that these weaknesses are common to both platform and roaming authentication.

ACCOUNT RECOVERY. When users lose access to their accounts, they need a convenient and reliable recovery mechanism. Participants from both groups raised questions about how to proceed in case they suddenly cannot access their accounts:

Table 3: Strengths of passwordless authentication reported by our participants.

	Strength	Group P	Group R
Both	Secure	69%	69%
	No password memorization issues	47%	52%
	Easy to use/Intuitive	38%	33%
	No passwords	16%	14%
	Easier than passwords	16%	12%
	No password creation	7%	5%
	No password updates	4%	2%
Platform	Fast	47%	12%
	Always available	11%	2%
	Privacy	11%	2%
	Easy to setup	4%	2%
Roaming	One solution for many accounts	4%	10%
	Good for lay users	2%	5%
	Security less reliant on smartphone	0%	2%

Note: Numbers in **Group P** and **Group R** are in-group percentages (%). As both groups mostly identify similar strengths of passwordless authentication, Section 6.1 reports them jointly and points out differences.

“If I lose the YubiKey, I will have to look for a replacement, block my accounts, and then register initially with my new YubiKey...” (P20, Group R)

AUTHENTICATOR REVOCATION. The problem of revocation closely relates to recovery. When the smartphone or the security key is stolen, users not only need to recover access, but they also want to revoke the stolen authenticator to prevent abuse. This problem applies to both platform and roaming authentication, but participants from Group P rarely mentioned revocation as an issue.

ACCOUNT DELEGATION. Users are familiar with sharing access to some of their accounts, e.g., streaming services. Furthermore, some users would like to give their partner access to a shared bank account. Account delegation is easy with passwords, but with passwordless authentication (especially platform authentication), there is nothing to share:

“In case another person needs to log in to your account for whatever reason that’s difficult, or if it’s an account that multiple people use [...]” (P34, Group R)

AVAILABILITY. Password-based authentication is available as long as the user remembers their password. In contrast, passwordless authentication can be temporarily unavailable, e.g., when a user forgets to carry their YubiKey (Group R) or the smartphone’s battery is empty (Group P). Participants of both groups worried about the potential inconvenience of authentication being unavailable.

6.2.2 Platform Weaknesses. Even though our study reports good usability of platform authentication, some participants wondered

Table 4: Weaknesses of passwordless authentication reported by our participants.

	Weakness	Group P	Group R
Both	Revocation/Recovery	11%	26%
	Complicated for lay users	13%	5%
	Privacy concerns	4%	10%
	Website compatibility	4%	2%
	Coerced authentication	0%	5%
	Account sharing	4%	0%
	Unfamiliarity	4%	0%
Platform	Use on multiple clients	27%	5%
	Technical problems	20%	5%
	Technology mistrust	13%	0%
	Empty battery	7%	0%
	Biometric security	4%	0%
Roaming	Loss/Destruction	20%	81%
	Something to carry	2%	36%
	Theft	4%	21%
	Costs	0%	12%
	Cumbersome	2%	7%

Note: Numbers in **Group P** and **Group R** are in-group percentages (%).

how to access their accounts from other devices. Others were concerned about hardware issues or did not trust the technology.

MULTIPLE CLIENTS. While passwords are applicable on most client devices, our participants did not know how to access their accounts on devices other than their smartphones, e.g., their personal computer (PC) or laptop. For example, one participant wrote:

“Authentication is bound to the device. For me, it is important that these accounts are accessible from everywhere [...]” (P57, Group P)

Similarly, some participants wondered how to access their accounts on public computers:

“Not usable on all devices, e.g., the PC in the library.” (P73, Group P)

MALFUNCTIONS. Participants were concerned about technical problems with the biometric sensor, preventing them from accessing their accounts. Especially participants who had used biometric authentication before to access their smartphones seemed to be aware of potential problems:

“Sometimes the fingerprint does not work so well when the hands are wet, for example (from my own experience with unlocking the iPhone).” (P71, Group P)

TECHNOLOGY MISTRUST. A few participants were skeptical of platform authentication due to its novelty. They were unsure if they could trust it, as Touch ID felt like a black box to them. For instance, one participant wrote:

“Less visibility into what’s actually happening. It feels like you’re giving up a lot of control.” (P56, Group P)

6.2.3 Roaming Weaknesses. Most of our participants gave good SUS scores to roaming authentication during our lab study, indicating that the general authentication flow is usable. However, our participants raised concerns about the requirement to purchase and carry additional hardware, which could get lost, destroyed, or stolen in the worst case.

ADDITIONAL HARDWARE. Roaming authentication requires additional hardware, which users must remember to carry. The associated physical effort is a hidden usability penalty [31]. While we provided all hardware to our participants in our study, users would need to bring their own authenticator in practice, which requires thought and effort.

LOSS/DESTRUCTION/THEFT. The roaming authenticator is often a small external device, so it can get lost, destroyed, or stolen. The immediate consequence is a lack of availability because users lose access to their accounts. Most participants in Group R (81%) mentioned this as an issue, for example:

“You must always have the [YubiKey] with you → can be lost, forgotten → no login possible. Actually, you always want to have it with you, but then you can lose it, in which case you can no longer log in.” (P80, Group R)

DEPLOYMENT COSTS. Another problem with requiring additional hardware is that it costs money. The NFC-capable FIDO2 roaming authenticators used for our study cost 55 USD at the time of our study [76], but cheaper models are available for 25 USD [75]. Nevertheless, some users are unwilling to pay this much for authentication, especially considering that the best practice is buying a second token as a backup.

“The [YubiKey] costs money, to begin with, and passwords are free. I think this will prevent many users from using it.” (P1, Group R)

For P34, however, the costs were no dealbreaker:

“Has a certain price (but would be worth it to me personally).” (P34, Group R)

7 DISCUSSION

We now discuss the main findings of our study. In Section 7.1, we discuss usage patterns in our participants’ responses that lead to account-specific adoption decisions. Section 7.2 revisits the weaknesses of FIDO2 and guides on how to alleviate major issues. Finally, we address the limitations of our work (Section 7.3) and discuss future work (Section 7.4).

7.1 Adoption Depends on Account Type

Most weaknesses presented in the previous section were only mentioned by a fraction of our participants, emphasizing their positive general impression and indicating that some of these weaknesses

only apply to specific usage patterns or accounts. After the participants stated their likelihood of using passwordless authentication for real-world online accounts, we asked them to explain what affected their decision. Table 5 and Table 6 summarize the characteristics of accounts for which participants (did not) want to use passwordless authentication. We observe that specific account characteristics affect how participants weigh the strengths and weaknesses of passwordless authentication. Our goal is to identify account characteristics that make an account more suitable for either platform or roaming authentication.

7.1.1 Account Sensitivity. We observe opposite trends regarding how users reason about adopting passwordless authentication for sensitive accounts.

SECURITY VS. AVAILABILITY. Some users reject roaming authentication for non-sensitive accounts, where availability concerns outweigh security benefits. These participants cited account sensitivity as a reason for adopting passwordless authentication but its absence as a reason against adoption:

“In principle, I would choose to use the authentication method for most accounts, except for accounts [...] to which I do not assign any importance.” (P32, Group R)

CONVENIENCE VS. MISTRUST. We identify a countertrend of users rejecting passwordless authentication for sensitive accounts, as they cannot put aside their technology mistrust in favor of the usability benefits. These participants cited “easy to use” and “fast” as strengths of passwordless authentication and generally considered it to be “secure”. However, they would instead rely on password-based authentication for sensitive accounts, trusting their memory more than the mechanisms of passwordless authentication:

“I consider the [Yubikey] to be secure. However, for important accounts, I would rather rely on my memory and that the login does not get into the wrong hands.” (P68, Group R)

7.1.2 Usage Frequency. During our study, there was a clear relationship between the adoption likelihood and the usage frequency of an account.

SECURITY VS. PHYSICAL EFFORT. Some users reject roaming authentication for frequently-used accounts because the physical effort outweighs the security benefits. We find that these participants described the YubiKey as “easy to use” and “secure”, but also as “something to carry”:

“For: Accounts that I do not use daily. [...] Not: For Insta and Co. Since I use them several times a day (even on the side or so), and the effort to always use the USB stick is not in proportion to the benefit, i.e., the danger if someone steals my password or hacks my account.” (P80, Group R)

SPEED VS. INITIAL EFFORT. We observe the opposite trend for platform authentication: Some users reject platform authentication for rarely-used accounts because the effort to set up passwordless authentication outweighs the reduced authentication times. This

Table 5: Account characteristics favoring adoption.

Yes, for ...	Group P	Group R
sensitive accounts	16%	38%
rarely used accounts	0%	7%
accounts that are not shared	4%	2%
frequently used accounts	4%	2%
business-related accounts	2%	2%
accounts where fast access is crucial	4%	0%
non-sensitive accounts	2%	0%
accounts mainly used on smartphone	2%	0%
accounts not used on other devices	2%	0%

Note: Numbers in **Group P** and **Group R** are in-group percentages (%).

trend is likely related to our previous observation that more participants perceived platform authentication as fast compared to roaming authentication:

“For: accounts where you have to log in more often (often on the same device). Against: social networks: login [only once]” (P72, Group P)

7.1.3 Mobility. We observe that some users reject passwordless authentication for accounts used outside of their homes because, in these scenarios, the chances of authentication unavailability increase. These participants worried about an “empty battery” (Group P) or described the need to remember to carry the roaming authenticator (Group R):

“For streaming services, I may log in from somewhere else than home, and then, if I don’t have the [YubiKey] with me, I won’t be able to log in.” (P30, Group R)

7.2 Roadmap to Addressing the Weaknesses of FIDO2

We now discuss the weaknesses identified in our study and propose how to address them while considering the findings of previous user studies in different passwordless authentication scenarios (Section 2). Some weaknesses, including account recovery, authenticator revocation, and account delegation, affect both platform and roaming authentication, but other issues directly stem from the concrete authentication mode.

7.2.1 Account Recovery. Recovery is an open problem in FIDO2 [23, 44, 51, 57], which the standards currently do not sufficiently address. The best practice is to register multiple authenticators per account and keep one as a backup, consolidating the issues of acquisition costs and additional hardware. However, this does not scale. Furthermore, our analysis shows that users need clear instructions on dealing with recovery as they did not realize that mitigations known from password-based authentication can also work for FIDO2. For instance, online services can allow users to request instructions for authentication reset via email and enroll new FIDO2 credentials, similar to a password reset form. The usability and acceptance of this form of recovery require further research.

Table 6: Account characteristics preventing adoption.

No, for ...	Group P	Group R
non-sensitive accounts	11%	21%
sensitive accounts	11%	10%
shared accounts	7%	10%
frequently used accounts	0%	10%
accounts used outside of home	4%	5%
business-related accounts	7%	2%
rarely used accounts	7%	2%
accounts supposed to be anonymous	0%	2%
accounts used on multiple devices	2%	0%

Note: Numbers in **Group P** and **Group R** are in-group percentages (%).

Some works propose to improve recovery in FIDO2 with extensions [33, 55, 74], certificates [15], or electronic IDs [62].

7.2.2 Authenticator Revocation. Authenticator revocation has been identified as a weakness in prior studies of FIDO2 roaming authentication [23, 44, 51, 57] but not in related studies of platform authentication. A possible explanation is that there already exist mitigations to deal with the loss of smartphones. For instance, platform authentication often requires local authentication, making it harder for attackers to access accounts with stolen smartphones. Additionally, users can remotely erase the data on a lost iPhone [6], further addressing revocation. Most roaming security keys do not support local authentication, although some models are equipped with biometric protection [24, 77]. However, as the FIDO2 standards do not sufficiently solve authenticator revocation, web services (i.e., FIDO2 relying parties) need to provide account management options to revoke access from specific authenticators.

7.2.3 Account Delegation. While giving your authenticator to someone else is possible, it only allows account delegation to one person at a time and does not work remotely [43, 44, 57]. As FIDO2 supports the registration of multiple authenticators for a single account [69], we argue that account delegation can become a strength of FIDO2 rather than a weakness once online services start using this feature.

7.2.4 Platform Authentication. Our study confirms that the main weakness of platform authentication is the use on multiple clients [43]. While FIDO2 allows users to add multiple authenticators per account, the individual registration of all devices for each account imposes an additional burden on users. Furthermore, it does not work in special authentication environments, e.g., public computers in a library where roaming authentication would be better suited.

Platform authentication requires trusting the involved hardware components to handle the key material. Even if the involved protocols and components are proven secure, they are often implemented as blackboxes, and some TPMs had vulnerabilities in the past [37, 49]. Many users mistrust such novel, unfamiliar technologies [23, 43, 44, 51, 57, 66], which is a dealbreaker for the acceptance of an authentication method. One potential solution is better education on how platform authentication, FIDO2, and public-key-based authentication methods work.

7.2.5 Roaming Authentication. The primary weaknesses of roaming authentication are by design: Roaming authenticators require additional hardware that users have to purchase [11, 23, 44], they have to be carried around [23, 44, 51, 57], and they can be lost, destructed and stolen [23, 43, 44, 51, 57]. These weaknesses can also apply to smartphones as physical devices, but more participants in Group R described these issues, indicating that the perceived risk is higher for roaming authentication. Furthermore, smartphones can mitigate these issues, which is infeasible with the limited capabilities of roaming authenticators: Some smartphone vendors allow users to locate their smartphones, such as via Apple Find My [4] or Android Find My Device [35], which mitigates the consequences of loss and theft. Additionally, platform authentication does not cause additional deployment costs, as it runs directly on the user's smartphone, which most users already own [54].

7.3 Limitations

This section addresses limitations of our work as a result of our recruitment and study design.

7.3.1 Task Instruction. Our instructional videos ensured that all participants had a basic understanding of passwordless authentication, resulting in a study sample that is likely better informed about FIDO2 than today's general public. This limits the ecological validity of our results for the first-time user experience of users unaware of passwordless authentication. A previous user study by Lassak et al. showed, however, that basic user training can address such misconceptions of first-time users [43]. In contrast, usability problems in the day-to-day user interaction of FIDO2 cannot be solved with education alone and are thus a more interesting scenario for our study. We argue that briefing our participants supports studying the usability of FIDO2 for informed lay users, which is increasingly representative given the ongoing proliferation of FIDO2 [46]. The instructional videos affected Group P and Group R in the same way, as both of our study groups watched identical videos except for the details of platform and roaming authentication. Thus, our participants all had identical conditions, allowing for fair comparison in our between-groups design.

7.3.2 Recruitment. Our sample was comparatively young, with a larger-than-average share of students and most participants in the 20-39 age range. Furthermore, not compensating participants for their participation in our study may have introduced a skew towards participants who are wealthy enough to participate in lab studies for free. Overall, our sample shows a slight deviation from the average population in terms of demographics but, as we argue, without significant effect on the measured items. Rather than selectively sampling participants to achieve a representative distribution, we instead controlled for factors potentially affecting usability and acceptance (Section 5.4). Additionally, our instructional videos helped balance prior knowledge, further mitigating the effects of our unfiltered sample. Neither our control variables nor age correlated significantly with usability or acceptance. While our sample only includes 10 participants (11%) older than 50, there have been related usability studies focusing on older adults' (older than 60) interaction with roaming authentication [19].

7.3.3 Mobile Operating System. We conducted our study on Apple iOS, which at the time of our study represented the most popular mobile operating system in North America [64]. As the FIDO2 user interaction is the same on iOS and Android, except for the wording and illustration in the web browser prompts in Safari and Chrome, we argue that our results also generalize to Android smartphones. Furthermore, none of the strengths, weaknesses, or adoption decisions stated by our participants directly relate to iOS but are equally relevant to FIDO2 on Android.

7.4 Outlook and Future Work

Before concluding our work, we dare a look into the future of FIDO2 and provide recommendations for future research, as well as for the FIDO Alliance and FIDO2 users.

7.4.1 User Education. Introducing our participants to the basics of passwordless authentication and the setup of FIDO2 might have led to improved usability compared to related studies, which should encourage the FIDO Alliance and website operators to strive for better FIDO2 education for the general public. Nevertheless, we found that our participants described several weaknesses that are actually solvable with mechanisms of the FIDO2 standards, e.g., account delegation or authenticator reset. The FIDO Alliance should prioritize explicitly informing users about these mechanisms of FIDO2 to mitigate such misconceptions [43], e.g., they could publish videos similar to our instructional videos (Section 4.2).

7.4.2 Compatibility. Our study shows that users consider a variety of client devices (e.g., laptops, smartphones, smart TVs, game consoles, wearables) when reflecting on their authentication use cases. As a consequence of our work, future authentication methods should be compatible with versatile authentication environments to qualify for the users' needs.

7.4.3 Platform vs. Roaming. Proper consideration of other FIDO2 weaknesses requires adding technical solutions to the FIDO2 standards or developing improved authenticators. For future research and industry, we encourage improving platform authentication because its main weaknesses (Section 6.2.2) are more easily addressable than those of roaming authenticators (Section 6.2.3).

Similarly, we suggest that relying parties (i.e., website operators) expand support for FIDO2 authentication. All modern smartphones support FIDO2 platform authentication nowadays, so FIDO2 is a viable second factor in addition to passwords or even a suitable single factor replacing passwords. Users can benefit from FIDO2's easy-to-use and secure authentication flow without too much overhead for website operators, as FIDO2's browser API is straightforward to implement. By relying on platform or roaming authentication, website operators also avoid storing users' passwords, eliminating a potential liability in case of data breaches.

A combination of smartphone platform authentication with the ideas of using smartphones as roaming authenticators for external devices [52, 57] would be promising to address most weaknesses identified in this study. For instance, Apple has announced *passkeys* for iOS 16 [5], replacing Touch ID for the Web as the iPhone's FIDO2 platform authenticator [8].

Passkeys are FIDO2 key credentials that are synchronized between the user's Apple devices. The usability results of our study

regarding FIDO2 platform authentication with Touch ID also apply to passkeys. Furthermore, passkeys address several platform authentication-related weaknesses identified for Touch ID, e.g., allowing the use of passkeys on multiple clients and enabling users to authenticate on other Apple devices (by connecting their own Apple device as a roaming authenticator). Additionally, the synchronization of passkeys in the iCloud Keychain simplifies FIDO2 account recovery. While passkeys are a promising solution within the Apple environment, the FIDO Alliance should strive for an open system that is usable on all platforms independent of a single vendor's infrastructure.

8 CONCLUSION

We conducted a between-groups lab study (N=87) of FIDO2 passwordless authentication, comparing roaming and platform authentication on smartphones. Our main goal was to identify the strengths and weaknesses of passwordless authentication on smartphones as perceived by lay users, focusing on account-specific adoption barriers. Our key findings for each research question are as follows:

RQ1. *"What is the usability and acceptance of FIDO2 passwordless authentication using platform and roaming authentication on smartphones?"* Both platform and roaming authentication show the potential of satisfactory day-to-day usability on smartphones, but less common events such as device malfunctions, account recovery, and account delegation impair the experience. Overall, users slightly prefer platform authentication.

RQ2. *"What benefits and concerns do users consider when using FIDO2 passwordless authentication on smartphones?"* Users appreciate platform and roaming authentication as simple and secure password replacements without the overhead of memorizing and managing passwords for each account. The primary weakness identified by users is the loss/theft/destruction of the authenticator and the associated burden of revoking and recovering access to each account. For roaming authentication, users criticize having to carry an additional device. In contrast, for platform authentication, users are concerned with accessing their accounts on other client devices and technical problems with the biometric sensor.

RQ3. *"Which account types do users want to secure using FIDO2 passwordless authentication on smartphones?"* While most users are likely to adopt passwordless authentication for their accounts, users prioritize usability, security, and availability differently depending on the account type. As a result, the weaknesses of passwordless authentication turn into dealbreakers for specific account types and usage patterns. Participants generally preferred platform over roaming authentication for non-sensitive accounts or frequently used accounts due to the excellent availability and the fast authentication time of the smartphone's integrated platform authenticator. Conversely, participants preferred roaming authentication for sensitive or rarely used accounts. Some users reject passwordless authentication for shared accounts as they do not see an option to delegate access to other users. Overall, most users prefer passwordless authentication for sensitive accounts. However, despite the usability benefits, some do not fully trust the technology and would therefore only use it for non-sensitive accounts.

In summary, although there is no one-size-fits-all authenticator for all account types, we recommend improving platform authentication, which has more easily addressable weaknesses than roaming authentication.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their helpful suggestions. Furthermore, we thank Annemarie Mattmann for her support during the qualitative data analysis, and Jasim Machkour for the helpful discussions on robust data analysis. This work has been co-funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center and the Federal Ministry of Education and Research of Germany in the project Open6GHub (grant number: 16KISK014).

AVAILABILITY

Together with this paper, we release a replication package with our evaluation scripts and the pseudonymized dataset generated in our study [71] consisting of usability and acceptance scores, the adoption likelihood for 11 account types, and 9 control variables for each of our 87 participants. We also provide our mockup website's source code [70] to facilitate future work.

REFERENCES

- [1] Fatima Alqubaisi, Ahmad Samer Wazan, Liza Ahmad, and David W Chadwick. 2020. Should We Rush to Implement Password-less Single Factor FIDO2 based Authentication?. In *12th Annual Undergraduate Research Conference on Applied Computing (URC)*.
- [2] Apple. 2020. About iOS 13 Updates. <https://support.apple.com/en-us/HT210393#133>. [Online; accessed 2022-06-07].
- [3] Apple. 2020. Safari 14 Release Notes. <https://developer.apple.com/documentation/safari-release-notes/safari-14-release-notes>. [Online; accessed 2022-06-07].
- [4] Apple. 2022. iCloud Find My. <https://www.apple.com/icloud/find-my>. [Online; accessed 2022-06-07].
- [5] Apple. 2022. iOS 16 Release Notes. <https://developer.apple.com/documentation/ios-ipados-release-notes/ios-16-release-notes/>. [Online; accessed 2022-09-14].
- [6] Apple. 2022. Locate a Lost or Stolen Device, Remotely Erase a Device. <https://support.apple.com/en-us/HT210515>. [Online; accessed 2022-06-07].
- [7] Apple Developers. 2020. Meet Face ID and Touch ID for the web. <https://developer.apple.com/videos/play/wwdc2020/10670/>. [Online; accessed 2022-06-07].
- [8] Apple Developers. 2022. Meet passkeys. <https://developer.apple.com/videos/play/wwdc2022/10092/>. [Online; accessed 2022-09-07].
- [9] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of usability studies* (2009).
- [10] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* (1995).
- [11] Joseph Bonneau, Cormac Herley, Paul C Van Oorschot, and Frank Stajano. 2012. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *2012 IEEE Symposium on Security and Privacy*.
- [12] John Brooke. 1996. SUS: A Quick and Dirty Usability Scale. *Usability Evaluation in Industry* (1996).
- [13] Stéphane Ciolano, Simon Parkin, and Paul Dunphy. 2019. Of Two Minds about Two-Factor: Understanding Everyday FIDO U2F Usability through Device Comparison and Experience Sampling. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS)*.
- [14] Jessica Colnago, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Lorrie Cranor, and Nicolas Christin. 2018. "It's not actually that horrible": Exploring Adoption of Two-Factor Authentication at a University. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*.
- [15] James Connors, Stephen Derbidge, Corey Devenport, Natalie Farnsworth, Kyler Gates, Stephen Lambert, Christopher McClain, Parker Nichols, and Daniel Zappala. 2022. Let's Authenticate: Automated Certificates for User Authentication. In *Network and Distributed Systems Security (NDSS) Symposium 2022*.
- [16] James Connors and Daniel Zappala. 2019. Let's authenticate: Automated cryptographic authentication for the web with simple account recovery. In *Who Are You?! Adventures in Authentication Workshop (WAY)*.
- [17] Harald Cramér. 1946. *Mathematical Methods of Statistics*. Princeton University Press.

- [18] Sanchari Das, Andrew Dingman, and L Jean Camp. 2018. Why Johnny Doesn't Use Two Factor: A Two-Phase Usability Study of the FIDO U2F Security Key. In *International Conference on Financial Cryptography and Data Security*.
- [19] Sanchari Das, Andrew Kim, Ben Jelen, Joshua Streiff, L Jean Camp, and Lesa Huber. 2019. Towards Implementing Inclusive Authentication Technologies for Older Adults. In *Who Are You? Adventures in Authentication Workshop (WAY)*.
- [20] Sanchari Das, Gianpaolo Russo, Andrew C Dingman, Jayati Dev, Olivia Kenny, and L Jean Camp. 2018. A Qualitative Study on Usability and Acceptability of Yubico Security Key. In *Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust*.
- [21] Sanchari Das, Bingxing Wang, Andrew Kim, and L Jean Camp. 2020. MFA is a Necessary Chore! Exploring User Mental Models of Multi-Factor Authentication Technologies. In *53rd Hawaii International Conference on System Sciences*.
- [22] Rachna Dhamija, J. D. Tygar, and Marti Hearst. 2006. Why Phishing Works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- [23] Florian M. Farke, Lennart Lorenz, Theodor Schnitzler, Philipp Markert, and Markus Dürmuth. 2020. "You still use the password after all"—Exploring FIDO2 Security Keys in a Small Company. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS)*.
- [24] Feitian. 2018. BioPass FIDO Series Biometric Security Keys. <https://www.ftsafe.com/products/FIDO/BIO>. [Online; accessed 2022-06-07].
- [25] Feitian. 2020. iePass FIDO Series iOS Security Key. <https://www.ftsafe.com/products/FIDO/iOS>. [Online; accessed 2022-06-07].
- [26] FIDO Alliance. 2017. FIDO U2F Raw Message Formats. <https://fidoalliance.org/specs/fido-u2f-v1.2-ps-20170411/fido-u2f-raw-message-formats-v1.2-ps-20170411.html>. [Online; accessed 2022-06-07].
- [27] FIDO Alliance. 2019. Android now FIDO2 certified, accelerating global migration beyond passwords. <https://fidoalliance.org/android-now-fido2-certified-accelerating-global-migration-beyond-passwords/>. [Online; accessed 2022-06-07].
- [28] FIDO Alliance. 2021. Client to Authenticator Protocol (CTAP). <https://fidoalliance.org/specs/fido-v2.1-ps-20210615/fido-client-to-authenticator-protocol-v2.1-ps-20210615.html>. [Online; accessed 2022-06-07].
- [29] FIDO Alliance. 2022. FIDO Alliance Member Companies & Organizations. <https://fidoalliance.org/members/>. [Online; accessed 2022-06-07].
- [30] FIDO Alliance. 2022. FIDO Alliance Specifications Overview. <https://fidoalliance.org/specifications/>. [Online; accessed 2022-06-07].
- [31] Mikhail Fomichev, Julia Hesse, Lars Almon, Timm Lippert, Jun Han, and Matthias Hollick. 2021. Fastzip: Faster and more secure zero-interaction pairing. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. 440–452.
- [32] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human-Computer Interaction* (2019).
- [33] Nick Frymann, Daniel Gardham, Franziskus Kiefer, Emil Lundberg, Mark Manulis, and Dain Nilsson. 2020. Asynchronous Remote Key Generation: An Analysis of Yubico's Proposal for W3C WebAuthn. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. New York, NY, USA. <https://doi.org/10.1145/3372297.3417292>
- [34] Dennis A. Gioia, Kevin G. Corley, and Aimee L. Hamilton. 2013. Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods* 16 (2013).
- [35] Google. 2022. Find, lock, or erase a lost Android device. <https://support.google.com/accounts/answer/6160491>. [Online; accessed 2022-09-08].
- [36] GoTrustID. 2022. Smart Badge Authenticator – Idem Card. <https://www.gotrustid.com/idem-card>. [Online; accessed 2022-06-07].
- [37] Seunghun Han, Wook Shin, Jun-Hyeok Park, and HyoungChun Kim. 2018. A bad dream: Subverting trusted platform module while you are sleeping. In *27th USENIX Security Symposium*.
- [38] Blake Ives, Kenneth R. Walsh, and Helmut Schneider. 2004. The Domino Effect of Password Reuse. *Commun. ACM* 47 (2004).
- [39] Maurice G. Kendall. 1938. A New Measure of Rank Correlation. *Biometrika* 30 (1938).
- [40] Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology*. Sage Publications, Inc.
- [41] Juan Lang, Alexei Czeskis, Dirk Balfanz, Marius Schilder, and Sampath Srinivas. 2016. Security Keys: Practical Cryptographic Second Factors for the Modern Web. In *International Conference on Financial Cryptography and Data Security*.
- [42] Markus Langer, Cornelius J König, and Andromachi Fiti. 2018. Information as a Double-Edged Sword: The Role of Computer Experience and Information on Applicant Reactions Towards Novel Technologies for Personnel Selection. *Computers in Human Behavior* (2018).
- [43] Leona Lassak, Annika Hildebrandt, Maximilian Golla, and Blase Ur. 2021. "It's Stored, Hopefully, on an Encrypted Server": Mitigating Users' Misconceptions About FIDO2 Biometric WebAuthn. In *30th USENIX Security Symposium*.
- [44] Sanam Ghorbani Lyastani, Michael Schilling, Michaela Neumayr, Michael Backes, and Sven Bugiel. 2020. Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication. In *2020 IEEE Symposium on Security and Privacy*.
- [45] Henry B. Mann and Donald R. Whitney. 1947. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* 18 (1947).
- [46] MDN. 2022. Web Authentication API: Browser compatibility. https://developer.mozilla.org/en-US/docs/Web/API/Web_Authentication_API#browser_compatibility. [Online; accessed 2022-06-07].
- [47] Nathan Mercer. 2018. Extending the capabilities of Windows Hello. <https://techcommunity.microsoft.com/t5/windows-blog-archive/extending-the-capabilities-of-windows-hello/ba-p/166534>. [Online; accessed 2022-06-07].
- [48] Sharan B. Merriam and Elizabeth J. Tisdell. 2015. *Qualitative Research: A Guide to Design and Implementation*. John Wiley & Sons.
- [49] Matus Nemec, Marek Sys, Petr Svenda, Dusan Klinec, and Vashek Matyas. 2017. The return of Coppersmith's attack: Practical factorization of widely used RSA moduli. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*.
- [50] Wataru Oogami, Hidehito Gomi, Shuji Yamaguchi, Shota Yamanaka, and Tatsuru Higurashi. 2020. Observation Study on Usability Challenges for Fingerprint Authentication Using WebAuthn-enabled Android Smartphones. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS)*.
- [51] Kentrell Owens, Olabode Anise, Amanda Krauss, and Blase Ur. 2021. User Perceptions of the Usability and Security of Smartphones as FIDO2 Roaming Authenticators. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS)*.
- [52] Kentrell Owens, Blase Ur, and Olabode Anise. 2020. A Framework for Evaluating the Usability and Security of Smartphones as FIDO2 Roaming Authenticators. In *Who Are You? Adventures in Authentication Workshop (WAY)*.
- [53] Karl Pearson. 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50 (1900).
- [54] Pew Research Center. 2021. Mobile Fact Sheet. <https://www.pewresearch.org/internet/fact-sheet/mobile/>. [Online; accessed 2022-06-07].
- [55] Florentin Putz, Steffen Schön, and Matthias Hollick. 2021. Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions. In *Emerging Technologies for Authorization and Authentication (ETAA)*. Springer.
- [56] R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- [57] Brian Rasmussen. 2021. *A Usability Study of FIDO2 Roaming Software Tokens as a Password Replacement*. Master's thesis. Brigham Young University.
- [58] Ken Reese, Trevor Smith, Jonathan Dutton, Jonathan Armknecht, Jacob Cameron, and Kent Seamons. 2019. A Usability Study of Five Two-Factor Authentication Methods. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS)*.
- [59] Joshua Reynolds, Nikita Samarin, Joseph Barnes, Taylor Judd, Joshua Mason, Michael Bailey, and Serge Egelman. 2020. Empirical Measurement of Systemic 2FA Usability. In *29th USENIX Security Symposium*.
- [60] Joshua Reynolds, Trevor Smith, Ken Reese, Luke Dickinson, Scott Ruoti, and Kent Seamons. 2018. A Tale of Two Studies: The Best and Worst of YubiKey Usability. In *2018 IEEE Symposium on Security and Privacy (SP)*.
- [61] Robert Rosenthal. 1991. *Meta-Analytic Procedures for Social Research*. Sage Publications, Inc.
- [62] Fabian Schwarz, Khue Do, Gunnar Heide, Lucjan Hanzlik, and Christian Rossow. 2022. FeDo: Recoverable FIDO2 Tokens Using Electronic IDs. In *2022 ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
- [63] Samuel S. Shapiro and Martin B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (1965).
- [64] Statista. 2021. Market share of mobile operating systems in North America from January 2018 to June 2021. <https://www.statista.com/statistics/1045192/share-of-mobile-operating-systems-in-north-america-by-month/>. [Online; accessed 2022-06-07].
- [65] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2015. "I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab. In *Eleventh Symposium On Usable Privacy and Security (SOUPS)*.
- [66] Leon van den Boogaard. 2022. *User understanding and user acceptance of biometric authentication on mobile phones*. Master's thesis. Radboud University.
- [67] Jinke D. Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A Simple Procedure for the Assessment of Acceptance of Advanced Transport Telematics. *Transportation Research Part C: Emerging Technologies* (1997).
- [68] Rand Wilcox. 2017. *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier, Academic Press.
- [69] World Wide Web Consortium. 2021. Web Authentication: An API for Accessing Public Key Credentials Level 2. <https://www.w3.org/TR/2021/REC-webauthn-2-20210408/>. [Online; accessed 2022-06-07].
- [70] Leon Würsching, Florentin Putz, Steffen Haesler, and Matthias Hollick. 2023. *FIDO2 The Smartphone: Mockup Website for Platform and Roaming Authentication on Smartphones*. <https://github.com/seemoo-lab/fido2-the-smartphone>

- [71] Leon Würsching, Florentin Putz, Steffen Haesler, and Matthias Hollick. 2023. *Lab Study Dataset: FIDO2 Platform and Roaming Authentication on Smartphones*. <https://doi.org/10.5281/zenodo.7572697>
- [72] Jeff Yan, Alan Blackwell, Ross Anderson, and Alasdair Grant. 2004. Password Memorability and Security: Empirical Results. *IEEE Security Privacy* 2 (2004).
- [73] Yubico. 2019. Yubico Launches the World's First Lightning-Compatible Security Key, the YubiKey 5Ci. <https://www.yubico.com/press-releases/yubico-launches-the-worlds-first-lightning-compatible-security-key-the-yubikey-5ci/>. [Online; accessed 2022-06-07].
- [74] Yubico. 2022. GitHub: webauthn-recovery-extension. <https://github.com/Yubico/webauthn-recovery-extension>. [Online; accessed 2022-06-07].
- [75] Yubico. 2022. Security Key NFC. <https://www.yubico.com/us/product/security-key-nfc-by-yubico/>. [Online; accessed 2022-06-07].
- [76] Yubico. 2022. YubiKey 5C NFC. <https://www.yubico.com/us/product/yubikey-5c-nfc/>. [Online; accessed 2022-06-07].
- [77] Yubico. 2022. YubiKey Bio - FIDO Edition. <https://www.yubico.com/us/product/yubikey-bio/>. [Online; accessed 2022-06-07].

A DEMOGRAPHICS

In this section, we extend our participants' descriptive data (Table 1) with further demographic information (Table 7).

Table 7: Our participants' demographic data. We report in-group percentages (and frequencies), Pearson's chi-square test, and Cramér's V as the effect size (ES).

	Variable	Group P (N = 45)	Group R (N = 42)	Statistic	ES
Gender	Female	44.4 (20)	28.6 (12)	$\chi^2(1) = 2.35$ $p = .125$	.16
	Male	55.6 (25)	71.4 (30)		
	Other	0 (0)	0 (0)		
	No answer	0 (0)	0 (0)		
Age	18-19	6.7 (3)	7.1 (3)	$\chi^2(5) = 4.94$ $p = .423$	.11
	20-29	48.9 (22)	69.0 (29)		
	30-39	28.9 (13)	14.3 (6)		
	40-49	2.2 (1)	0 (0)		
	50-59	11.1 (5)	7.1 (3)		
	60-69	2.2 (1)	2.4 (1)		
Education	Still in school	4.4 (2)	0 (0)	$\chi^2(6) = 3.89$ $p = .692$	.09
	Middle school	8.9 (4)	9.5 (4)		
	High school	37.8 (17)	26.2 (11)		
	Bachelor	35.6 (16)	45.2 (19)		
	Master	8.9 (4)	11.9 (5)		
	Doctorate	2.2 (1)	2.4 (1)		
	Other	2.2 (1)	4.8 (2)		

B QUESTIONNAIRE

We provide a translation of our questionnaire, with descriptive question names for convenience. The actual questionnaire only had non-descriptive section names ("Category 1" and so on) so as to not influence the participants. We gave the questionnaire to the participants in the native language of the country where we ran the study.

USABILITY. Please mark the answer that reflects your immediate response to each statement. Please do not think too long about each statement and be sure to provide an answer to all statements.

(Five responses ranging from "Strongly disagree" to "Strongly agree")

(1) I think that I would like to use this system frequently. (2) I found the system unnecessarily complex. (3) I thought the system

was easy to use. (4) I think that I would need the support of a technical person to be able to use this system. (5) I found the various functions in this system were well integrated. (6) I thought there was too much inconsistency in this system. (7) I would imagine that most people would learn to use this system very quickly. (8) I found the system very cumbersome to use. (9) I felt very confident using the system. (10) I needed to learn a lot of things before I could get going with this system.

ACCEPTANCE. Now please evaluate the system. To do this, read each pair of words carefully and make one cross per line.

(Five responses in between the words)

(1) Useless vs. Useful (2) Pleasant vs. Unpleasant (3) Bad vs. Good (4) Nice vs. Annoying (5) Effective vs. Superfluous (6) Irritating vs. Likeable (7) Assisting vs. Worthless (8) Undesirable vs. Desirable (9) Raising Alertness vs. Sleep-Inducing

PRIVACY CONCERNS. Please indicate how strongly you agree with the following statements

(Seven responses ranging from "Strongly disagree" to "Strongly agree")

(1) I am concerned that companies are collecting too much personal information about me. (2) I am concerned about my privacy. (3) To me it is important to keep my privacy intact. (4) Novel technologies are threatening privacy increasingly.

TECHNOLOGY AFFINITY. In the following questionnaire, we will ask you about your interaction with technical systems. The term "technical systems" refers to apps and other software applications, as well as entire digital devices (e.g., mobile phone, computer, TV, car navigation).

(Six responses from "Completely disagree" to "Completely agree")

(1) I like to occupy myself in greater detail with technical systems. (2) I like testing the functions of new technical systems. (3) I predominantly deal with technical systems because I have to. (4) When I have a new technical system in front of me I try it out intensively. (5) I enjoy spending time becoming acquainted with a new technical system. (6) It is enough for me that a technical system works; I don't care how or why. (7) I try to understand how a technical system exactly works. (8) It is enough for me to know the basic functions of a technical system. (9) I try to make full use of the capabilities of a technical system.

iOS FAMILIARITY. (1) Do you use an iPhone for private or business purposes? *(Yes / No)*

OPEN-ENDED QUESTIONS. (1) How would you describe your general experience with the presented authentication method? *(Free text response)* (2) Which advantages do you see in the usage of the presented authentication method? *(Free text response)* (3) Which disadvantages do you see in the usage of the presented authentication method? *(Free text response)*

ADOPTION. In the following, think of a website or account that you use yourself. Assume that the service to which this account belongs supports the presented authentication method. How likely is it that you would use the presented authentication method for this account?

(For items 1-11, there are five responses ranging from "very unlikely" to "very likely", with an additional response for "Not available")

Table 8: Codebooks created during the qualitative analysis (Section 4.6) of the open-ended text questions.

(a) Strengths	
Usability	⟨Easy to use/Intuitive⟩ ⟨Good for lay users⟩ ⟨Fast⟩ ⟨Easy to setup⟩
Availability	⟨One solution for many accounts⟩ ⟨Always available⟩
Cognitive effort	⟨No passwords⟩ ⟨No password creation⟩ ⟨No password updates⟩ ⟨Easier than passwords⟩ ⟨No password memorization issues⟩
Security	⟨Secure⟩ ⟨Security less reliant on smartphone⟩ ⟨Privacy⟩ ⟨Security less reliant on website⟩
(b) Weaknesses	
User experience	⟨Slow⟩ ⟨Complicated for lay users⟩ ⟨Cumbersome⟩
Deployability	⟨Costs⟩ ⟨Website compatibility⟩
Availability	⟨Technical problems⟩ ⟨Something to carry⟩ ⟨Empty battery⟩ ⟨Loss/ Destruction⟩
Mental models	⟨Unfamiliarity⟩ ⟨Revocation/Recovery⟩ ⟨Account sharing⟩ ⟨Technology mistrust⟩ ⟨Use on multiple clients⟩
Security	⟨Theft⟩ ⟨Coerced authentication⟩ ⟨Biometric security⟩ ⟨Privacy concerns⟩
(c) General Expression	
Conclusion	⟨Good⟩ ⟨Good, but⟩ ⟨Bad, but⟩ ⟨Bad⟩
Usability	⟨Easy to use / Understandable⟩ ⟨Cumbersome⟩ ⟨Fast⟩ ⟨Complicated Safari UI⟩
Security	⟨Secure⟩ ⟨Too easy, can feel insecure⟩ ⟨Privacy concerns⟩
Cognitive effort	⟨No password memorization issues⟩ ⟨Easier than a password manager⟩ ⟨(Needs) no passwords⟩
Availability	⟨Forgetting/Loss/Theft/Destruction⟩
Mental models	⟨Technology mistrust⟩ ⟨Important⟩ ⟨Innovative⟩ ⟨More transparency/ information required⟩
(d) Adoption reasons	
Importance	⟨Sensitive accounts⟩ ⟨Non-sensitive accounts⟩
Business Stuff	⟨Business-related accounts⟩
Anonymity	⟨Accounts supposed to be anonymous⟩
Multiple Devices	⟨Accounts not used on other devices⟩ ⟨Accounts mainly used on smartphone⟩ ⟨Accounts used on other devices⟩
Account Sharing	⟨Shared accounts⟩ ⟨Accounts that are not shared⟩
Frequency	⟨Rarely used accounts⟩ ⟨Frequently used accounts⟩
Speed	⟨Accounts where fast access is not important⟩ ⟨Accounts where fast access is crucial⟩
Mobility	⟨Accounts used outside of home⟩

(1) Social network account (2) Streaming service account (3) Messenger service account (4) Travel portal account (5) Gaming portal account (6) Internet service provider account (7) Google/Apple account (8) Private email account (9) Business email account (10) Online shop account (11) Online banking account (12) Please briefly explain for which accounts you would decide to use or not use the presented authentication method. ⟨Free text response⟩

2FA FAMILIARITY. Which two-factor authentication methods have you previously used?

(1) Text messages ⟨Checkbox⟩ (2) TAN lists ⟨Checkbox⟩ (3) Code generators ⟨Checkbox⟩ (4) Smartphone apps ⟨Checkbox⟩ (5) Hardware keys ⟨Checkbox⟩ (6) Other: ⟨Free text response⟩

DEMOGRAPHICS. (1) Please state your gender. ⟨Male / Female / No answer / Other⟩ (2) How old are you? ⟨10-19 / 20-29 / 30-39 / 40-49 / 50-59 / 60-69 / 70-79 / ≥80⟩ (3) Please state your highest educational degree: ⟨Still in school / Middle school / High school / Bachelor's degree / Master's degree or Diploma / Doctorate / Other⟩ (4) Please state your area of work or area of studies. ⟨Free text response⟩ (5) Is there anything else you would like to tell us? ⟨Free text response⟩

C CODEBOOKS

This section lists the codebooks (Table 8) that we created during the qualitative analysis of the four open-ended text-questions.

Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions

Citation	F. Putz, S. Schön, and M. Hollick. “Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions.” In: <i>Emerging Technologies for Authorization and Authentication</i> . ETAA '21. Cham, Switzerland: Springer International Publishing, 2021, pp. 17–32. ISBN: 978-3-030-93746-1. DOI: 10.1007/978-3-030-93747-8_2
Contribution	This publication is discussed in Chapter 2. My contribution is described in the section “ <i>Collaborations and My Contribution</i> ” on page x.
Rights notice	© Springer Nature Switzerland AG 2021. Reproduced with permission from Springer Nature. The final authenticated version is available online at https://doi.org/10.1007/978-3-030-93747-8_2 .

Future-Proof Web Authentication: Bring Your Own FIDO2 Extensions

Florentin Putz, Steffen Schön, and Matthias Hollick

TU Darmstadt, Germany
{fputz, sschoen, mhollick}@seemoo.de

Abstract. The FIDO2 standards for strong authentication on the Internet define an extension interface, which allows them to flexibly adapt to future use cases. The domain of establishing new FIDO2 extensions, however, is currently limited to web browser developers and members of the FIDO alliance. We show how researchers and developers can design and implement their own extensions for using FIDO2 as a well-established and secure foundation to demonstrate innovative authentication concepts or to support custom deployments. Our open-source implementation targets the full FIDO2 stack, such as the Chromium web browser and hardware tokens, to enable tailor-made authentication based on the power of the existing FIDO2 ecosystem. To give an overview of existing extensions, we survey all published FIDO2 extensions by manually inspecting the source code of major web browsers and authenticators. Their current design, however, hinders the implementation of custom extensions, and they only support a limited number of extensions out of the box. We discuss weaknesses of current implementations and identify the lack of extension pass-through as a major limitation in current FIDO2 clients.

Keywords: Security · Authentication · Key Management · Hardware Token · Passwordless · WebAuthn

1 Introduction

Stronger forms of authentication than passwords can protect online accounts from phishing attacks and mitigate the impact of data breaches. The FIDO2 standards [8] are now implemented in all major web browsers and allow users to securely log in to websites without passwords, using either a hardware security key (e.g., a YubiKey [28]) or a built-in authenticator in modern smartphones or laptops. Website operators can easily implement support for FIDO2 using the WebAuthn JavaScript API [25], which also supports custom extensions to implement special use cases.

Extensible web standards such as X.509 [2] or TLS [3] have played a key role during the advancement of the Internet so far. A standards organization such as the IETF cannot predict all possible future use cases of the standard beforehand, but custom extensions provide flexibility to the otherwise fixed specification and prevent fragmentation via the creation of new competing standards to satisfy



new use cases. The behavior of a protocol can then be modified without the effort to get the modifications accepted into the main standard.

Why extend FIDO2? At its core, FIDO2 is a well-established and secure platform for accessing public-key credentials, with little restrictions imposed on the authenticator. In the future, however, innovative authentication approaches might require modifications to the standard, e.g., by transmitting or receiving additional information to the authenticator. FIDO2 extensions can be used to support new use cases, implement additional features, and mitigate shortcomings of the standard.

The potential of FIDO2 extensions becomes clear by looking at the extensions that have been proposed so far: A major weakness of FIDO2 is the lack of efficient recovery options, which has been criticized in previous user studies [15,5,1]. Yubico proposed a new FIDO2 extension to address this problem, by implementing an efficient way to automatically register backup credentials, which can be used in case the authentication token gets lost [30].

Developers and researchers can also use FIDO2 extensions to prototype and demonstrate new authentication designs within the existing FIDO2 ecosystem on real-world web browsers. Companies can use FIDO2 extensions to adapt the protocol for tailor-made authentication in internal deployments.

Web browsers, however, currently do not support any third-party FIDO2 extensions. Furthermore, there is no clear development path to implement custom extensions in current web browsers. Even FIDO2 extensions which do not require any special processing by the web browser need explicit browser support instead of automatically being forwarded to the authenticator. Although extensions are an important part of any standard, the extensibility of the FIDO2 standards unfortunately has not received much attention as of today. This calls for a detailed analysis of FIDO2 extensions:

- As our core contribution, we show how to design and implement tailor-made authentication based on the power of the existing FIDO2 ecosystem (Section 5). Our source code and additional documentation is available at <https://seemoo.de/s/fido2ext>.
- We survey all publicly known FIDO2 extensions by manually inspecting the source code of major web browsers and authenticators. (Section 3 and Section 4).
- We describe limitations of current FIDO2 extension implementations and identify the lack of extension pass-through as a major weakness that inhibits the development of innovative FIDO2 extensions (Section 6).

2 Background

2.1 FIDO2

Strong and passwordless authentication has been standardized by the FIDO industry alliance and the World Wide Web Consortium (W3C) in form of the

FIDO2 standards [8]. The top part of Figure 1 shows the FIDO2 system model, consisting of relying party (RP), client, and authenticator. In essence, FIDO2 is a challenge-response protocol, which RPs such as websites can use to access public key credentials, managed for the user by an authenticator. The authenticator can be embedded into the client device or attached as an external hardware token via USB, NFC, or Bluetooth. It mainly supports the following two operations:

1. **authenticatorMakeCredential**: The authenticator generates a new key pair, binds it to the requesting RP, and returns the public key and optionally an attestation signature. Websites use this operation to implement account registration.
2. **authenticatorGetAssertion**: The authenticator returns an assertion signature over a challenge from the RP. Websites use this operation to implement account login.

The client (e.g., a web browser) forwards the RP’s requests to the authenticator, which stores all key material. FIDO2 consists of two specifications, WebAuthn (Web Authentication) [25] and CTAP (Client to Authenticator Protocol) [6], which succeed the older Universal 2nd Factor (U2F) standard [7]. All major web browsers implement FIDO2 [17] and there exists a large ecosystem of commercial hardware tokens, e.g., Yubico’s YubiKey [28].

2.2 FIDO2 Extensions

The FIDO2 authentication protocols WebAuthn and CTAP2 can be extended to support special use cases. Figure 1 shows the general protocol flow of a FIDO2 extension, as specified by §9 of the WebAuthn specification [25]. The RP initiates an extension request with the client, which then processes the request and responds to the RP. WebAuthn distinguishes between different types of extensions. An extension is always a *client extension*, as it involves communication with the client. The extension specification defines the form of the JSON *client extension input/output* and what kind of *client extension processing* the client needs to perform to transform the inputs to the outputs.

If the extension also involves processing on the authenticator, it is additionally an *authenticator extension* and defines the form of the concise binary object representation (CBOR) *authenticator extension input/output* and what kind of *authenticator extension processing* the authenticator needs to perform to transform the inputs to the outputs. In this case, the client also needs to know how to convert the JSON *client extension input* to the CBOR *authenticator extension input*, and how to convert the CBOR *authenticator extension output* back to the JSON *client extension output*.

FIDO2 extensions also are *registration extensions* and/or *authentication extensions*, depending on which part of the protocol they affect. We describe further implementation details in Section 5.

Each FIDO2 extension has an extension identifier string, which can be registered in the IANA “WebAuthn Extension Identifiers” registry [12]. This registry

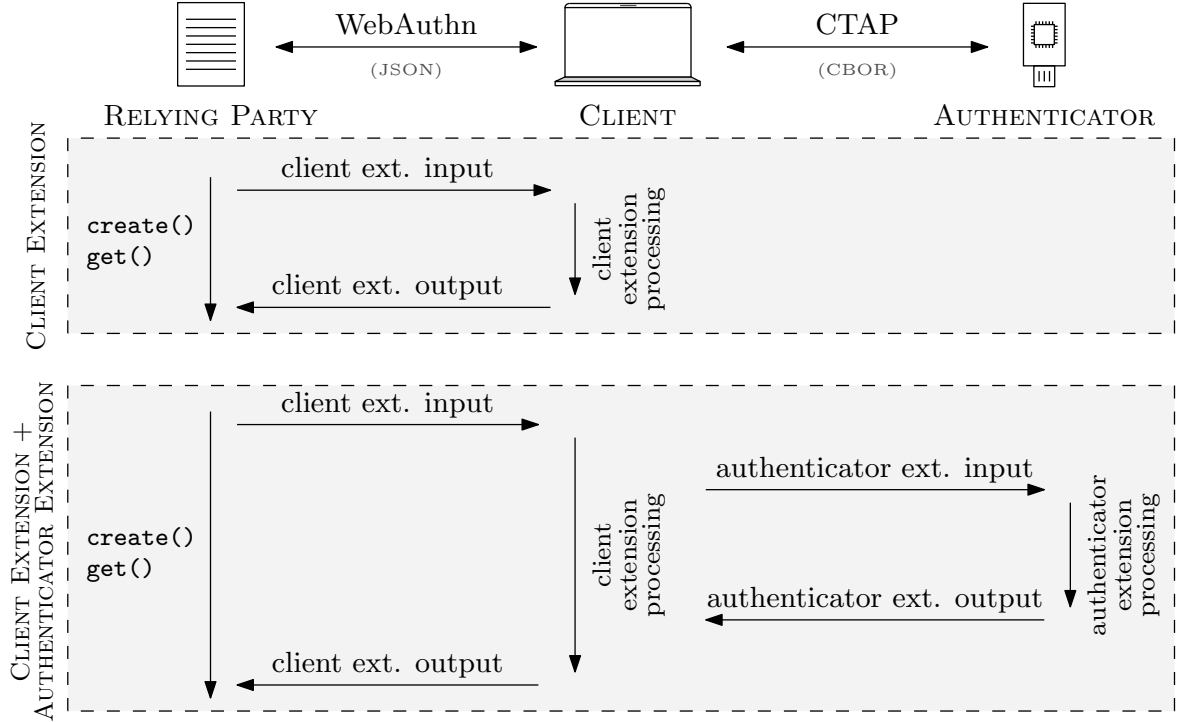


Fig. 1. Protocol flow of FIDO2 extensions. Extensions can communicate between RP, client, and authenticator (*top*). An extension is always a client extension (*middle*) but can also additionally be an authenticator extension (*bottom*).

has been defined and established by RFC 8809 [11] and contains an up-to-date list of all currently registered WebAuthn extension. We describe all currently available FIDO2 extensions in Section 3.

Authenticators supporting CTAP 2.1 must implement the `hmac-secret` extension and the `credProtect` extension if they support some form of user verification (§9 in CTAP2 [6]). Clients can use the CTAP `authenticatorGetInfo` method to detect which extensions an authenticator supports. We give an overview of which FIDO2 extensions are supported by popular browsers and authenticators in Section 4.

2.3 Extension Pass-Through

In general, clients and authenticators do not need to support any WebAuthn extensions and can simply ignore them as they are all optional. RPs must therefore be prepared to deal with any subset of requested extensions being ignored, as this must never fail the WebAuthn transaction. To increase compatibility with unknown extensions, clients can choose to directly pass through extension inputs to the authenticator and outputs to the RP, without any additional client processing. The WebAuthn standard defines conversion rules between JSON and CBOR to facilitate this [26]. Authenticators must be prepared to ignore such inputs in case such a direct pass-through results in invalid inputs.

To increase compatibility, extensions that do not require special client processing should define the authenticator input and outputs in such a way that direct pass-through results in semantically correct values. Clients that implement pass-through facilitate innovation as they allow extensions to work when only the authenticator explicitly supports them.

3 Survey: Existing FIDO2 Extensions

Our survey on FIDO2 extensions consists of two parts: (1) This section is the first part, where we identify all extensions that have been published so far. (2) Section 4 is the second part, where we determine which clients and authenticators support these extensions. Table 1 shows the results of our survey.

Most extensions are specified in the FIDO2 standards themselves (“standard extensions”), such as the `hmac-secret` extension for using a symmetric key with a FIDO2 credential [6]. Additionally, there are three non-standard extensions which are not specified in the official FIDO2 standards. First, there is Yubico’s `recovery` extension [30], which has a separate specification but no web browser implementation yet. Second, there is Google’s `caBLE` extension [19], which is part of the Google Chrome browser but has no public specification. Third, there is the `googleLegacyAppIdSupport` extension [13], which is implemented in some browsers but has no formal specification yet. We briefly introduce the standard extensions and then go over the non-standard extensions in more detail, as they are currently less well-known.

3.1 Standard Extensions

We refer to the extensions defined in the FIDO2 standards as *standard extensions*. At the time of writing, the IANA registry [12] contains eight registered extension identifiers, which are all specified in WebAuthn Level 1 [24]. WebAuthn Level 2 defines three additional extensions and also plans to register them (§12.4 in [25]). Furthermore, CTAP 2.1 defines five extensions and plans to register them as well (§14.1 in [6]). One of them is the `hmac-secret` extension, which allows requesting a symmetric secret from an authenticator to encrypt and decrypt user data. This extension can support new use cases where a static secret is necessary, such as disk encryption or unlocking password managers like KeePass. The other standard extensions enable features such as compatibility with legacy Universal 2nd Factor (U2F) credentials, specifying a credential protection policy, or storing arbitrary data associated with a credential. The FIDO2 standards contain further information on each extension as well as a formal specification.

3.2 Recovery Extension (Yubico)

Recovery and backups still are open problems within the FIDO2 standards, as the current best practice of manually registering multiple tokens on each RP has low usability. Yubico aims to solve this problem using the `recovery` extension,

which currently has a draft specification available online [30]. This extension allows automatically registering a backup authenticator for recovery purposes without needing to have the backup authenticator physically available at the time of registration, so that it can be permanently stored at a secure location. Their proposal is based on a key agreement scheme where the primary authenticator generates nondeterministic public keys, but only the backup authenticator can derive the corresponding private keys. This enables a usable recovery process while maintaining the FIDO2 privacy protections of unlinkable public keys. Frymann et al. generalized the procedure to Asynchronous Remote Key Generation (ARKG) and proved the cryptographic security of such protocols [9].

3.3 CaBLE Extension (Google)

Google proposed caBLE (Cloud-Assisted Bluetooth Low Energy), which allows using a smartphone as a FIDO2 roaming authenticator via BLE [19,20]. The caBLE protocol consists of two phases: A pairing phase using a QR code, which encodes a nonce to derive key material. Afterwards, the smartphone broadcasts advertisements via BLE, which the web browser recognizes to initiate a handshake phase to establish a CTAP2 channel. There is no public specification available for caBLE and it is currently implemented only in the Chrome web browser and in Android smartphones via the proprietary Google Play Services. The Chromium implementation hints at a WebAuthn registration extension with client input `cableRegistration` and an authentication extension with client input `cableAuthentication`. The communication with the smartphone is not implemented using the standard FIDO2 extension API, but using a custom CTAP2 transport instead.

3.4 GoogleLegacyAppIdSupport Extension (Google)

Google proposed the `googleLegacyAppIdSupport` extension [13] to provide compatibility with some Android factory images, which only support the U2F JavaScript API and cannot be patched with WebAuthn support. Google Chrome plans to deprecate the U2F API, but credentials created using WebAuthn are normally not backwards compatible with the U2F JavaScript API. This extension allows creating WebAuthn credentials which work with both WebAuthn and U2F, by using a hard-coded U2F AppID specific to Google Accounts, limited to `*.google.com` domains.

4 Survey: Compatibility of FIDO2 Extensions

While the first part of our survey identified all currently known FIDO2 extensions, this section contains the second part of our survey, where we determine the client and authenticator support for all FIDO2 extensions. Since it is standard compliant for a client to ignore any extensions, the currently available feature trackers for web browser WebAuthn implementations do not indicate support

Table 1. FIDO2 extension compatibility as of July 2021.

Identifier	Chrome Edge	Firefox Safari	libfido2 python-fido2	Windows Hello YubiKey 5	SoloKey OpenSK	Specification	Reference	IANA Registration?
appid	✓	✓	✓	✗	✓	WebAuthn2 §10.1 [25]	✓	✓
appidExclude	✓	✓	✗	✗	✗	WebAuthn2 §10.2 [25]	✗	✗
authnSel	✗	✗	✗	✗	✗	WebAuthn1 §10.4 [24]	✓	✓
biometricPerfBounds	✗	✗	✗	✗	✗	WebAuthn1 §10.9 [24]	✗	✗
credBlob	✓	✓	✗	✓	✓	CTAP2 §12.2 [6]	✗	✗
credProps	✓	✓	✗	✗	✗	WebAuthn2 §10.4 [25]	✗	✗
credProtect	✓	✓	✗	✓	✓	CTAP2 §12.1 [6]	✗	✗
exts	✗	✗	✗	✗	✗	WebAuthn1 §10.5 [24]	✓	✓
hmac-secret	✓	✓	✓	✓	✓	CTAP2 §12.5 [6]	✗	✗
largeBlob	✓	✓	✗	✗	✗	WebAuthn1 §10.5 [24]	✗	✗
largeBlobKey	✓	✓	✗	✓	✓	CTAP2 §12.3 [6]	✗	✗
loc	✗	✗	✗	✗	✗	WebAuthn1 §10.7 [24]	✓	✓
minPinLength	✗	✗	✗	✗	✓	CTAP2 §12.4 [6]	✗	✗
txAuthSimple	✗	✗	✗	✗	✗	WebAuthn1 §10.2 [24]	✓	✓
txAuthGeneric	✗	✗	✗	✗	✗	WebAuthn1 §10.3 [24]	✓	✓
uvi	✗	✗	✗	✗	✗	WebAuthn1 §10.6 [24]	✓	✓
uvm	✓	✓	✗	✗	✗	WebAuthn2 §10.3 [25]	✓	✓
caBLE	✓	✓	✗	✗	✗		[19]	✗
googleLegacyAppIdSupport	✓	✓	✗	✗	✗		[13]	✗
recovery	✗	✗	✗	✗	✗		[30]	✗

for individual extensions. Table 1 gives an overview of FIDO2 extension support for popular web browsers, client libraries, and authenticators. In the following, we explain our methodology and results in more detail.

4.1 Web Browsers

There is no automatic way to query all extensions supported by a web browser, so we need to manually inspect each browser’s source code to see which FIDO2 extensions they implement. We checked the Chromium source code¹ (which forms the base of Google’s Chrome browser and Microsoft’s Edge browser), the Gecko source code² (which corresponds to Mozilla’s Firefox browser), and the WebKit source code³ (which forms the base of Apple’s Safari browser). Table 1 shows

¹ Chromium version *canary 93.0.4570.0* → Google Chrome 93, Microsoft Edge 93.

² Gecko version *nightly 91.0a1* → Mozilla Firefox 91.

³ WebKit version *611.2.7.1* → Apple Safari 14.1.1.

all WebAuthn and CTAP2 extensions currently supported by these browsers. Chrome supports 11 FIDO2 extensions, including the non-standard caBLE extension, while Firefox and Safari only support two extensions each. All browsers support the `appid` extensions for compatibility with legacy U2F credentials.

Extension pass-through is an important feature that allows FIDO2 clients to transparently forward extension inputs and outputs between RP and authenticator, even for unknown extensions. The WebAuthn specification notes that extension pass-through can “facilitate innovation, allowing authenticators to experiment with new extensions and RPs to use them before there is explicit support for them in clients.” [26]. While inspecting the web browser source codes, however, we noticed that no web browser currently supports extension pass-through. This makes it very difficult to establish custom extensions, as users will not be able to use them.

4.2 Client Libraries

Yubico’s `python-fido2` library [27] can be used to implement custom FIDO2 clients that interface with CTAP authenticators. It currently supports six FIDO2 extensions as of version 0.9.1. This library is especially interesting for researchers and developers as it provides an easy interface for implementing custom CTAP2 extensions. Yubico also provides the C library `libfido2` [29] for interacting with FIDO2 authenticators, with bindings available to many other languages. This library supports four different FIDO2 extensions as of version 1.7.0.

4.3 Authenticators

To query the extensions supported by different FIDO2 authenticators, we use the `authenticatorGetInfo` CTAP2 request to retrieve a list of supported extensions in the `extensions` response field. As not all extensions are specified to announce themselves this way, we also manually inspect the authenticator’s source code if available. We looked at Yubico’s YubiKey 5 NFC, YubiKey 5C NFC, the open source Solo⁴ C implementation (which powers the SoloKey, the NitroKey, and the OnlyKey), and Google’s open source OpenSK⁵ authenticator implementation written in Rust. All tokens only support the `credProtect` and the `hmac-secret` extensions.

In addition to these roaming authenticators, it is also possible to use platform authenticators. Microsoft provides Win32 headers for communicating with the Windows Hello platform authenticator and with roaming authenticators in general [18]. The actual implementation of these APIs is proprietary, but developers can use these headers to support FIDO2 authentication on Windows. As of `WEBAUTHN_API_VERSION_3`, Windows implements three FIDO2 extensions. It also partially implements `hmac-secret`, but obtaining the secret is not supported yet.

⁴ Solo version 4.1.2.

⁵ OpenSK version 1.0.0.

4.4 Summary

The survey in the last two sections shows that FIDO2 extensions have the potential to solve major weaknesses of the FIDO2 standards, such as the lack of efficient recovery options. Support of these extensions by web browsers, clients, and authenticators is currently weak, however, as shown in Table 1. Some extensions defined in the FIDO2 standards have no support at all. Extension pass-through would improve the compatibility of FIDO2 clients with current and future FIDO2 extensions, which would facilitate researchers and developers to create their own extensions, to add new features or to solve shortcomings of FIDO2.

5 Design and Implementation of Custom Extensions

In order to showcase the possibilities of FIDO2 extensions and the required steps to implement one, we define a proof-of-concept extension that we implement on all parts of the FIDO2 stack: RPs (web applications), clients (browser and non-browser), and authenticators. Implementing a custom FIDO2 extension is not straightforward and often undocumented, so we try to facilitate the process for other researchers and developers. We start by defining our extension and then describe the process of implementing this extension for each affected component. Our source code as well as supplemental material is available online⁶.

There are two primary requirements for our proof-of-concept extension. First, it should use all capabilities that FIDO2 extensions offer, namely inputs, outputs, and usage in both registration (`makeCredential`) and authentication (`getAssertion`). Second, it should be simple in order to allow readers to focus on the implementation of an extension in general instead of the functionality of our example. Thus we choose to define and implement the **greeter** extension, which implements a concept that is well known in communications.

Definition 1 (Greeter Extension). *The **greeter** extension is an authenticator extension, which allows the authenticator to respond to the RP with a greeting message. We define an input string, which is the name of the sender (e.g., “John”). We also define an output string, which is the greeting message (e.g., “Hello John”). This extension requires no special client extension processing, as the client can simply pass through all inputs/outputs to/from the authenticator. The authenticator extension processing consists of constructing the greeting message from the input string and returning it. Our greeter extension uses the identifier **greeter** in both CTAP and WebAuthn.*

5.1 Relying Party

The RP initiates the FIDO2 authentication protocol and can include extensions in its requests. To demonstrate this, we implement our FIDO2 extension in a proof-of-concept web application, which uses the native browser

⁶ Online repository with our source code and additional documentation on how to implement your own FIDO2 extensions: <https://seemoo.de/s/fido2ext>

JavaScript WebAuthn APIs supported by all modern browsers [16]. The process of using extensions for registration (`makeCredential`) and authentication (`getAssertion`) is very similar. To register a credential using the browser API, we use `navigator.credentials.create()`. This function takes a parameter object with the `publicKey` member in case of WebAuthn.

The `publicKey` object accepts an `extensions` attribute that contains the extension identifier including inputs: `extensions: {"greeter": "John"}`. In this case we do not use a backend to generate the challenge and various other values. In practice, those values should not be computed on the client side. After the extension was successfully processed by browser and authenticator, the outputs of the extension can be accessed using `credential.getClientExtensionResults()`, which returns an object with the extension identifiers as keys and the respective extension outputs as values.

5.2 Web Browser

Web browsers are the most common FIDO2 client and are next in the FIDO2 stack. They receive the authentication request from the RP and communicate with the authenticator. As web browsers do not currently support FIDO2 extension pass-through, we need to modify the browser to support our custom extension. We have implemented the `greeter` extension in the Chromium browser, as it is the foundation of the most popular web browser on the market, Google Chrome. Unfortunately, Chromium does not implement FIDO2 extension support in a modular way, which means that we need to modify different components of the browser stack in order to implement support for our `greeter` extension. Chromium does not provide any documentation for this, so we had to find out the correct approach by inspecting the implementation of other FIDO2 extensions in the source code.

Figure 2 shows how Chromium processes and sends FIDO2 extensions to authenticators. Let us assume a web application with the JavaScript code described in the previous section. Chromium’s rendering engine Blink contains the V8 JavaScript engine, which runs this code and dispatches the command `navigator.credentials.create()` to the corresponding Blink module implementing this Web Platform API. This command is part of the general Credential Management API⁷, which gets called first. This gets treated as a WebAuthn request in `CredentialsContainer::create()`, because we specified `publicKey` in the code above.

The JavaScript request contains a parameter object following the `PublicKeyCredentialCreationOptions` WebIDL specification. This object also contains the client extension inputs according to `AuthenticatorExtensionsClientInputs`, which we need to modify to include our own extension input. In Chromium, these parameters from JavaScript get converted into Mojo structures in `authenticator.mojom`⁸,

⁷ Chromium Credential Management API (*Blink*): `third_party/blink/renderer/modules/credentialmanager/`

⁸ Chromium WebAuthn Mojo (*Blink*): `third_party/blink/public/mojom/webauthn/`

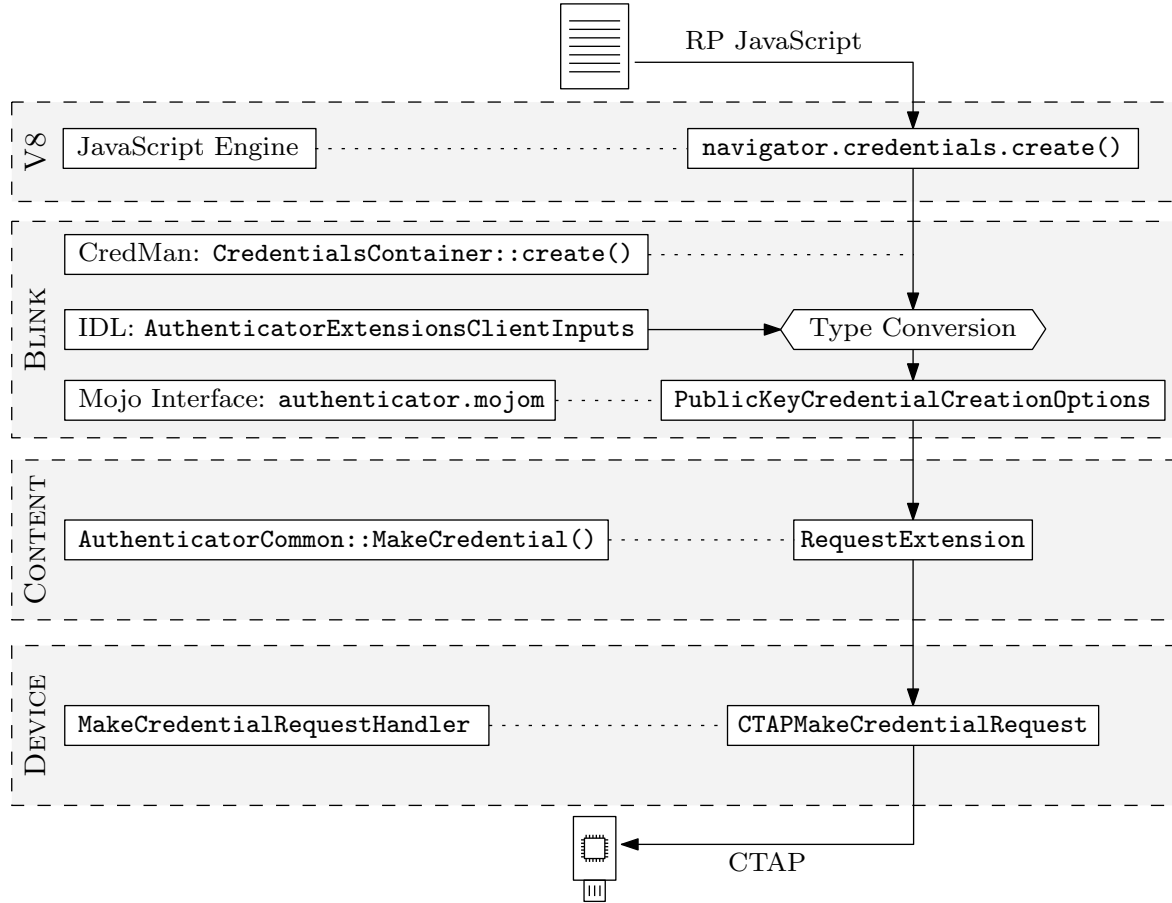


Fig. 2. Chromium’s FIDO2 client implementation, showing the data and processing flow for sending extensions to an authenticator.

which are used for IPC between different parts of Chromium. In our case, the relevant Mojo structure is `PublicKeyCredentialCreationOptions`. We need to change both the Mojo structure and the converter from IDL to Mojo. Afterwards, the call arrives at `AuthenticatorCommon::MakeCredential()` inside Chromium’s `content` layer⁹. We need to modify this function to include our extension as a `RequestExtension`. This function then calls the `MakeCredentialRequestHandler` inside Chromium’s `device` layer¹⁰. We modify the `CTAPMakeCredentialRequest`, which converts our request into a CTAP message. Afterwards, this gets sent to the authenticator, which responds back to us.

The response path is similar and traverses back to the `content` layer and finally to the `Blink` layer, where the client extension output gets converted into the WebIDL `AuthenticationExtensionsClientOutputs` and passed back to the V8 JavaScript engine.

⁹ Chromium Web Authentication (*content*): `content/browser/webauth/`

¹⁰ Chromium CTAP (*device*): `device/fido/`

5.3 Client Library

Besides browsers, it is also possible to implement a FIDO2 client in desktop applications. The `python-fido2` library [27] by Yubico provides the client functionality to communicate with an authenticator using CTAP2 in Python. Although not officially documented, the library can be extended with custom FIDO2 extensions due to its modular design. We have implemented our custom extension by inheriting from the `Ctap2Extension` class, which is located in `fido2/ctap2/extensions.py`. There are two abstract methods, `process_create_input` and `process_create_output`, which need to be implemented in order to use extension inputs and outputs. The subclass does not have to be registered anywhere, as `Fido2Client` considers all subclasses of `Ctap2Extension` when parsing a request or response.

5.4 Authenticator

The final component involved in FIDO2 is the authenticator. When implementing a FIDO2 extension, there are a number of open source hardware security token implementations available that we can build on:

- SoloKeys Solo 1 [21]
- SoloKeys Solo 2 [23]
- Google OpenSK [10]

While Solo 2 is still in early development and thus might not be suitable yet, Solo 1 and OpenSK are in a state that can be used to implement custom extensions. We have implemented the **greeter** extension in Solo 1, but describe the process for both Solo 1 and OpenSK. Neither project has a modular software architecture for implementing FIDO2 extensions, as all extensions are hard-coded at different code locations across the CTAP stack. Solo 1, however, provides documentation on how to implement custom extensions [22]. For OpenSK, we figured out the correct approach by inspecting the implementation of other FIDO2 extensions in the source code. In general, when implementing an extension for an authenticator, three modifications are required independent of the CTAP implementation:

- Announcing the support for the extension in the CTAP 2 `getInfo` function.
 - In Solo: `ctap_get_info()`.
 - In OpenSK: `process_get_info()`.
- Parsing and processing the extension inputs.
 - In Solo: `struct CTAP_extensions, ctap_parse_extensions()`.
 - In OpenSK: `struct MakeCredentialExtensions, MakeCredentialExtensions::try_from(), create_make_credential_parameters_with_cred_protect_policy()`.
- Building the extension outputs.
 - In Solo: `ctap_make_extensions()`.
 - In OpenSK: `process_make_credential()`.

6 Discussion

6.1 Extension Pass-Through

In the previous sections, we showed that FIDO2 extensions can address shortcomings of the current FIDO2 standards and allow them to easily adapt to new use cases. In practice, the applicability of FIDO2 extensions is limited due to browser design decisions. In order to facilitate the development and usage of innovative FIDO2 extensions, we argue that web browsers should pass-through unknown extensions to authenticators, as the WebAuthn specification suggests [26].

Some custom FIDO2 extensions might pose the risk of exposing too much fine-grained data to RPs, which would allow them to discriminate between authenticators. This is the reason why Chromium does not support the official `uvi` and `uvm` extensions [14]. This potential problem can be approached by letting the user opt-in to pass-through custom FIDO2 extensions, e.g., via a configuration option in `chrome://settings/securityKeys` or via the permission API that is already used for camera or microphone access.

6.2 Supporting FIDO2 Extensions via Browser Extensions

Apart from supporting pass-through, web browsers can also facilitate the development of innovative FIDO2 extensions by making it easier to implement them. As we have shown before, Chromium’s support for FIDO2 extensions is currently hard-coded in many different parts of the browser stack. This could be improved using a modular architecture. Another interesting approach would be to allow implementing custom FIDO2 extensions as regular web browser extensions. Chromium already has an extensive web extensions API, which could be extended with an API for FIDO2 extensions. This would also facilitate the use case of internal company deployments, as it is already possible in an enterprise setting to provision web browsers with a set of extensions.

6.3 Outlook

Without support for custom FIDO2 extensions in web browsers, it is unlikely that FIDO2 extensions will be utilized to a similar extent than X.509 or TLS extensions (see Section 7). The domain of developing new FIDO2 extensions currently is rather limited to developers of web browsers or members of the FIDO Alliance, like Google is already doing with their `caBLE` extension. Third-party developers have difficulties of implementing and testing their custom FIDO2 extensions in web browsers. In practice, the primary use case of custom extensions is probably not the public Internet but rather internal deployments with a limited target audience. In this case, the argument of possible authenticator discrimination is not applicable anymore as these internal deployments already have custom authentication policies in place.

The previously mentioned restriction on the usage of custom extension might also explain the limited support for implementing custom extensions in FIDO2

authenticators and clients. Documentation is often limited, if available at all. The software architectures are not designed to be modular with regards to FIDO2 extensions. One instance of a well designed modular extension architecture is the `python-fido2` client library, although it still requires modification of the internal source code and lacks public documentation.

7 Related Work

FIDO2 is not the first protocol that allows the definition of extensions in order to extend the capabilities of the specified protocol. There are a number of other protocols or data formats that support extensions and successfully integrate them in real world usage, two of which we describe in more detail:

- **X.509 Certificates:** In the context of public key infrastructures, the use of X.509 [2] certificates is very common. X.509 allows the inclusion of additional fields in certificates, revocation lists, etc. through extensions. In practice, these extensions are heavily used to address issues that cannot be solved using the specification alone. An example is the subject alternative name extension, which allows the declaration of additional identities of the certificate subject, as the basic certificate only allows a common name.
- **TLS:** The transport layer security protocol provides a secure communication layer for otherwise insecure protocols such as HTTP. Since TLS 1.2 [3], additional data can be communicated in the handshake of the protocol using extensions [4]. These can, for example, be used to let clients specify the domain of the requested resources in addition to the IP address using the server name indication extension. In the modern web, servers often do not only serve a single domain but a number of domains. If different certificates are used for the different domains, the server cannot know which certificate to include in the response without the server name indication extension.

In summary, extensible protocols are easier to adapt to new use cases, which prevents fragmentation. This is especially important for authentication protocols such as FIDO2, which so far have not managed to get widespread adoption to replace passwords.

8 Conclusion

In this paper, we showed that the extensibility of the FIDO2 standards is an important feature to remain flexible and future-proof. We demonstrated how to design and implement new FIDO2 extensions in web browsers, client libraries, and authenticators. This allows researchers and developers to use the existing FIDO2 infrastructure as a secure foundation for custom authentication deployments or for demonstrating new authentication approaches with real-world browsers and

hardware. We surveyed all existing FIDO2 extensions and determined their respective support by web browsers and authenticators. A small number of extensions have widespread support, but most extensions have only minimal support or no support at all.

Current FIDO2 client implementations limit FIDO2’s extensibility due to missing features and complex software architectures. We also showed that web browsers should (1) implement a configuration option that enables FIDO2 extension pass-through, and (2) consider extending the regular web browser extensions API to allow implementing custom FIDO2 extensions as, e.g., Chrome extensions. Both steps would empower researchers and developers to create their own extensions.

Acknowledgements. This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center. We thank the anonymous reviewers for reviewing this paper and for their helpful comments.

References

1. Ciolino, S., Parkin, S., Dunphy, P.: Of two minds about two-factor: Understanding everyday FIDO U2F usability through device comparison and experience sampling. In: Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019). pp. 339–356. USENIX Association, Santa Clara, CA (Aug 2019)
2. Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., Polk, W.: Internet X.509 public key infrastructure certificate and certificate revocation list (CRL) profile. RFC 5280, RFC Editor (May 2008), <http://www.rfc-editor.org/rfc/rfc5280.txt>
3. Dierks, T., Rescorla, E.: The transport layer security (TLS) protocol version 1.2. RFC 5246, RFC Editor (Aug 2008), <http://www.rfc-editor.org/rfc/rfc5246.txt>
4. Eastlake, D.: Transport layer security (TLS) extensions: Extension definitions. RFC 6066, RFC Editor (Jan 2011), <http://www.rfc-editor.org/rfc/rfc6066.txt>
5. Farke, F.M., Lorenz, L., Schnitzler, T., Markert, P., Dürmuth, M.: “you still use the password after all” – exploring FIDO2 security keys in a small company. In: Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020). pp. 19–35. USENIX Association (Aug 2020)
6. FIDO Alliance: Client to authenticator protocol (CTAP) (Jun 2021), <https://fidoalliance.org/specs/fido-v2.1-ps-20210615/fido-client-to-authenticator-protocol-v2.1-ps-20210615.html>
7. FIDO Alliance: FIDO U2F raw message formats (Apr 2017), <https://fidoalliance.org/specs/fido-u2f-v1.2-ps-20170411/fido-u2f-raw-message-formats-v1.2-ps-20170411.html>
8. FIDO Alliance: FIDO2: Moving the world beyond passwords using WebAuthn & CTAP (Jun 2020), <https://fidoalliance.org/fido2>
9. Frymann, N., Gardham, D., Kiefer, F., Lundberg, E., Manulis, M., Nilsson, D.: Asynchronous remote key generation: An analysis of yubico’s proposal for W3C WebAuthn. Association for Computing Machinery, New York, NY, USA (Oct 2020). <https://doi.org/10.1145/3372297.3417292>
10. Google: OpenSK. <https://github.com/google/OpenSK> (2020)

11. Hodges, J., Mandyam, G., Jones, M.B.: RFC 8809: Registries for web authentication (WebAuthn). RFC 8809, RFC Editor (Aug 2020), <https://www.rfc-editor.org/rfc/rfc8809.txt>
12. IANA: Web authentication (WebAuthn) registries (Aug 2020), <https://www.iana.org/assignments/webauthn/webauthn.xhtml>
13. Kreichgauer, M.: Intent to deprecate and remove: U2F API (cryptotoken). https://groups.google.com/a/chromium.org/g/blink-dev/c/xHC3AtU_65A (2021)
14. Langley, A.: Re: Issue 1097972: Support WebAuthn uvi & uvm extension (Jun 2020), <https://bugs.chromium.org/p/chromium/issues/detail?id=1097972#c3>
15. Lyastani, S.G., Schilling, M., Neumayr, M., Backes, M., Bugiel, S.: Is FIDO2 the kingslayer of user authentication? a comparative usability study of FIDO2 passwordless authentication. In: 2020 IEEE Symposium on Security and Privacy (SP). pp. 842–859. IEEE Computer Society (May 2020). <https://doi.org/10.1109/SP40000.2020.00047>
16. MDN: Web authentication API. https://developer.mozilla.org/en-US/docs/Web/API/Web_Authentication_API (2021)
17. MDN: Web authentication API: Browser compatibility (Mar 2021), https://developer.mozilla.org/en-US/docs/Web/API/Web_Authentication_API#Browser_compatibility
18. Microsoft: Win32 headers for WebAuthn. <https://github.com/microsoft/webauthn> (2021)
19. Mooney, N.: Addition of a network transport. <https://github.com/w3c/webauthn/issues/1381> (2020)
20. Protalinski, E.: You can now use your android phone as a 2FA security key for google accounts. VentureBeat (Apr 2019), <https://venturebeat.com/2019/04/10/you-can-now-use-your-android-phone-as-a-2fa-security-key-for-google-accounts>
21. SoloKeys: Solo 1: open security key supporting FIDO2 & U2F over USB + NFC. <https://github.com/solokeys/solo> (2018)
22. SoloKeys: Tutorial: Writing an extension for the solo stick. <https://github.com/solokeys/solo/blob/b86f0ee4e563f0b5ceb69770a6d6f64e42a688b6/docs/tutorial-getting-started.md> (2020)
23. SoloKeys: Solo 2 monorepo. <https://github.com/solokeys/solo2> (2021)
24. W3C: Web authentication: An API for accessing public key credentials - level 1. W3C recommendation (Mar 2019), <https://www.w3.org/TR/2019/REC-webauthn-1-20190304/>
25. W3C: Web authentication: An API for accessing public key credentials - level 2. W3C recommendation (Apr 2021), <https://www.w3.org/TR/2021/REC-webauthn-2-20210408/>
26. W3C: Web authentication: An API for accessing public key credentials - level 3. W3C first public working draft (Apr 2021), <https://www.w3.org/TR/2021/WD-webauthn-3-20210427/>
27. Yubico: python-fido2 (Mar 2018), <https://github.com/Yubico/python-fido2>
28. Yubico: Discover YubiKey 5. strong authentication for secure login. <https://www.yubico.com/products/yubikey-5-overview> (Jul 2021)
29. Yubico: libfido2 (Jul 2021), <https://developers.yubico.com/libfido2>
30. Yubico: webauthn-recovery-extension. <https://github.com/Yubico/webauthn-recovery-extension> (2021)

The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity

Citation	L. Janzen, F. Putz, M.-A. Kaufhold, K. Straub, and M. Hollick. "The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity." In: <i>Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems</i> (Apr. 26–May 1, 2025). CHI '25. Yokohama, Japan: ACM, Apr. 2025. DOI: 10.1145/3706598.3714324
Contribution	This publication is discussed in Chapter 3. My contribution is described in the section " <i>Collaborations and My Contribution</i> " on page x.
Rights notice	© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is included in this dissertation in accordance with ACM author rights. The definitive Version of Record was published in <i>Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems</i> (CHI '25), DOI: https://doi.org/10.1145/3706598.3714324 .

The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity

Leon Janzen
Secure Mobile Networking Lab
TU Darmstadt
Darmstadt, Germany
ljanzen@seemoo.de

Florentin Putz
Secure Mobile Networking Lab
TU Darmstadt
Darmstadt, Germany
fputz@seemoo.de

Kolja Straub
Secure Mobile Networking Lab
TU Darmstadt
Germany
kstraub@seemoo.de

Marc-André Kaufhold
Science and Technology for Peace and
Security
TU Darmstadt
Germany
kaufhold@peasec.tu-darmstadt.de

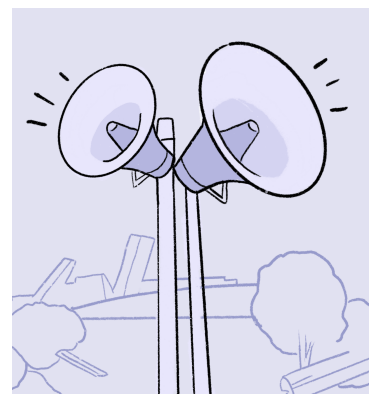
Matthias Hollick
Secure Mobile Networking Lab
TU Darmstadt
Germany
mhollick@seemoo.de



(a) You are taking a walk in **your city**.



(b) Suddenly, you feel several strong tremors and notice that everything around you is shaking.



(c) A siren sounds with a loudspeaker announcement: "Attention! The city has been hit by an earthquake. Please seek safety."

Figure 1: Crisis Scenario. This comic board introduced our participants to the assumed crisis scenario of our study, an earthquake. We individualized the first comic (Figure 1a) by replacing **your city** with the participant's city.

ABSTRACT

Using smartphone apps during crises is well-established, proving critical for efficient crisis response. However, such apps become futile without an Internet connection, which is a common issue during crises. The ongoing 6G standardization explores the capability to provide local cellular connectivity for areas cut off from the Internet in crises. This paper introduces to the HCI community the concept of cellular *island connectivity* in isolated areas, promising

a seamless transition from normal operation to island operation with local-only cellular connectivity. It presents findings from a survey ($N = 857$) among adult smartphone users from major German cities regarding their smartphone usage preferences in this model. Results show a shift in app demand, with users favoring general-purpose apps over dedicated crisis apps in specific scenarios. We prioritize smartphone services based on their criticality, distinguishing between apps essential for crisis response and those supporting routines. Our findings provide operators, developers, and authorities insights into making user-centric design decisions for implementing island-ready 6G communication.

CHI '25, April 26-May 1, 2025, Yokohama, Japan

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI Conference on Human Factors in Computing Systems (CHI '25)*, April 26-May 1, 2025, Yokohama, Japan, <https://doi.org/10.1145/3706598.3714324>.

CCS CONCEPTS

• **Computer systems organization** → **Availability**; *Reliability*; **Fault-tolerant network topologies**; *Cloud computing*; *Peer-to-peer architectures*; • **Computing methodologies** → *Self-organization*; • **Human-centered computing** → **Empirical studies in HCI**; • **Information systems** → *Distributed storage*; *Social networking sites*; • **Networks** → *Cloud computing*; **Mobile networks**; *Public Internet*; *Data center networks*; *Logical / virtual topologies*.

KEYWORDS

User-Centric, Smartphone Interaction, Island-Ready, 6G Communication, Smartphone Usage, Island Connectivity, Local Communication, Island Networking, Local-First, Decentralized 6G, Distributed Core Network, Crisis, Resilience

ACM Reference Format:

Leon Janzen, Florentin Putz, Kolja Straub, Marc-André Kaufhold, and Matthias Hollick. 2025. The User Perspective on Island-Ready 6G Communication: A Survey of Future Smartphone Usage in Crisis-Struck Areas with Local Cellular Connectivity. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*, April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA, 22 pages. <https://doi.org/10.1145/3706598.3714324>

1 INTRODUCTION

For the past 20 years, the use of social media during emergencies and natural disasters has dramatically increased [92], thus receiving substantial attention within the research fields of Human-Computer Interaction (HCI) and Computer-Supported Collaborative Work (CSCW). As more and more people own a smartphone and carry it with them at all times [28, 105], smartphones have become a central tool for users to cope with crisis scenarios. While social media and dedicated crisis apps have been well-researched in the community, e.g., by [33, 109], the role of other general-purpose smartphone apps in crisis response has not received much attention [101]. To motivate further efforts in this field, our first research question is:

RQ1: Which apps do citizens prefer to use for crisis-specific use cases?

One area of concern for smartphone apps during crisis response is that Internet outages often accompany crises, but many smartphone apps require an Internet connection for full functionality. For example, the resilience strategy of the German federal network agency considers natural disasters, large-scale cyber attacks, and destruction during armed conflicts as threat scenarios for telecommunications networks [12]. In such scenarios where crisis-struck areas are also affected by Internet outages, users cannot use smartphone apps that require Internet connectivity. Furthermore, cellular services, such as emergency calls, telephony, and cell broadcasts, are unavailable as they require connection to the cellular core network that is most likely hosted outside the affected area.

With the upcoming standardization of the sixth-generation technology standard for cellular networks (6G), there are proposals to distribute the core network, primarily to achieve low latencies [19, 21, 72, 82]. However, decentralizing 6G also increases the probability that crisis-struck areas have a core network in proximity,

which operators could use to maintain local cellular Internet connectivity within the affected area. Section 3 defines such areas as *islands*, as they form a self-contained network isolated from the outside Internet, and introduces the conceptual model of *island connectivity*. We envision users on islands using their smartphone apps locally, giving them access to crisis-relevant apps and letting them communicate with friends on the island. Other apps and services, however, would not be available on islands, e.g., smartphone apps that require remote server connections and telephony with people outside the island. Considering the findings of our narrative review [85], understanding the user-smartphone interaction on islands with local cellular connectivity is an unstudied research problem. We contribute to this field with our second research question:

RQ2: Which apps do citizens prefer to use in isolated areas with local connectivity?

We designed and conducted a survey ($N = 857$) to understand smartphone usage with island connectivity and answer research questions **RQ1** and **RQ2**. Section 4 discusses our study design and sample representative of general adult smartphone users in major German cities. Our findings indicate that there is demand for general-purpose apps in crisis response, e.g., users prefer messengers and news apps for communication and information. We studied which apps users prefer with island connectivity, finding a demand increase for some apps, e.g., messengers, news, and crisis apps, and a demand decrease for others, e.g., sports, shopping, and travel apps. Section 5 describes our findings in detail.

Putting the vision of *island connectivity* into practice poses many challenges and involves many stakeholders: Storage and compute capabilities must be available in a distributed manner, app developers must adapt their backend communication infrastructure to equip smartphone apps for island operation, and mobile network operators (MNOs) must prepare 6G to transition into island operation when a crisis-struck area is disconnected from the Internet. These efforts must be either justified by user demand or regulated by authorities. For instance, the resilience strategy of the German federal network agency requests that "citizens should be able to use a certain range of basic services even in the event of widespread [outages]" [12]. While they name emergency calls and cell broadcasts, they do not regulate how to prioritize apps when faced with limited capacity. Our third research question addresses this gap:

RQ3: How should operators prioritize smartphone services in isolated areas with local connectivity?

We construct a *smartphone service typology for island connectivity*, i.e., systematically classifying smartphone apps and cellular services according to their criticality for crisis response on islands. Our typology, discussed in detail in Section 6, distinguishes smartphone services that are crucial for crisis response, e.g., making emergency calls, receiving cell broadcasts, and using messengers, and those that enable crisis routines, e.g., finance apps, social media, and telephony. We compile our survey results and the typology into design implications that support stakeholders in making user-centric design decisions while realizing island-ready 6G communication.

2 RELATED WORK AND RESEARCH GAP

This section reviews related user studies and highlights research gaps in prior work that we address in our work.

2.1 App Usage During Crises

The first decade of crisis informatics research had a strong emphasis on social media, providing qualitative empirical case studies [83], identifying usage and role patterns [91, 104], designing algorithms for processing big crisis data [18, 54], and evaluating open source and social media analytics systems for enhanced situational awareness [3, 59, 60]. Many quantitative representative surveys examined citizens' usage patterns, perceived usefulness, and expectations towards social media in crises [93] and compared perceptions across European countries, North America, and China [94, 113].

However, considering the narrow focus of crisis informatics, researchers have suggested looking beyond social media [103]. While general-purpose apps and social media sometimes include functionality for community-oriented crisis response [109], multiple official so-called crisis and warning apps have been established across various countries of the Global North [47]. These often provide location-based warnings and tips on how to behave before, during, and after an emergency, and sometimes facilitate communication with emergency services [26]. Existing qualitative representative surveys with German citizens show increasing adoption of crisis apps, highlighting their perceived usefulness and the importance of different functionalities [45]. Researchers have examined the motivation and usability factors for using crisis apps [36, 108], including nudging approaches to enhance preparedness [46]. However, considering the findings of our narrative review [85], there is a lack of research examining the citizens' preferences for different apps for crisis-specific use cases. We contribute to this area by studying which apps citizens prefer to use for crisis-specific use cases and testing the demand for non-crisis apps during crises (RQ1).

2.2 Communication Networks During Crises

Existing studies on social media and crisis apps often lack attention to the underlying infrastructure requirements, such as energy and telecommunications, which are especially important in isolated areas. Not least, the 2021 Ahr Valley flooding showcased the sociotechnical complexity of warnings in a developed country like Germany. On the one hand, the acute warning period was too short; on the other hand, communication and electricity infrastructure failed during the floods and thus impaired further crisis communication and rescue operations [35]. Data from an online survey with citizens reveals that 35 % of the respondents from the federal state of North Rhine-Westphalia and 29 % from Rhineland-Palatinate did not receive any warning, highlighting the importance of effective administrative decision-making processes and resilient warning systems. Of those warned, 85 % did not expect severe flooding, and 46 % reported a lack of situational knowledge on protective behavior [111]. Another representative survey revealed that German citizens demand behavioral instructions (83 %), information on the reason for the failure (80 %), and the expected outage duration (78 %) from the respective infrastructure provider [58].

2.3 Related Notions of Island Connectivity

The idea of *islands* has a long history in networking and disaster research [17, 84]. Our interpretation of an island is that a region affected by a crisis and simultaneously disconnected from the outside Internet can maintain local Internet connectivity for users, e.g., to app servers hosted on the network edge, allowing users to continue using general-purpose apps on their smartphones. Section 3 elaborates on the conceptual model of island connectivity. In the following, we distinguish island connectivity from related notions.

A central problem of modern ICT for crisis communication is their dependence on centralized infrastructures [56]. Thus, it is essential to harden centralized infrastructures against natural hazards [67] and explore decentralized crisis communication. Related proposals exist to enhance the resilience of cellular networks, e.g., with self-organizing small cell networks [119] or network-in-a-box approaches [88]. These approaches are often limited to a specific part of the cellular network or require time to set up during a crisis. In contrast, island connectivity considers the entire 6G system, and we envision that island-ready 6G can seamlessly switch to island operation in case of a crisis. While related work studied decentralized and disruption-tolerant networks in smart rural areas [53, 68], the prevalence of Internet of Things (IoT) devices will be exceptionally high in smart cities, including uncrewed aerial vehicles (UAVs), smart cars, smart lanterns, and smartphones, for the establishment of multi-hop networks [49, 106]. Since applications and services will need to switch between different modes of centralized and decentralized operation, probably with varying functionalities available, HCI has to ensure a seamless user experience during these transitions [42]. Yet, we did not find any study of citizens' preferences regarding different app categories in isolated areas embracing the potential of island connectivity. To fill this gap, we investigate which apps users prefer to use in crisis-struck areas isolated from the Internet (RQ2).

2.4 Practical Aspects of Island Connectivity

Putting the vision of *island connectivity* into practice poses many challenges and involves many stakeholders. App developers must adapt their backend communication infrastructure to ready smartphone apps for island operation. MNOs must prepare 6G to transition into island operation when a crisis-struck area is disconnected from the Internet. These efforts must be either justified by user demand or regulated by authorities. To our knowledge, the user demand for islands is an unstudied research problem. Thus, the research community lacks an understanding of which app categories citizens prefer to use in isolated areas. Furthermore, authorities should regulate how MNOs prioritize smartphone apps when faced with limited capacity. However, no existing regulation guides MNOs in this endeavor [12]. We address this gap by constructing a smartphone service typology (RQ3).

In summary, **our study contributes to understanding citizens' smartphone usage in crisis-struck isolated areas**, motivating the demand for general-purpose apps for crisis-specific use cases (RQ1), reporting the demand for smartphone apps in isolated areas with local connectivity (RQ2), and suggesting prioritization

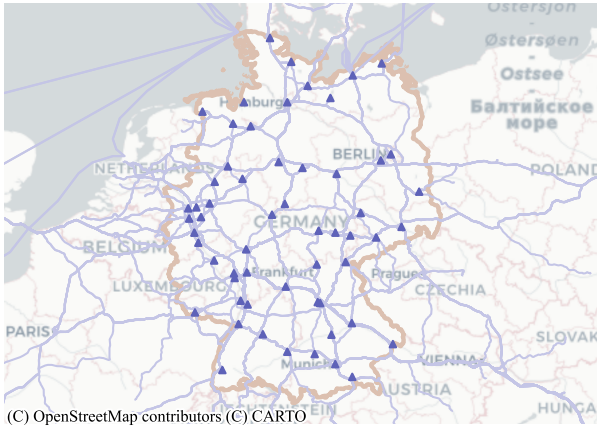


Figure 2: Internet Infrastructure in Germany. A beige line shows the German border, blue triangles highlight known physical interconnection facility locations, e.g., point of presences (PoPs), Internet exchange points (IXPs), or cable landing sites, and purple lines show the routes of submarine cables and the assumed terrestrial cable routes, following roadways, rails, and power lines [6, 31].

of smartphone services for these scenarios (RQ3). Section 3.4 complements our review of prior work by distinguishing our *island connectivity* model from related concepts.

3 THE CONCEPTUAL MODEL OF ISLAND CONNECTIVITY

In this section, we introduce the conceptual model of *island connectivity*, which is our study’s specific application area.

3.1 Background: Internet Connectivity

The Internet is a network of networks that builds a physical and logical graph around the globe [14]. The physical Internet comprises infrastructure nodes, e.g., Internet exchange points (IXPs), colocation centers, and point of presences (PoPs), connected by fiber conduits, e.g., terrestrial long-haul and submarine cables [6]. The Internet Protocol (IP) suite defines abstraction layers to enable logical end-to-end connections while abstracting from the concrete physical connections in use, e.g., allowing end-to-end communication between hosts based on IP addresses [75].

When users open a smartphone app that requires Internet connectivity to connect to an app server, the app usually stores the server’s domain name or IP address. The smartphone first connects to a 4G/5G base station to establish a connection to the app server, from where the user data is augmented with control plane data and sent to the MNO’s core network [1]. If the core network successfully processes the control plane data, e.g., handling authentication and session management, the user data proceeds to the Internet. Alternatively, smartphones can connect to the Internet via a Wi-Fi router. On the Internet, the user data is routed toward the app server’s IP address until it reaches its destination.

Current cellular networks have a mostly central architecture with limited core network gateways per country, which can cause path inflation [118], i.e., a diversion from the shortest route due to the requirement of passing the core network. Path inflation is especially pronounced if the app server is hosted on the local network edge close to the user but the closest core network gateway is remote. The physical Internet backbone in Germany is considered critical infrastructure and, therefore, is not publicly available [13]. However, Figure 2 visualizes how the Internet Geographic Database (iGDB) [6] interpolates connections between known Internet infrastructure, assuming that cables are often installed along existing networks such as roadways, rail, and power lines [31].

3.2 Islands: Crisis-Struck Areas Isolated From the Internet

The communication infrastructure is generally expected to be resilient against outages. Still, cellular network outages are actively considered in the resilience strategy of the German federal network agency [12], and previous outages have isolated individual areas [74]. Another complication is that user traffic has to traverse the core network before being forwarded to the Internet, but the chances of a cellular core network being available in the affected area are negligible due to the low geo-redundancy of cellular core network gateways [118]. These gateway locations of the German MNOs, i.e., 1&1 AG, Deutsche Telekom AG, Telefónica Deutschland Holding AG, Vodafone GmbH, are unknown. Still, observations from other countries [98] suggest only a handful of gateways per MNO exist in Germany. As a result, users in affected areas likely cannot use 4G/5G connectivity to reach these servers despite the locally intact physical Internet and servers being available on the local network edge. We aggregate the notion of a crisis-struck area that is isolated from the outside Internet in the definition of *islands*:

An **island** is a crisis-struck area isolated from the outside Internet, preventing citizens from using smartphone apps that require an Internet connection to remote servers.

Several terms exist in the literature to describe deviations from regular operation. This paper uses the term *crisis* as a hypernym for crises, emergencies, catastrophes, and disasters [32]. Islands can be of various shapes and sizes, ranging from small rural areas to major cities. One example is the 2021 flooding in Western Europe, where the floods destroyed fiber links installed under roads and bridges and, thus, disconnected large parts of the Ahr Valley from the Internet for days to weeks [74]. While satellite connectivity has successfully been used in crisis-struck areas to provide citizens with Internet access, the achievable bandwidth with a few satellite links is insufficient to reconnect an entire region to the Internet [78]. Another common approach is microwave transmissions, closing the gap of broken fiber links with point-to-point directional antennas [27]. However, their applicability depends on the type of failure responsible for the isolation, and setting up directional antennas can take several days.

3.3 Island Connectivity: Local Fall-Back Cellular Networks

With multi-access edge computing, operators increasingly aim to move content and computational power closer to the user by hosting servers on the network edge, e.g., co-located with IXPs or in a data center within the user's proximity [8]. These advances reduce latency as Internet traffic destined within the user's proximity instead of being routed to a remote data center. Still, they do not fully apply to cellular network users due to *path inflation* [62, 118, 120]. With 6G on the horizon, there are proposals [19, 21, 72, 82] to tackle the bottleneck of path inflation by distributing core network functions, e.g., co-located at suitable Internet nodes or generally within the data centers of major cities.

A higher geo-redundancy of core network gateways would improve latencies and the chances of *islands* having a core network available. As a result, island users could use crisis-relevant cellular services, i.e., making emergency calls, receiving cell broadcasts, and contacting close ones on the island by text or phone call. In addition, a core network on the island facilitates cellular Internet connectivity, thus enabling users to use smartphone apps for crisis response. In this context, app servers must be deployed within the island's boundaries, aligning with the multi-access edge computing trend. To this end, we distinguish *local* and *remote* connections, the former staying within the island's bounds, while the latter destined at a server outside the island. We distinguish two operation modes of the cellular network: *normal operation* in non-crisis times and *island operation*, providing island connectivity on islands.

Island connectivity enables users on an island to use the cellular network within the island's bounds, facilitating the use of cellular services (e.g., emergency calls and cell broadcasts) within the island and access to servers on the islands's network edge (e.g., multi-access edge computing).

With island connectivity, users can access the cellular network within that island's boundaries, i.e., they can use apps with app servers hosted on the local edge but not apps with remote app servers; they can communicate with friends on the island but not with friends outside the island. One requirement of island connectivity is an intact power supply, another critical infrastructure that can fail during natural disasters or be affected by dedicated cyber attacks. The German federal network agency reports that many MNOs already have emergency and backup power systems. Still, these are often severely limited, providing only a few hours of electricity. For the future, it recommends considering renewable energies as on-site backups and for end users to have a mobile power supply in stock to recharge mobile devices in crisis scenarios [12]. Implementing the recommendations from these guidelines ensures that island users can use island connectivity when available [1].

The realization of island connectivity poses various challenges to operators, developers, and authorities. To our knowledge, the HCI community has not studied the user-smartphone interaction in the specific application area of islands with local cellular connectivity. This paper aims to understand citizens' smartphone usage on islands and derive design criteria for apps and 6G to facilitate user-centric design decisions while realizing island connectivity.

3.4 Distinction from Related Technical Concepts

The design principles underlying the Internet are simple and mostly unchanged since its inception [75], as the principles of routing, reliability, and resolution allowed the Internet to grow from a few computers to today's global scale. Originally, the Internet was island-ready due to the mostly static configurations and its decentralization, but over time, lost its island capabilities since essential functions, e.g., domain name resolution, were centralized. Some technical concepts propose to re-integrate decentralized capabilities into the Internet, e.g., peer-to-peer systems as application layer overlays [96], content delivery networks (CDNs) [86], the decentralized web [90], and campus networks [97]. Out of these concepts, only campus networks are related to the cellular network. Campus networks are private wireless communication networks connecting the users of an organization in a dedicated geographical area, e.g., a university or an industrial site [22]. The main objective of 5G campus networks is to use the advantages of cellular connectivity, i.e., higher speed and lower latency. Campus networks are designed for static settings, i.e., the operator registered the campus network with the authorities and got permission to operate the campus network in a specified frequency band in a specified geographic area. In addition, only a pre-defined set of SIM cards can connect to the campus network. As such, the concept of campus networks is clearly distinct from island connectivity by two factors: First, island connectivity applies to the *public* cellular network and should be usable for all users. Second, island-ready 6G communication can *dynamically* switch from normal operation to island operation when a core network replica detects a disconnection from the outside Internet. Nevertheless, non-public networks can benefit from the advances of island connectivity, e.g., when multiple campus networks hosted at different sites that are usually connected through the Internet become isolated due to outages.

3.5 Relevance of Island Scenarios

The limited period of island scenarios, infrastructure requirements, and recent proposals for alternative technologies to recover connectivity provoke the question of how relevant island connectivity is for the research community. First, information and communication technology (ICT) is a critical infrastructure [9, 24, 25], underscoring the crucial value of connectivity for crisis response. In particular, the cellular network plays a pivotal role in crisis response, providing Internet connectivity, emergency calls, and cell broadcasts.

We argue that **island connectivity can have a critical impact during the first few days of a crisis or even across multiple months in the case of a large-scale conflict**, such as in the ongoing Russo-Ukrainian War, where critical infrastructures are operational targets by cyber or physical attacks [102, 117]. Past crises underscore the scenario's plausibility and that alternative recovery means do not suffice to reconnect entire areas. Understanding how citizens prefer to use their smartphones on islands is crucial to supporting the ongoing development of 6G. Operators, developers, and authorities face many challenges in implementing this vision, and contributions from the HCI community can help facilitate user-centric design decisions.

4 METHOD

Six researchers and one designer were involved in the design phase of this study, cooperating in multiple workshops to combine knowledge from the fields of resilient cellular networks, HCI, and CSCW.

4.1 Material

Our goal was to study the user-smartphone interaction in island connectivity. We realized early in the design phase of our study that the novelty and complexity of island connectivity as a conceptual model might be hard to grasp. To avoid overloading participants, we decided to fix a concrete example for islands in our study. Thus, **our survey investigates smartphone usage in major German cities**, simplifying the abstract concept of island connectivity to the concrete instantiation of local connectivity in an isolated urban area. We motivate this choice by the increasing share of the worldwide population living in urban areas, as 71 % of the German population lived in metropolitan areas in 2022 [29]. In addition, cities play a central role in the Internet, and most of the physical Internet infrastructure aligns around cities [6]. As such, cities are prime candidates to host replicas of the distributed 6G core network, which aligns with the concept of resilient cities [51, 61, 121].

4.2 Questionnaire Design

Section A holds a full translated transcript of the questionnaire. It comprises five parts that incrementally familiarize participants with our study scenario.

Part 1: Consent and Screening. We informed participants about the procedure and goals of the study before they gave their informed consent to participate in our study. We collected age, gender, education, income, and federal state (Q1-Q5) to steer recruitment towards representativity. To recruit citizens with prior smartphone experience, we screened whether the participants had lived in a major German city for at least one year and owned a smartphone during that time (Q6). For eligible participants, we collected the city name and the time the participants lived there (Q7-Q8). Participants confirmed they had owned and used a smartphone during that time (Q9). We implemented the city name into the subsequent questions to tailor the questionnaire to the participant, e.g., for participants who had lived in Berlin, the caption of the comic depicted in Figure 1a said: "You are taking a walk in *Berlin*."

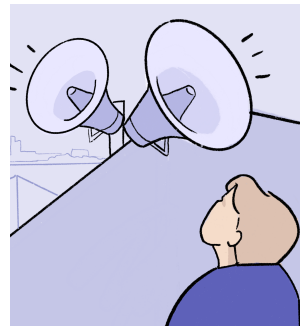
Part 2: Everyday Smartphone Usage. We asked the participants to estimate their daily smartphone screen time in hours (Q10), guiding them to find the corresponding Android and Apple iOS statistics. We studied everyday smartphone usage with positional voting [99], where participants chose a subset of 23 categories and positioned the chosen app categories by their daily usage frequency (Q11). There is no generally accepted smartphone app categorization in the literature, so we based our categorization on the Apple App Store [7]. We refined it to fit the survey's scope by splitting the *Social Networking* category into *Social Media* and *Messengers*, combining *Photo and Video* and *Graphics and Design* into *Creativity*, introducing categories for *Crisis* and *Web Browsers*, and omitting *Business apps*, *Developer Tools*, *Kids*, *Lifestyle*, and *Safari Extensions*.



(a) On your way home, you want to let your friends know that you are okay. As you have no signal, you cannot send the message.



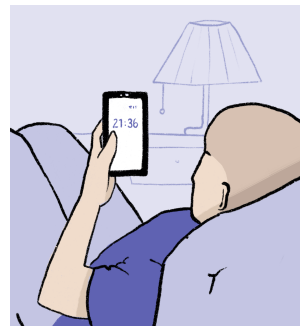
(b) Even sending via Wi-Fi from a nearby restaurant does not work.



(c) A siren sounds again with a loudspeaker announcement: "Attention! The city has been disconnected from the Internet. Therefore, communication is currently not possible."



(d) The mobile network operator was able to partially restore the network, so communication within the city is possible again, but not to the outside.



(e) At home, you want to use the restored connectivity to inform your friends within the city that you are okay.

Figure 3: Local Connectivity Model. This comic board introduced our participants to the conceptual model of local connectivity in isolated cities. The message in Figure 3a translates to "I'm fine!". The text in Figure 3d translates to "A few hours later."

Part 3: Crisis Smartphone Usage. We used a comic board (Figure 1) to introduce the assumed crisis scenario, an earthquake in the participant's hometown. We asked the participants to hypothetically answer the remaining survey with this crisis scenario in mind. Comicboarding [80] is a technique to break down complex information when co-designing with children. We chose comic boards to explain the study scenario to our non-expert study sample, anticipating the complexity of the local connectivity model in isolated cities. To study **RQ1**, we chose seven crisis-specific use cases (Q12) and asked which apps participants would use to handle them: Receiving warnings (U1), getting safety tips (U2), contacting emergency services (U3), looking up information (U4), communicating to others (U5), navigating in the city (U6), and distracting themselves (U7). We chose these use cases because they were well-established in related work: [45, 93, 95] study U1 - U5, [101, 115] study U6, and [101] studies U7. Each participant was assigned three of the seven use cases at random to reduce their effort, resulting in 348 to 388 answers per use case (median 364; IQR 19.5).

Part 4: Crisis Smartphone Usage with Island Connectivity. We showed a second comic board (Figure 3) to introduce island connectivity into the study scenario: Internet connectivity is recovered but limited to the participant's city, i.e., communication is only possible within the city, and content is available if hosted on local servers within the city but not from remote servers hosted outside the city. We showed a video (38 s) to clarify the scope and limits of island connectivity. We provided an alternative video description for participants unable to watch videos or encountering technical problems. We asked participants to answer the remaining questionnaire hypothetically, assuming the crisis was ongoing for several days with local connectivity. For **RQ2**, we repeated the positional voting category for smartphone usage in isolated cities (Q13).

Part 5: Control Variables and Data Quality. We collected the following control variables (Q14-Q17): Prior experience with crises, prior experience with cellular/landline Internet outages, prior activity in relevant emergency forces, e.g., as a firefighter, police officer, or first responder, and the affinity for technology interaction (ATI) scale [37] with an implemented instructed response item (IRI) [77]. Finally, participants could indicate technical issues and additional comments in open statements (Q18-Q19). The survey ended with a self-reported single item (SRSI) UseMe [15] (Q20), informing the participants that the survey was an integral part of our research and asking them if we should consider their responses for our studies.

After evaluating the results of our survey, we augmented the findings of **RQ1** and **RQ2** with applicable technology standards [1] and resilience strategies [12] to construct a smartphone service typology for island connectivity (**RQ3**).

4.3 Data Analysis

We used the statistical software R 4.4.1 [89] and an alpha level of 0.05 for the quantitative data analysis. We corrected for multiple tests with the Bonferroni method [2] to evaluate significant associations. We used the well-known Borda count [76] to analyze our participants' positional voting results. We used inductive coding for the open statements (Q18 - Q19) and all "Other" responses. Two researchers independently constructed initial codebooks and

merged them through discussions. Two researchers deductively coded all questions according to that codebook, reaching satisfactory inter-coder reliability (mean Krippendorff's Alpha = 0.9985, minimum 0.997) [65]. We solved the remaining coding differences through discussions. We generally report absolute and relative numbers to classify our results. We report median values and the inter-quartile range (IQR) as these statistics are more robust against outliers. We report the effect sizes for tests to find significant and relevant associations. Upon acceptance of this paper, **we will release a replication package that contains the survey dataset, our evaluation scripts, and the coding system** for readers to verify, reproduce, and replicate our findings.

4.4 Recruitment and Study Sample

We chose GapFish to handle recruitment because they are an ISO-certified panel provider cooperating with more than 500 000 panelists from Germany, Austria, and Switzerland [40]. In coordination with GapFish, we aimed to recruit a representative sample of smartphone users in major German cities based on age, gender, education, and income. To avoid effort for unfitting participants, GapFish invited panelists over 18 who owned a smartphone and lived in a German city with more than 100 000 citizens. We collected demographic information early in the survey (Q1-Q5) to avoid effort for participants who did not fit the quotas. Of the 2187 participants who started the survey, 986 participants completed the questionnaire. After removing participants who failed SRSI UseMe ($n = 64$), failed the IRI ($n = 58$), stated invalid education ($n = 4$), stated an invalid major city ($n = 2$), or commented hate speech ($n = 1$), we report $N = 857$ participants. Table 4 describes our study sample regarding demographic and control variables. The depicted categories simplify the identification of semantic groups, e.g., younger participants or participants with medium technology affinity.

Representativity. We recruited a study sample representing smartphone users from major German cities with the following limitations: Our study sample does not include participants under 18 or over 69 and has a smaller share of participants with low or high education compared to the German population. GapFish only distinguished male and female panelists in their database, so we could only steer gender representativity regarding men and women. Still, four participants openly identified as non-binary. All participants were panelists who had lived in a major German city for at least one year and had owned a smartphone during that time. We acknowledge these deficits and argue that **our study sample is representative of adult smartphone users in major German cities** considering the above limitations.

Significant associations. We found significant associations between our demographic and control variables using Pearson's chi-squared test [87] after correcting for multiple tests with the Bonferroni method [2]. Income and education showed a significant association ($\chi^2(6) = 136.94, p < 0.001, V = 0.28$) as did age and screen time ($\chi^2(4) = 98.2, p < 0.001, V = 0.239$). These associations represent medium effect sizes, according to Cohen [20]. We found further significant associations with small effect sizes between age and education ($\chi^2(4) = 33, p < 0.001, V = 0.139$), citizen duration and age ($\chi^2(4) = 67.6, p < 0.001, V = 0.119$), citizen duration and

screen time ($\chi^2(4) = 27, p < 0.001, V = 0.126$), and gender and technology affinity ($\chi^2(4) = 19.7, p < 0.001, V = 0.107$).

4.5 Biases and Limitations

Naturally, every study has limitations due to the decisions made in the design process. In the following, we motivate selected design decisions and discuss which biases and limitations they might cause.

Study sample. We hired an ISO-certified panel provider to recruit and compensate suitable participants. While recruiting only panelists to complete our questionnaire might introduce a bias against non-panelists, it facilitated steering our study sample towards representativeness for smartphone users from major German cities regarding gender, age, education, and income, thus preventing sampling and selection biases [11, 48]. Only participants who had lived in a major German city for at least one year and had used a smartphone during that time could complete our questionnaire. This screening might introduce a survivorship bias [34] against Germans who have not lived in a major German city for at least one year and German citizens who have not owned a smartphone during that time. We opted to screen out these groups to align the sample with our study subject, smartphone usage in isolated cities.

Questionnaire design. We ordered the questionnaire to incrementally familiarize participants with our study subjects to avoid question order bias [100]. Considering the risks of survey time fatigue [70], we determined the dependent variables Q11 - Q13 as early as possible, showed each participant only three random of the seven use cases in Q11, and included two comic boards [80] and a video before Q13. On average, it took participants 12 min 38 s (median 10 min 26 s; IQR 6 min 11 s) to complete our survey. We recorded many outliers, with the fastest completion taking 2 min 24 s and the slowest completion 10 h 43 min 20 s. Following the suggestions of [41], we did not remove speeding participants. We implemented a combination of single-choice and multiple-choice selections, number inputs, positional voting, yes/no questions, Likert scales, and open statements to mitigate extreme or neutral response biases and an acquiescence bias [81]. Only Q15 used an unbalanced scale, but for this question and, generally, where possible, we provided "No answer" or "Other" options to mitigate a conformity bias [66]. We used neutral, non-leading questions and explanations to avoid anchoring biases [112]. Interviewing the general public about a complex research scenario can introduce a cultural bias [50], so we used comic boards [80] to convey complex information, an explanatory video, and feedback from the pilot study (Section 4.6). We hosted the questionnaire with *LimeSurvey* [73] to ensure accessibility on various end devices and provided alternative descriptions for the video and comic boards to mitigate technical biases [30].

Data quality. We opted for an online survey to gather knowledge from a broad study sample and focus on quantitative results. Ensuring data quality in online surveys can be challenging because participants might be inattentive while filling in the survey, not understand the study subject, or face technical issues. We anticipated these challenges during in our study design by equipping the questionnaire with an SRSI UseMe and an IRI, which led to the removal of 122 inattentive participants. In addition, we implemented comic boarding and an accessible explanatory video with a transcript and

description. While one participant (P162) acknowledged that the study subject "was visionary and hard to empathize with," other participants described the questionnaire as "easy to follow" ($n = 5$), positively mentioning the structure, video, and comics ($n = 5$). At the end of the survey, we asked participants if they had faced technical issues. 98.95 % did not describe technical problems, and five participants (0.58 %) had issues sorting the app categories in Q11 and Q13. Four participants (0.47 %) had technical issues with the video, one of which (P402) mentioned that it was "great there was a description below the video." One participant (P60) proposed expanding Q16 with options for prior and retired emergency forces and hospital staff, which we suggest for future surveys.

4.6 Ethical Considerations

We tested our survey in a pilot study with $N = 29$ participants. Their valuable feedback led to several improvements, including a shorter survey length, more precise questions, and clarifications.

We adhered to local ethical and legal standards while designing, conducting, and evaluating this study. Our university's institutional review board reviewed and approved this study. GapFish compensated the panelists for their participation in our survey. As stated in the survey, we did not report participants who failed the IRI or self-reported with the SRSI UseMe to GapFish. Despite GapFish recruiting only adult panelists, two participants reported an age under 18 and were excluded before completing the questionnaire. We collected demographic information in clusters to protect participants' privacy. We informed all participants about the survey's purpose and data collection while adhering to the General Data Protection Regulation. We warned participants that the survey's subject included crises, which might be a sensible topic for them. Participants agreed to consent before starting the survey, and they could abort their participation at any time without explanations.

5 RESULTS

This section reports our survey results, addressing each research question individually. We supplement our findings with quotes from the open statements to highlight interesting aspects.

5.1 App Demand for Crisis-Specific Use Cases

Our first goal is to understand which apps citizens intuitively prefer to use for crisis-specific use cases (RQ1). Notably, the list of use cases covered by our study is not exhaustive, and other use cases might yield different results. We decided to include U1 - U7 because related work [45, 93, 95, 101, 115] studied similar use cases but not with a focus on citizens of major German cities.

Immediately after introducing the crisis scenario with the comic board depicted in Figure 1, we showed each participant three of the seven use cases. We asked them which apps they would use in these scenarios. Table 1 visualizes the results with a row per use case (U1 - U7). On average, participants selected 2.03 app categories per use case (median 1.67; IQR 1.67). The most often selected app category across the $857 \cdot 3 = 2571$ votes was *crisis apps*, selected 1066 times and scoring highest for receiving warnings (U1) and getting safety tips (U2). For the remaining use cases, participants preferred other apps over crisis apps: *News apps* for looking up information

Table 1: App Demand for Crisis-Specific Use Cases (RQ1). This heat map depicts the seven use cases (U1 - U7) and the ten most selected app categories. For each use case, it shows the percentages for the top three app categories, where bold text highlights the most preferred app category. The bottom row M_c shows the mean frequency of how often a category c was mentioned across all use cases. The rightmost column M_u shows the mean number of categories selected for use case u .

Use case \ Category ¹	N_u	Crisis	Messengers	News	Social Media	Web Browsers	Navigation	Weather	Reference	Music	Entertainment	M_u	None
Warnings (U1)	380	65 %	27 %	42 %								2.12	5 %
Safety tips (U2)	356	61 %		39 %		27 %						2.28	5 %
Emergency call (U3)	375	48 %	37 %			13 %						1.56	16 %
Information (U4)	348	53 %	32 %	60 %								2.42	2 %
Communication (U5)	364	22 %	72 %		35 %							1.88	6 %
Navigation (U6)	388	20 %	14 %				66 %					1.64	3 %
Distraction (U7)	360		37 %		27 %					27 %		2.42	9 %
M_c	367.29	41 %	35 %	29 %	19 %	15 %	14 %	9 %	7 %	4 %	3 %		

¹ The following app categories did not make the top 10 regarding M_c and were omitted in the heat map for better readability (in descending order according to M_c): Utilities, Games, Education, Books, Health & Fitness, Medical, Travel, Finance, Food & Drink, Creativity, Productivity, Shopping, Sports.

(U4), *messengers* for communication (U5) and distraction (U7), and *navigation apps* for navigating in the city (U6). While crisis apps scored highest for contacting emergency services (U3), many participants preferred emergency calls for this use case, selecting "No app" or using the "Other" option to describe telephony (16 %). Overall, this highlights that it takes more than crisis apps to serve all crisis-specific use cases and that users prefer using general-purpose apps in crisis scenarios.

5.2 App Demand on Islands

For **RQ2**, we focus on two app types: Apps that our study sample ranked high in island operation and apps with strong trends from island operation to normal operation. Our questionnaire featured two positional voting questions to learn about the participants' app demands in normal operation (Q11) and island operation (Q13). In this context, *normal operation* refers to a non-crisis scenario with global Internet connectivity, and *island operation* refers to the scenario where only island connectivity within the island is feasible. Notably, many methods exist to compute ranks from positional votes. We used the Borda count [76], a widely recognized and straightforward method, as it generally produces a broadly acceptable consensus based on individual rankings. Since using a different method could yield different results, we also publish our raw data to enable other researchers to apply alternative methods. Table 2 compares the positional voting results of all apps in normal and island operation, and Figure 4 takes a closer look at the vote distribution for apps with strong trends.

Apps with Unchanged Demand In Island Operation. Ten apps scored in the same rank for normal operation (Table 2) and island operation (Table 2) or with a difference of one rank. Table 2 does not show arrows for these apps for better readability. *Messengers* scored highest in both operation modes. Most participants ranked

messengers near Rank 1 in Q11 and Q13, resulting in a density peak at the lower left corner of Figure 4a. From RQ1, we understand that citizens prefer messengers to communicate with other people for communication and distraction. However, island connectivity limits communication to people in the city. For example, P457 commented, "As I do not have any friends or acquaintances in Düsseldorf, [...] I would not want to communicate with anyone within the city in that scenario." We found similar unchanged but low demands for *creativity* apps, causing the density peak in the upper right corner of Figure 4b. Most participants who voted for creativity apps in Rank 23 in Q11 voted for them in Rank 22 in Q13. Apps related to weather, navigation, entertainment, music, utilities, reference, games, and health & Fitness showed a similarly unchanged demand.

Apps with Increased Demand in Island Operation. Five apps scored higher for island operation than normal operation. As a crisis-specific medium, *crisis apps* climbed 18 ranks from normal operation to island operation. This is reflected in Figure 4c, where a density peak in the bottom-right corner indicates that crisis apps were voted near Rank 1 in island operation but near Rank 23 in normal operation. However, the second peak in the top-right corner indicates a share of our study sample with low demand for crisis apps regardless of the operation mode. We observe a trend of participants moving from Ranks 6 - 10 in normal operation to Ranks 1 - 5 in island operation, visible as a white area in Figure 4c. This trend suggests that participants who occasionally use crisis apps in their daily lives use them more frequently on islands, which aligns with our findings regarding **RQ1** reporting high demands for crisis apps to receive warnings and safety tips, contact emergency services, and seek information. We found weak positive demand changes for apps related to news, food & drink, books, and medical.

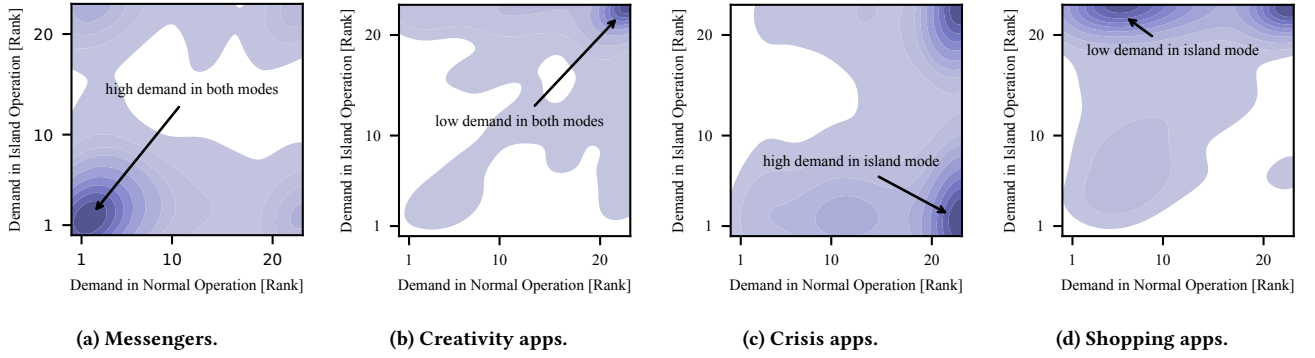


Figure 4: Positional Voting Distribution for Apps with Strong Trends. Each figure compares the vote distribution of the app demand in normal operation (Q11) and island operation (Q13) in a bivariate kernel density estimation (KDE) plot [110]. The annotated density peaks underscore strong trends identified in Table 2. For instance, the soaring demand for crisis apps in island operation compared to normal operation (Figure 4c) causes a density peak in the lower-right corner, as participants voted for crisis apps near Rank 23 (low demand) in Q11 and near Rank 1 (high demand) in Q13. Participants who prefer not to use crisis apps in normal nor island operation account for the smaller density peak in the upper-right corner.

Table 2: App Demand on Islands (RQ2). This table compares how apps ranked in normal operation (Table 2) and island operation (Table 2). Dotted arrows indicate trends of two ranks or more, and bold blue arrows highlight strong trends.

(a) Q11: Normal Operation	(b) Q13: Island Operation
App Category	App Category
1 Messengers	1 Messengers
2 Social Media	2 News
3 Web Browsers	3 Crisis
4 Weather	4 Social Media
5 News	5 Weather
6 Shopping	6 Web Browsers
7 Navigation	7 Navigation
8 Entertainment	8 Entertainment
9 Music	9 Music
10 Finance	10 Utilities
11 Utilities	11 Reference
12 Reference	12 Finance
13 Productivity	13 Food & Drink
14 Games	14 Games
15 Food & Drink	15 Productivity
16 Travel	16 Books
17 Health & Fitness	17 Health & Fitness
18 Education	18 Medical
19 Books	19 Shopping
20 Sports	20 Education
21 Crisis	21 Travel
22 Medical	22 Creativity
23 Creativity	23 Sports

Apps with Decreased Demand in Island Operation. Eight apps scored lower for island operation than normal operation. The demand for *shopping apps* plummeted the most, described by the density peak in the upper-left corner of Figure 4d. Most participants who voted for shopping apps in the top-10 ranks in Q11 moved to the lowest ranks in Q13. Similarly, shopping apps show a secondary peak as some participants showed no demand for shopping apps at all. Social media, web browsers, finance, productivity, education, travel, and sports apps showed a similar demand decrease.

Apps with High Demand in Island Operation. Three apps complement the top-5 ranks for island operation besides the aforementioned messengers and crisis apps. *News apps* climbed from Rank 5 in Q11 to Rank 2 in Q13. From **RQ1**, we understand that our participants prefer news apps for looking up crisis-relevant information as well as receiving warnings and safety tips. The increased demand in island operation underscores the importance of news apps for islands. *Social media* descends from Rank 2 in normal operation to Rank 4 in island operation, as the use of social media is limited to the island in island connectivity. The results of **RQ1** show that participants primarily prefer social media for communication and distraction. We find different demands for messengers and social media among our study sample, which confirms our design decision to split Apple’s *Social Networking* app category into *Social Media* and *Messengers*, as the demands of these app categories seem to differ on islands. *Weather apps* can play a crucial role in crisis-struck isolated areas, especially considering natural disasters.

5.3 Smartphone Services Typology for Island Connectivity

Our contribution to **RQ3** is a smartphone service typology for island connectivity. In this context, *smartphone services* generalize the functionality provided by smartphone apps and services of the cellular network. We construct this typology by combining

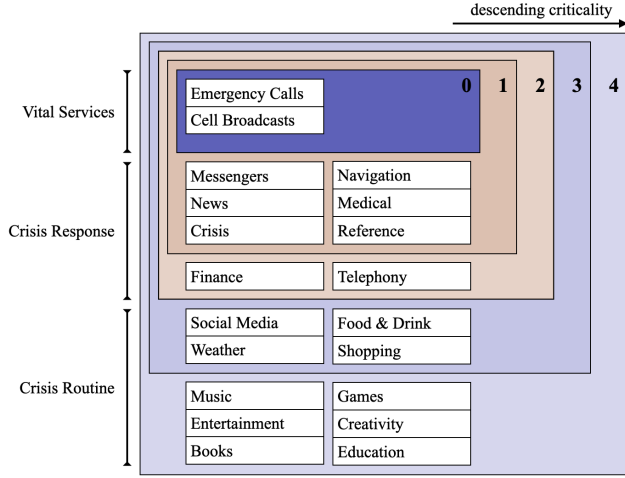


Figure 5: Smartphone Service Typology for Island Connectivity (RQ3). This prioritization organizes smartphone services in tiers based on their criticality for islands. Tier 0 contains vital services for islands, Tiers 1 - 2 focus on immediate crisis response, and Tiers 3 - 4 focus on crisis routine.

the fieldwork-informed knowledge from our survey with practice-informed design knowledge from authorities [12, 16] and applicable technology standards [1]. Figure 5 visualizes how the typology classifies smartphone services into five tiers, prioritizing how critical the included services' availability is on islands. The typology is based on our findings for research questions **RQ1** and **RQ2**, which we augment with demographic tendencies observable in our study sample. Figure 6 depicts a selection of tendencies and Table 3 shows the median ranks of all categories by semantic groups. While Tiers 0 and 1 focus on the immediate crisis response, distinguishing paramount smartphone services and crucial apps, Tiers 2 and 3 consider cities in crisis routine, identifying required enablers and apps for day-to-day crisis life. Tier 4 is about citizens' well-being during crises, discussing the potential of nice-to-have apps during crisis recovery. In the following, we discuss each tier individually:

Tier 0: Crisis-Critical Cellular Services. Making *emergency calls* and receiving *cell broadcasts* are central pillars of the German federal network agency's resilience strategy [12]. European and federal laws regulate the availability of emergency calls and cell broadcasts, e.g., users can make emergency calls without money on their account and use another MNO's network [79], highlighting its crucial importance for crisis response. The results of our survey support this assessment, as despite not having asked about it explicitly, 16 % of the participants reported they would not use smartphone apps but telephony to make an emergency call. The services of Tier 0 should be the absolute minimum that every MNO maintains on islands.

Tier 1: Services for Immediate Crisis Response. We found that *messengers* are the most demanded app category on islands, covering crisis-specific communication and distraction use cases. Communicating with close ones affected by the crisis can alleviate

stress and avoid extensive movement through the crisis-struck area [4, 107]. Messengers show similar rank distributions for all considered variables (Table 3). All semantic groups in our study sample showed similarly high demand for messengers (median Rank 2), e.g., see Figure 6 to compare rank distributions of groups based on age and education. Our participants preferred *news apps* to look up information. While providing access to global or generally external newspapers is hard, if not infeasible, at least local newspapers should be available on islands. As the name suggests, *crisis apps* are dedicated to crisis scenarios, including islands. While we classify cell broadcast as a Tier 0 service, which is paramount to broadcasting urgent information to all citizens, crisis apps can provide more detailed warnings and additional information about the crisis [47]. Similarly to messengers, the rank distribution of news and crisis apps shows minimal variance between semantic groups with medians around Rank 3. *Navigation apps* can help citizens navigate isolated areas depending on the crisis. Local map data can be hosted on the island, and local crisis scenarios usually do not affect GPS satellites. Furthermore, should GPS be unavailable, e.g., due to a targeted cyber attack, the radio-based positioning of 6G can be utilized [10]. These apps become especially useful when augmented with crisis-relevant details, e.g., available and blocked roads, water sources, or places with Wi-Fi connectivity [42]. *Medical apps* can be vital for patients relying on dedicated apps, e.g., monitoring insulin for diabetes patients. Maintaining the availability of medical apps can prevent crisis-unrelated emergencies of patients, avoiding additional stress on first responders [52]. We observed several tendencies regarding medical apps in our study sample: Participants with a short or medium residence time (median Rank 8) and younger participants (median Rank 9) showed lower demands for medical apps compared to participants with long residence times (+6 ranks), and older participants (+4 ranks), respectively. *Reference apps* can provide crisis-relevant information for citizens, e.g., how to filter water or prepare food without electricity [69].

Tier 2: Services with Critical Purpose. *Finance apps* give citizens access to their finances, which is a critical infrastructure in many countries [114]. Our study does not find an increased demand for finance apps on islands. Still, payment apps and, ultimately, the critical infrastructure of finances must be available on islands to enable an organized day-to-day crisis life, which poses substantial challenges as we elaborate in Section 6.2. We observe higher demand for finance apps in high-income and medium-income participants (median Rank 7, both) than low-income participants (+4 ranks). *Telephony* enables citizens to make phone calls within islands and is another enabler of crisis routines [101].

Tier 3: Services for Crisis Routine. Our survey shows a high demand for *social media* and *weather apps* on islands. Since the terrorist attacks of 9/11, the HCI and CSCW communities have observed an increasing use of social media in crises [92]. Related studies investigated use cases for social media during crisis response, e.g., to organize community efforts [104] or to use media from crisis-struck regions as input for situational awareness [116]. As such, individual use cases for social media during crisis response

exist even though social media's primary purpose is in crisis routine. *Weather apps* have become an integral part of our everyday lives, not necessarily opening the weather app itself but reading snippets of weather information on lock screens, info panels, or menu bars. *Food & drink* and *shopping apps* might not seem as crucial for crisis routines as other apps. However, food supply is a critical infrastructure in Germany, and shopping apps can mean shops offering crisis-relevant equipment.

Tier 4: Services for Crisis Well-Being. *Music apps* play a central role for citizens seeking distraction from the crisis [122]. In our sample, younger adults had a higher demand for music apps during crises (median Rank 5) than middle-aged adults (−2 ranks) and older adults (−7 ranks). *Entertainment apps* for TV shows or movies, *books* for reading, and *games* for gaming are further sources of distraction [43]. While our survey found little demand for *creativity apps* and *education apps* on islands, their availability can benefit crisis routine. Among our study sample, we observe high demand for education apps among younger participants and participants with low technology affinity (median Rank 4, both) compared to older participants (−5.5 ranks) and participants with high technology affinity (−4 ranks), respectively. Further, we observed a tendency linking lower demand for education apps with lower education (Table 3).

Limitations of this Classification. The above typology omits some apps for different reasons. *Web browsers* stand out from all other apps from a technical perspective, considering that web browsers give users access to the World Wide Web with a virtually infinite number of servers [75]. Notably, our system model does not distinguish app-based and web-based Internet traffic. Island operation assumes that app servers on the network edge are available for users even when an island is disconnected from the outside Internet. The same holds for web servers, i.e., web pages with a web server hosted on the island will be available while web pages hosted on remote web servers will not. Website operators can change their backend communication infrastructure similar to app operators, enabling the use of their websites on islands. Therefore, the availability of apps and websites is somewhat equivalent on islands, making web browsers orthogonal to our classification. However, considering the enormous number of websites and the limited capacity of edge servers, island-ready websites will be feasible only for a few website operators and most interesting to local content providers such as local newspapers or communities. *Productivity apps* and *utilities* usually work offline and would be available to island users if downloaded before the crisis. The applicability of *travel* and *sports apps* with island connectivity is limited as no information from the outside Internet is available.

5.4 Qualitative Findings

App Prioritization. Three participants commented that not all app categories are equally important in crisis scenarios. One participant (P136) commented, "Several considered apps will never have a practical use nor will they be used during a crisis." Similarly, P902 found the available app categories "too many and in parts

absurd," asking, "Why should I read a book or listen to music during an earthquake?". This statement contrasts our observed need for distraction during crises, as 27% preferred *music apps* for distraction (RQ1). In addition, *music* stayed on Rank 14 while *books* gained three ranks from normal to island operation (RQ2). P475 demanded that "priorities should be set, since, e.g., games and apps for ordering food should be considered low-priority/useless," which indicates demand for the smartphone service typology we propose for island connectivity (RQ3).

Non-App Technologies. Thirteen participants (1.52%) used the "Other" option of Q12 to indicate they preferred telephony over apps for the use cases of emergency calls ($n = 7$), information ($n = 1$), communication ($n = 4$), and distraction ($n = 1$). This is complemented by the comments of four participants who "only use [their] smartphone for phone calls, SMSes, and utilities" (P148). Similarly, two participants criticize the overreliance on apps and smartphones, commenting that "one should never solely depend on smartphones" (P121) and regretting the shift away from traditional landline phones that were equipped to function during power outages (P742). These comments suggest prioritizing *telephony* higher than Tier 2 in the smartphone service typology (RQ3), potentially as a vital Tier 0 service. However, the small share of participants preferring telephony over general-purpose apps and the large share preferring general-purpose apps during crisis response motivate us to prioritize Tier 1 apps over telephony services.

Scenario Realism. Three participants commented that the earthquake scenario is implausible in Germany, to which one participant (P384) added that "flooding would be a better-suited example." This indicates that geographic factors should be considered when prioritizing apps for crisis scenarios. Six participants found the described island operation doubtful; e.g., one participant (P69) stated, "The described scenario is wrong because the Internet does not work locally." These comments suggest that some participants were inattentive while reading the survey or did not understand the study subject, as the main idea of island connectivity is maintaining local Internet connectivity, which the questionnaire explains elaborately.

6 DISCUSSION

This section positions our findings relative to prior work and discusses the practical implications for operators, developers, and authorities towards island-ready 6G communication.

6.1 Contextualizing Our Findings

Considering our German sample, we found demand for other apps than crisis apps for crisis-specific use cases (RQ1). While crisis apps were most demanded for warnings (U1), safety tips (U2), and emergency calls (U3), participants preferred messengers for communication (U5) and distraction (U7), news apps for looking up information (U4), and navigation apps for navigation (U6). These results support the findings of related work that dedicated crisis apps enjoy an increasing relevance for warnings, safety tips, and emergency calls and that crisis apps are perceived as more helpful sources for crisis communication than other channels. While in 2019, social media was still perceived as a quite or very helpful (50%)

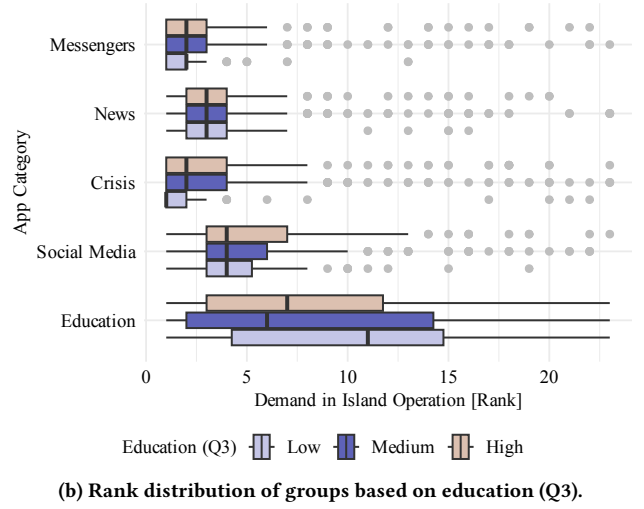
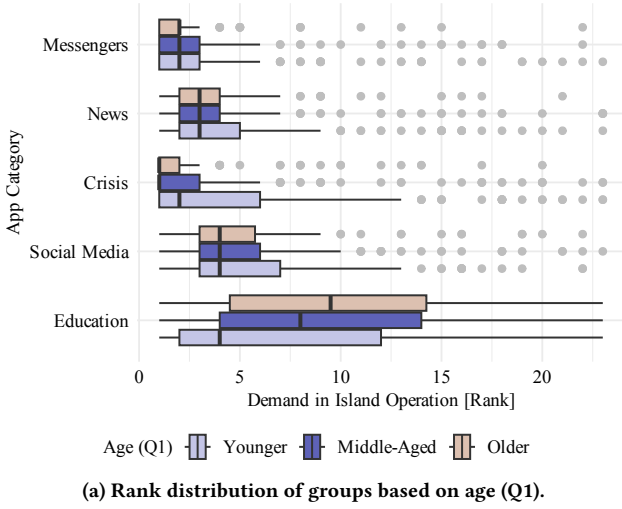


Figure 6: Rank Distribution Tendencies of Semantic Groups. This figure compares the app demand of semantic groups based on age (Figure 6a) and education (Figure 6b). Each figure features box plots to show the app demand in island operation (Q13) of a group (e.g., younger adults or participants with high education) for a selected app (e.g., messengers or crisis apps). Table 3 gives an overview of all considered variables.

information source [44], this perception decreased considerably until 2021 (only 28 %) in line with an increasing fear of false information disseminated through social media (54 % in 2019 vs. 76 % in 2021) and concerns about their reliability during emergencies (38 % in 2019 vs. 48 % in 2021) [93]. Furthermore, in both cited studies, messengers were the most used type of social media to search for and share information. This change in social media perception and use aligns with our findings that not social media but crisis, messengers and news apps were mentioned among the most important sources for warnings, safety tips and information. However, this work’s novel contribution compares use cases with general-purpose apps [109], highlighting the critical complementary roles of news and messenger apps.

In isolated areas with island connectivity (RQ2), German citizens report high demands for messengers, news apps, crisis apps, and social media apps. The demand for crisis apps was exceptionally high compared to normal operation, while the demand for shopping and travel apps dropped the most. Thus, being informed about the ongoing event alongside social exchange via messengers and social media constitutes essential smartphone-based activities in an island setting, emphasizing the need to provide them with decentralized local cellular connectivity [49, 106]. We assume that leisure activities such as entertainment, music, and gaming are still among important coping and distraction strategies, which, however, the widely available types of gaming, personal computers, and television devices [57] might fulfill. Overall, while prior user research examined the impact of infrastructure outages [35, 111], social media [94, 113], and warnings apps [36, 108] in crises, this work extends upon them by considering the concept of island connectivity.

We develop a smartphone service typology for island connectivity (RQ3) that combines smartphone apps and cellular services into tiers ordered by the criticality of their availability. Our typology

gives recommendations to MNOs, app operators, and authorities, e.g., to maintain emergency calls and cell broadcasts before ministering telephony and to prioritize the availability of messenger apps on islands over social media apps. These align with the resilience strategy of the German federal network agency that demands MNO to ensure essential services with the highest priority [12]. They also describe that MNOs should prioritize smartphone apps when the cellular network faces limited capacity, which aligns with applicable resilience definitions [51, 61]. We complement these general suggestions with a concrete typology that enables stakeholders to prioritize smartphone services.

6.2 Design Implications

We combined the crowd-sourced understanding gained from our survey with practice-informed knowledge of applicable standards and federal guidelines to derive a smartphone service typology. With the ongoing development of 6G, local connectivity on islands is still a vision, and many challenges remain to be solved by MNOs, smartphone app developers, and authorities. While the vision of island connectivity aligns well with the decentralized web [90] and local-first software [64], we emphasize that **the realization of islands and island connectivity poses non-trivial challenges for many stakeholders and requires a substantial review of current standards**. Proposing solutions to these challenges is outside the scope of this paper, but we outline design implications for selected involved parties in the following. With these, we support stakeholders in making user-centric design decisions during the realization of island connectivity as a part of 6G communication.

Towards Island-Ready 6G. The vital requirement to realize the vision of island connectivity in crisis-struck areas is the *distribution*

of the core network. Equipping every major city with a core network replica would improve resilience and maximize synchronization efforts and costs [21]. A trade-off could be to equip each metropolitan area with a core network replica or to co-locate each Internet gateway with a complete core network instance. The cellular network should be able to *transition into and out of island operation*. When an area is disconnected from the outside Internet, the core network replica should detect the isolation and enter island operation. Alternatively, Internet connections could follow a local-first approach [64] and access fallback to the Internet only when the requested content is unavailable on the local network edge. The *availability of critical smartphone services* should be prioritized depending on the crisis and the cellular network's capacity. Ideally, users could use all smartphone apps on islands, but MNOs might have to prioritize available services due to limited capacity. In that case, the smartphone service typology (RQ3) suggests making Tier 0 cellular services the main priority. At the same time, the availability of apps in Tiers 1 - 4 should be prioritized with descending priority, e.g., enabling citizen-to-citizen communication with messenger apps before enabling entertainment apps whose primary use case is distraction.

Towards Island-Ready Smartphone Apps. Our findings regarding RQ1 indicate that users would like to continue using general-purpose apps during crises, in addition to or instead of dedicated crisis apps. Therefore, apps should be available in island-ready 6G without requiring users to download a separate app version for island operation. Ideally, apps should be able to *transition into island operation* immediately when the area is disconnected from the outside Internet, which likely requires changes to the communication design of how apps connect to remote servers. Multi-access edge computing and local-first communication patterns [64] could play a central role in this endeavor, e.g., the app on a user's smartphone could connect to an app server on the local network edge, which (a) forwards traffic to a remote app server in normal operation or (b) enables local use of the app in island operation. For island operation, one specific challenge is ensuring security and privacy, as many apps rely on centralized trust infrastructure and identity providers, making it important to explore *local or decentralized authentication methods*. For messengers, which were most important to our participants overall, stronger emphasis needs to be placed on ensuring confidential end-to-end encrypted (E2EE) communication without requiring access to centralized authentication systems. Considering that the challenges of local connectivity will impact user interaction, *apps in island operation should be user-friendly*, i.e., how to announce the transition into island operation, how to establish trust locally, and how to explain to users the limitations of island connectivity for that app [42]. When the Internet connection is recovered, apps should be able to *transition out of island operation*, which likely requires the application of disruption-tolerant networking principles [38, 39, 71].

Towards Island-Ready Regulations. We envision that smartphone apps and 6G will be island-ready, supporting citizens during crisis response. Enabling island users to download smartphone apps

requires a *distributed crisis app store* on each island. To this end, app stores could be distributed similarly to app servers, providing a selection of apps and content suitable for potential crisis scenarios. Distributed app stores would further facilitate considering geographic factors when preparing a set of crisis-relevant apps for crisis scenarios. The responsibility for these app stores is an open problem, as is their content. On the one hand, the threat scenarios differ for each city, so the criticality of apps can also vary from city to city. This observation aligns with one participant's statement (P384), criticizing that "the scenario earthquake is implausible for Bremen. [...] A more suitable example would be flooding." For example, the Tier 3 app category weather apps should include tide forecasts for areas close to the sea, while areas in the mountains need rain and snow forecasts. On the other hand, managing app stores inflicts a lot of effort on municipalities, and leaving distributed crisis app stores under the authority of app operators or network providers might scale better.

Towards Island-Ready Users. Although our sample is limited to adult smartphone users from major German cities, previous research has identified different national risk cultures [23]. *State-oriented* risk cultures emphasize the state's responsibility for crisis prevention and management. *Individual-oriented* risk cultures describe that citizens assume their duties of being informed, prepared, and aware concerning risks. *Fatalistic* risk cultures are characterized by low confidence in their respective individual and state problem-solving potentials. While an appropriate level of both *citizen and state preparedness* and *response capabilities* is necessary, it seems essential to examine educational, political, and technological measures to prepare users and policymakers for the scenario of island connectivity. This shapes *transparency of smartphone service prioritization* to mitigate potential frustration upon the unavailability of functionality and *trust into critical infrastructure resilience*, including the expected outage durations [58].

6.3 Outlook and Future Work

In this paper, we have introduced the concept of island connectivity in crisis-struck areas to the HCI community. We acknowledge that this concept is visionary, and future work needs to solve many challenges before its realization.

Future HCI research. Our survey studied how adult smartphone users (over 18 and under 70) from major German cities prefer using smartphones in isolated cities. We decided to screen out other groups to align our study sample with our study subject. Nevertheless, insights from groups under-represented in our study can complement our findings very well, and we explicitly encourage future surveys that build upon our findings by studying other samples, e.g., participants from rural areas or from other countries. Upon acceptance of this paper, we will publish a replication package to facilitate such future studies. Our goal was to understand users' preferences about smartphone usage on islands with a focus on statistical and quantitative evaluation. Future qualitative studies can help refine our quantitative and statistical findings. Interviewing experts from app development, MNOs, and authorities can augment our user-centric findings with insights from other involved parties.

Surveys closer to real-world crisis scenarios, e.g., conducted in an affected area during crisis response, can support or contradict our findings. Connecting HCI research with one of the few available crisis-related datasets [5, 63] can create valuable insights.

Future 6G research. The design implications discussed in Section 6.2 give a first glance at the challenges left to be solved by future work. Most importantly, including distributed core networks in 6G is a paramount enabler of local connectivity in isolated areas. Future research can support app developers in making their smartphone apps island-ready by identifying suitable communication patterns for a swift transition into island operation.

7 CONCLUSION

The work reported in this paper makes two central contributions: Introducing the conceptual model of island connectivity and conducting the first survey ($N = 857$) of user preferences regarding smartphone usage on islands with local-only connectivity. With the upcoming standardization of 6G, it is crucial to understand the user perspective of island connectivity to support operators, developers, and authorities in making user-centric design decisions while realizing island connectivity.

We describe how 6G can facilitate island connectivity for crisis-struck areas and introduce the concept of island connectivity. We apply statistical and quantitative methods to learn that users prefer apps other than crisis apps for some crisis-specific use cases, e.g., messengers for communicating with other citizens and news apps to look up information (RQ1). App demands differ greatly in isolated areas compared to everyday life (RQ2). While news and crisis apps surge in demand, shopping and travel apps plummet. These findings are based on our survey whose design we describe in detail, proactively considering biases and limitations. We combine our survey's insights with applicable standards to develop a smartphone service typology (RQ3), organizing apps in tiers with descending criticality for isolated cities, distinguishing between apps crucial for crisis response and apps enabling crisis routine. From this classification, we derive implications for mobile network operators, app developers, and authorities, facilitating user-centric design decisions in the realization process of island-ready 6G.

ACKNOWLEDGMENTS

This work has been co-funded by the Federal Ministry of Education and Research of Germany in the projects Open6GHub (grant number: 16KISK014) and CYLENCE (grant number: 13N16636), and by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72]. We thank the participants of our pilot study and the anonymous reviewers for their helpful suggestions, which helped improve this work. We thank Lea Holaus for the comics (Figures 1 and 3), Alexander Heinrich for supporting the qualitative data analysis, and Sebastian Perrig and Nicholas Scharowski for the helpful discussions on survey design.

AVAILABILITY

We provide artifacts to verify, reproduce, and replicate our findings: With this paper, we publish our survey dataset and evaluation scripts [55] to verify and reproduce our results and the LimeSurvey survey structure to replicate our study.

REFERENCES

- [1] 3GPP. 2023. *System architecture for the 5G System (5GS), Technical Specification (TS) 23.501, version 18.4.0*. Technical Specification. 3rd Generation Partnership Project (3GPP). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>
- [2] Hervé Abdi. 2007. Bonferroni and Šidák corrections for multiple comparisons. *Encyclopedia of measurement and statistics* 3, 01 (2007), 2007.
- [3] Firoj Alam, Ferda Ofli, and Muhammad Imran. 2020. Descriptive and visual summaries of disaster events using artificial intelligence techniques: case studies of Hurricanes Harvey, Irma, and Maria. *Behaviour & Information Technology* 39, 3 (2020), 288–318. Publisher: Taylor & Francis.
- [4] Hadi Alizadeh, Ayyoob Sharifi, Safiyeh Damanbagh, Hadi Nazarnia, and Mohammad Nazarnia. 2023. Impacts of the COVID-19 pandemic on the social sphere and lessons for crisis management: a literature review. *Natural Hazards* 117, 3 (2023), 2139–2164.
- [5] Flor Álvarez, Lars Almon, Patrick Lieser, Tobias Meuser, Yannick Dylla, Björn Richerzhagen, Matthias Hollick, and Ralf Steinmetz. 2018. Conducting a large-scale field test of a smartphone-based communication network for emergency response. In *Proceedings of the 13th Workshop on Challenged Networks*. ACM, New Delhi, India, 3–10.
- [6] Scott Anderson, Loqman Salamatian, Zachary S Bischof, Alberto Dainotti, and Paul Barford. 2022. iGDB: connecting the physical and logical layers of the internet. In *Proceedings of the 22nd ACM Internet Measurement Conference*. ACM, Madrid, Spain, 433–448.
- [7] Apple Inc. 2024. Categories and Discoverability - App Store. <https://developer.apple.com/app-store/categories/> [Accessed 2024-09-12].
- [8] Bahareh Bahrami, Mohammad Reza Khayyambashi, and Seyedali Mirjalili. 2023. Edge server placement problem in multi-access edge computing environment: models, techniques, and applications. *Cluster Computing* 26, 5 (2023), 3237–3262. Publisher: Springer.
- [9] BBK. 2024. Informationstechnik und Telekommunikation. https://www.bbk.bund.de/DE/Themen/Kritische-Infrastrukturen/Sektoren-Branchen/Informationstechnik-Telekommunikation/informationstechnik-telekommunikation_node.html [Accessed 2024-09-12].
- [10] Ali Behravan, Vijaya Yajnanarayana, Musa Furkan Keskin, Hui Chen, Deep Shrestha, Traian E. Abruđan, Tommy Svensson, Kim Schindhelm, Andreas Wolfgang, Simon Lindberg, and Henk Wymeersch. 2023. Positioning and Sensing in 6G: Gaps, Challenges, and Opportunities. *IEEE Vehicular Technology Magazine* 18, 1 (2023), 40–48.
- [11] Jelke Bethlehem. 2010. Selection bias in web surveys. *International statistical review* 78, 2 (2010), 161–188. Publisher: Wiley Online Library.
- [12] BNA. 2022. *Resilience of Telecommunications Networks*. Technical Report. Bundesnetzagentur (BNA). https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Strategiepapier_Resilienz_eng.pdf?__blob=publicationFile&v=2
- [13] BNA. 2024. Informationsplattform Schienengüterverkehr (ISA). <https://isa.bundesnetzagentur.de/home/#/> [Accessed 2024-09-12].
- [14] Robert Braden. 1989. *Rfc1122: Requirements for internet hosts-communication layers*. Technical Report. RFC Editor.
- [15] Florian Brühlmann, Serge Petralito, Lena F. Aeschbach, and Klaus Opwis. 2020. The quality of data collected online: An investigation of careless responding in a crowdsourced sample. *Methods in Psychology* 2 (2020), 100022. <https://doi.org/10.1016/j.metip.2020.100022>
- [16] BSI. 2023. *ZwiBACK-Studie: Zwischenfälle in Bezug auf Bedenken und Achtsamkeit bei der Cybersicherheit*. Technical Report. Bundesamt für Sicherheit in der Informationstechnik (BSI). https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Studien/ZwiBACK/ZwiBACK-Studie.pdf?__blob=publicationFile&v=3
- [17] Patrice Calégar, Frédéric Guidec, Pierre Kuonen, and Daniel Kobler. 1997. Parallel island-based genetic algorithm for radio network design. *J. Parallel and Distrib. Comput.* 47, 1 (1997), 86–90.
- [18] Carlos Castillo. 2016. *Big crisis data: social media in disasters and time-critical situations*. Cambridge University Press, New York, US.
- [19] Songyan Chen, Junjie Chen, and Hongjun Li. 2024. Joint optimization of UPF placement and traffic routing for 5G core network user plane. *Computer Communications* 216 (2024), 86–94. <https://doi.org/10.1016/j.comcom.2023.12.029>
- [20] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates, New York, US.
- [21] Marius Corici, Hemant Zope, Hauke Buhr, Christian Scheich, and Thomas Magedanz. 2024. Shortcuts: A Simple Mechanism for Reducing the Data Path Delay in Beyond 5G and 6G Networks. In *27th Conference on Innovation in Clouds, Internet and Networks (ICIN)*. IEEE, Paris, France, 47–51.
- [22] Marius-Julian Corici, Fabian Eichhorn, Roland Bless, Michael Gundall, Daniel Lindenschmitt, Bastian Bloessl, Marina Petrova, Lara Wimmer, Ronny Kreuch, Thomas Magedanz, et al. 2023. Organic 6G networks: Vision, requirements, and research approaches. *IEEE Access* 11 (2023), 70698–70715.

- [23] Alessio Cornia, Kerstin Dressel, and Patricia Pfeil. 2016. Risk cultures and dominant approaches towards disasters in seven European countries. *Journal of Risk Research* 19, 3 (2016), 288–304. Publisher: Taylor & Francis.
- [24] Cybersecurity and Infrastructure Security Agency. 2024. Communications Sector. <https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors/communications-sector> [Accessed 2024-09-12].
- [25] Cybersecurity and Infrastructure Security Agency. 2024. Information Technology Sector. <https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors/information-technology-sector> [Accessed 2024-09-12].
- [26] Irina Dallo and Michèle Marti. 2021. Why should I use a multi-hazard app? Assessing the public's information needs and app feature preferences in a participatory process. *International journal of disaster risk reduction* 57 (2021), 102197. Publisher: Elsevier.
- [27] GC Deepak, Alexandros Ladas, Yusuf Abdulrahman Sambo, Haris Pervaiz, Christos Politis, and Muhammad Ali Imran. 2019. An overview of post-disaster emergency communication systems in the future networks. *IEEE Wireless Communications* 26, 6 (2019), 132–139. Publisher: IEEE.
- [28] Destatis. 2024. Ausstattung privater Haushalte mit Informations- und Kommunikationstechnik - Deutschland. <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/Tabellen/liste-infotechnik-d.html#115502> [Accessed 2024-11-07].
- [29] Destatis. 2024. Großstadregionen im Wandel. <https://www.destatis.de/DE/Themen/Querschnitt/Demografischer-Wandel/Aspekte/demografie-grossstadregionen.html> [Accessed 2024-09-12].
- [30] Don A Dillman, Jolene D Smyth, and Leah Melani Christian. 2014. *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons, Hoboken, US.
- [31] Ramakrishnan Durairajan, Paul Barford, Joel Sommers, and Walter Willinger. 2015. InterTubes: A study of the US long-haul fiber-optic infrastructure. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*. ACM, London, United Kingdom, 565–578.
- [32] Frances L. Edwards. 2022. Crisis, Emergency, Disaster, and Catastrophe Defined. In *Global Encyclopedia of Public Administration, Public Policy, and Governance*. Ali Farazmand (Ed.). Springer International Publishing, Cham, 2793–2797. https://doi.org/10.1007/978-3-030-66252-3_2878
- [33] Kathrin Eismann, Oliver Posegga, and Kai Fischbach. 2016. Collective behaviour, Social Media, and disasters: a Systematic literature Review.. In *ECIS*. Association for Information Systems, Istanbul, Turkey, 1–20.
- [34] Dirk M. Elston. 2021. Survivorship bias. *Journal of the American Academy of Dermatology* 10 (2021), 1–2. <https://doi.org/10.1016/j.jaad.2021.06.845>
- [35] Alexander Fekete and Simone Sandholz. 2021. Here comes the flood, but not failure? Lessons to learn after the heavy rain and pluvial floods in Germany 2021. *Water* 13, 21 (2021), 3016. Publisher: MDPI.
- [36] Diana Fischer-Prefßler, Dario Bonaretti, and Kai Fischbach. 2022. A protection-motivation perspective to explain intention to use and continue to use mobile warning systems. *Business & Information Systems Engineering* 64, 2 (2022), 167–182. Publisher: Springer.
- [37] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human-Computer Interaction* 35, 6 (2019), 456–467. Publisher: Taylor & Francis.
- [38] Wei Gao and Guohong Cao. 2011. User-centric data dissemination in disruption tolerant networks. In *2011 Proceedings IEEE INFOCOM*. IEEE, Shanghai, China, 3119–3127.
- [39] Wei Gao, Guohong Cao, Arun Iyengar, and Mudhakar Srivatsa. 2011. Supporting cooperative caching in disruption tolerant networks. In *2011 31st International Conference on Distributed Computing Systems*. IEEE, Minneapolis, USA, 151–161.
- [40] GapFish GmbH. 2024. GapFish – Knowing what really matters. <https://gapfish.com> [Accessed 2024-09-12].
- [41] Robert Greszki, Marco Meyer, and Harald Schoen. 2015. Exploring the effects of removing “too fast” responses and respondents from web surveys. *Public Opinion Quarterly* 79, 2 (2015), 471–503. Publisher: Oxford University Press US.
- [42] Steffen Haesler, Ragnark Mogk, Florentin Putz, Kevin T. Logan, Nadja Thiessen, Katharina Kleinschnitger, Lars Baumgärtner, Jan-Philipp Stroscher, Christian Reuter, Michele Knodt, and Matthias Hollick. 2021. Connected Self-Organized Citizens in Crises: An Interdisciplinary Resilience Concept for Neighborhoods. In *CSCW '21 Companion: Conference Companion Publication of the 2021 on Computer Supported Cooperative Work and Social Computing*. ACM, Virtual Event, USA, 62–66. <https://doi.org/10.1145/3462204.3481749>
- [43] Yunita Astriani Hardayati and Mustikasari Mustikasari. 2019. The implementation of relaxation and distraction techniques on adolescents experiencing anxiety in earthquake prone areas. *International Journal of Nursing and Health Services (IJNHS)* 2, 3 (2019), 9–15.
- [44] Jasmin Haunschild, Marc-André Kaufhold, and Christian Reuter. 2020. Sticking with Landlines? Citizens' and Police Social Media Use and Expectation During Emergencies. In *Proceedings of the International Conference on Wirtschaftsinformatik (WI) (Best Paper Social Impact Award)*. AIS Electronic Library (AISel), Potsdam, Germany, 1–16. https://doi.org/10.30844/wi_2020_o2-haunschild
- [45] Jasmin Haunschild, Marc-André Kaufhold, and Christian Reuter. 2022. Perceptions and Use of Warning Apps – Did Recent Crises Lead to Changes in Germany?. In *Mensch und Computer - Tagungsband*. ACM, New York, 25–40. <https://doi.org/10.1145/3543758.3543770>
- [46] Jasmin Haunschild, Selina Pauli, and Christian Reuter. 2023. Preparedness Nudging for Warning Apps? A Mixed-Method Study Investigating Popularity and Effects of Preparedness Alerts in Warning Apps. *International Journal on Human-Computer Studies (IJHCS)* 172 (2023), 102995. <https://doi.org/10.1016/j.ijhcs.2023.102995>
- [47] Andrin Hauri, Kevin Kohler, and Benjamin Scharte. 2022. *A comparative assessment of mobile device-based multi-hazard warnings: Saving lives through public alerts in europe*. Technical Report. ETH Zurich.
- [48] Douglas D. Heckathorn. 2011. Comment: Snowball versus Respondent-Driven Sampling. *Sociological Methodology* 41, 1 (2011), 355–366. <https://doi.org/10.1111/j.1467-9531.2011.01244.x> _eprint: <https://doi.org/10.1111/j.1467-9531.2011.01244.x>
- [49] Michael Heise, Martin Pietsch, Florian Steinke, Maximilian Bauer, and Burak Yilmaz. 2022. Optimized UAV Placement for Resilient Crisis Communication and Power Grid Restoration. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, Novi Sad, Serbia, 1–5. <https://doi.org/10.1109/ISGT-Europe54678.2022.9960494>
- [50] Joseph Heinrich. 2015. Culture and social behavior. *Current opinion in behavioral sciences* 3 (2015), 84–89. Publisher: Elsevier.
- [51] Matthias Hollick, Anne Hofmeister, Jens Ivo Engels, Bernd Freisleben, Gerrit Hornung, Anja Klein, Michèle Knodt, Inke Lorenz, Max Mühlhäuser, Peter Pelz, and others. 2019. The emergency responsive digital city. In *2019 World Congress on Resilience, Reliability and Asset Management (WCRAM)*. <http://tubiblio.ulb-tu-darmstadt.de/121649>. WCRAM, Singapore, 136–143.
- [52] Muzammil Hussain, Ahmed Al-Haiqi, AA Zaidan, BB Zaidan, Miss Laiha Mat Kiah, Nor Badrul Anuar, and Mohamed Abdulnabi. 2015. The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations. *Computer methods and programs in biomedicine* 122, 3 (2015), 393–408.
- [53] Jonas Höchst, Lars Baumgärtner, Franz Kuntke, Alvar Penning, Artur Sterz, Markus Sommer, and Bernd Freisleben. 2023. Mobile Device-to-Device Communication for Crisis Scenarios Using Low-Cost LoRa Modems. In *Disaster Management and Information Technology: Professional Response and Recovery Management in the Age of Disasters*. Hans Jochen Scholl, Eric E. Holdeman, and F. Kees Boersma (Eds.). Springer International Publishing, Cham, 235–268. https://doi.org/10.1007/978-3-031-20939-0_12
- [54] Muhammad Imran, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. 2015. Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)* 47, 4 (2015), 1–38. Publisher: ACM New York, NY, USA.
- [55] Leon Janzen, Florentin Putz, Kolja Straub, Marc-André Kaufhold, and Matthias Hollick. 2025. *Survey Dataset: The User Perspective on Island-Ready 6G Communication*. <https://doi.org/10.5281/zenodo.14812894>
- [56] Marc-André Kaufhold. 2024. Exploring the evolving landscape of human-centred crisis informatics: current challenges and future trends. *i-com - Journal of Interactive Media* 2 (2024), 155–163. <https://doi.org/10.1515/icom-2024-0002>
- [57] Marc-André Kaufhold, Julian Bäumler, and Christian Reuter. 2022. The Implementation of Protective Measures and Communication of Cybersecurity Alerts in Germany - A Representative Survey of the Population. In *Mensch und Computer 2022 - Workshopband (Mensch und Computer 2022 - Workshopband)*. Gesellschaft für Informatik, Darmstadt, 1–12. <https://doi.org/10.18420/muc2022-mci-ws01-228>
- [58] Marc-André Kaufhold, Margarita Grinko, Christian Reuter, Marén Schorch, Amanda Langer, Sascha Skudelný, and Matthias Hollick. 2019. Potentiale von IKT beim Ausfall kritischer Infrastrukturen: Erwartungen, Informationsgewinnung und Mediennutzung der Zivilbevölkerung in Deutschland. In *Proceedings of the International Conference on Wirtschaftsinformatik (WI)*. AIS, Siegen, Germany, 1054–1068. <https://aisel.aisnet.org/wi2019/track09/papers/4/>
- [59] Marc-André Kaufhold, Thea Riebe, Markus Bayer, and Christian Reuter. 2024. 'We Do Not Have the Capacity to Monitor All Media': A Design Case Study on Cyber Situational Awareness in Computer Emergency Response Teams. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI) (Best Paper Award) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3613904.3642368>
- [60] Marc-André Kaufhold, Nicola Rupp, Christian Reuter, and Matthias Habdank. 2020. Mitigating Information Overload in Social Media during Conflicts and Crises: Design and Evaluation of a Cross-Platform Alerting System. *Behaviour & Information Technology (BIT)* 39, 3 (2020), 319–342. <https://doi.org/10.1080/144929X.2019.1620334>
- [61] Ladan Khalooupour, Yanpeng Su, Florian Raskob, Tobias Meuser, Roland Bless, Leon Wirsching, Kamyar Abedi, Marko Andjelkovic, Hekma Chaari, Pousali Chakraborty, Michael Kreutzer, Matthias Hollick, Thorsten Strufe, Norman

- Franchi, and Vahid Jamali. 2024. Resilience-by-Design Concepts for 6G Communication Networks. <https://arxiv.org/abs/2405.17480> eprint: 2405.17480.
- [62] Wolfgang Kiess and Ashiq Khan. 2014. Centralized vs. distributed: On the placement of gateway functionality in 5G cellular networks. In *2014 IEEE Global Communications Conference*. IEEE, Texas, United States of America, 4788–4793.
- [63] Jooho Kim and Makarand Hastak. 2018. Social network analysis: Characteristics of online social networks after a disaster. *International journal of information management* 38, 1 (2018), 86–96.
- [64] Martin Kleppmann, Adam Wiggins, Peter Van Hardenberg, and Mark McGranaghan. 2019. Local-first software: you own your data, in spite of the cloud. In *Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software*. ACM, New York, US, 154–178.
- [65] Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology*. Sage Publications, Inc., London, UK.
- [66] Ivar Krumpal. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & quantity* 47, 4 (2013), 2025–2047. Publisher: Springer.
- [67] Franz Kuntke, Lars Baumgärtner, and Christian Reuter. 2023. Rural Communication in Outage Scenarios: Disruption-Tolerant Networking via LoRaWAN Setups. In *Proceedings of Information Systems for Crisis Response and Management (ISCRAM)*. Department of Information Systems & Quantitative Analysis, Omaha, US, Omaha, US, 1–13. https://idl.iscram.org/files/kuntke/2023/2581_Kuntke_etal2023.pdf
- [68] Franz Kuntke, Merve Bektas, Laura Buhleier, Ella Pohl, Rebekka Schiller, and Christian Reuter. 2023. How Would Emergency Communication Based on LoRaWAN Perform? Empirical Findings of Signal Propagation in Rural Areas. In *Proceedings of Information Systems for Crisis Response and Management (ISCRAM)*. Department of Information Systems & Quantitative Analysis, Omaha, US, Omaha, US, 1–8. https://idl.iscram.org/files/kuntke/2023/2586_Kuntke_etal2023.pdf
- [69] Jeremy Lakeman, Matthew Lloyd, Romana Challans, Angus Wallace, Paul Gardner-Stephen, Milan Stute, and Matthias Hollick. 2017. A practical and secure social media facility for internet-deprived populations. In *2017 IEEE Global Humanitarian Technology Conference (GHTC)*. IEEE, San Jose, USA, 1–9.
- [70] Glyn Lewis and Simon Wessely. 1992. The epidemiology of fatigue: more questions than answers. *Journal of epidemiology and community health* 46, 2 (1992), 92. Publisher: BMJ Publishing Group.
- [71] Feng Li and Jie Wu. 2009. MOPS: Providing content-based service in disruption-tolerant networks. In *2009 29th IEEE international conference on distributed computing systems*. IEEE, Montreal, Canada, 526–533.
- [72] Yuanzhe Li, Xiao Ma, Mengwei Xu, Ao Zhou, Qibo Sun, Ning Zhang, and Shangguang Wang. 2021. Joint placement of UPF and edge server for 6G network. *IEEE Internet of Things Journal* 8, 22 (2021), 16370–16378. Publisher: IEEE.
- [73] LimeSurvey. 2024. LimeSurvey - Free Online Survey Tool. <https://www.limesurvey.org/de> [Accessed 2024-09-12].
- [74] MASTD. 2021. Mobilfunk und Breitbandversorgung im Ahrtal weiter im Fokus. <https://mastd.rlp.de/service/presse/detail/mobilfunk-und-breitbandversorgung-im-ahrtal-weiter-im-fokus> [Accessed 2024-09-12].
- [75] James McCauley, Scott Shenker, and George Varghese. 2023. Extracting the essential simplicity of the Internet. *Commun. ACM* 66, 2 (2023), 64–74.
- [76] Iain McLean, Arnold B. Urken, and Fiona Hewitt. 1995. *Classics of Social Choice*. University of Michigan Press, Michigan, US.
- [77] Adam W Meade and S Bartholomew Craig. 2012. Identifying careless responses in survey data. *Psychological methods* 17, 3 (2012), 437. Publisher: American Psychological Association.
- [78] François Michel, Martino Trevisan, Danilo Giordano, and Olivier Bonaventure. 2022. A first look at starlink performance. In *Proceedings of the 22nd ACM Internet Measurement Conference*. ACM, Nice, France, 130–136.
- [79] Thea Palsgaard Møller, Annette Kjør Ersbøll, Janne Schurmann Tolstrup, Doris Østergaard, Søren Viereck, Jerry Overton, Fredrik Folke, and Freddy Lippert. 2015. Why and when citizens call for emergency help: an observational study of 211,193 medical emergency calls. *Scandinavian journal of trauma, resuscitation and emergency medicine* 23 (2015), 1–10.
- [80] Neema Moraveji, Jason Li, Jiarong Ding, Patrick O’Kelley, and Suze Woolf. 2007. Comicboards: using comics as proxies for participatory design with children. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, San Jose, US, 1371–1374.
- [81] Simon Moss. 2008. Acquiescence bias. <http://www.psych-it.com.au/Psychlopedia/article.asp> [Accessed 2024-09-12].
- [82] Shreyasee Mukherjee, Ravishankar Ravindran, and Dipankar Raychaudhuri. 2018. A distributed core network architecture for 5G systems and beyond. In *Proceedings of the 2018 Workshop on Networking for Emerging Applications and Technologies*. ACM, Budapest, Hungary, 33–38.
- [83] Alexandra Olteanu, Sarah Vieweg, and Carlos Castillo. 2015. What to expect when the unexpected happens: Social media communications across crises. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. ACM, Vancouver, Canada, 994–1009.
- [84] Francesco Palmieri, Ugo Fiore, Aniello Castiglione, Fang-Yie Leu, and Alfredo De Santis. 2013. Analyzing the internet stability in presence of disasters. In *Security Engineering and Intelligence Informatics: CD-ARES 2013 Workshops: MoCrySEn and SeCIHD, Regensburg, Germany, September 2-6, 2013. Proceedings* 8. Springer, Regensburg, Germany, 253–268.
- [85] Guy Paré, Marie-Claude Trudel, Mirou Jaana, and Spyros Kitsiou. 2015. Synthesizing information systems knowledge: A typology of literature reviews. *Information & management* 52, 2 (2015), 183–199.
- [86] Mukaddim Pathan, Rajkumar Buyya, and Athena Vakali. 2008. Content delivery networks: State of the art, insights, and imperatives. *Content Delivery Networks* 9 (2008), 3–32.
- [87] Karl Pearson. 1992. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. In *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. Vol. 50. Springer, New York, USA, 11–28.
- [88] Matteo Pozza, Ashwin Rao, Hannu Flinck, and Sasu Tarkoma. 2018. Network-in-a-box: A survey about on-demand flexible networks. *IEEE Communications Surveys & Tutorials* 20, 3 (2018), 2407–2428.
- [89] R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> [Accessed 2024-09-12].
- [90] Aravindh Raman, Sagar Joglekar, Emiliano De Cristofaro, Nishanth Sastry, and Gareth Tyson. 2019. Challenges in the decentralised web: The mastodon case. In *Proceedings of the internet measurement conference*. ACM, Amsterdam, Netherlands, 217–229.
- [91] Christian Reuter, Oliver Heger, and Volkmar Pipek. 2013. Combining real and virtual volunteers through social media. In *Proceedings of the 10th International ISCRAM Conference*. KIT, Karlsruhe, Germany, Baden-Baden, Germany, 780–790.
- [92] Christian Reuter, Amanda Lee Hughes, and Marc-André Kaufhold. 2018. Social Media in Crisis Management: An Evaluation and Analysis of Crisis Informatics Research. *International Journal on Human-Computer Interaction (IJHCI)* 34, 4 (2018), 280–294. <https://doi.org/10.1080/10447318.2018.1427832>
- [93] Christian Reuter, Marc-André Kaufhold, Tom Biselli, and Helene Pleil. 2023. Increasing adoption despite perceived limitations of social media in emergencies: Representative insights on German citizens’ perception and trends from 2017 to 2021. *International Journal of Disaster Risk Reduction* 96 (2023), 103880. <https://doi.org/10.1016/j.ijdrr.2023.103880>
- [94] Christian Reuter, Marc-André Kaufhold, Stefka Schmid, Thomas Spielhofer, and Anna Sophie Hahne. 2019. The Impact of Risk Cultures: Citizens’ Perception of Social Media Use in Emergencies across Europe. *Technological Forecasting and Social Change (TFSC)* 148, 119724 (2019), 1–17. <https://doi.org/10.1016/j.techfore.2019.119724>
- [95] Christian Reuter, Marc-André Kaufhold, Thomas Spielhofer, and Anna Sophie Hahne. 2017. Social Media in Emergencies: A Representative Study on Citizens’ Perception in Germany. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 1–19. <https://doi.org/10.1145/3134725> Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [96] Matei Ripeanu. 2001. Peer-to-peer architecture case study: Gnutella network. In *Proceedings first international conference on peer-to-peer computing*. IEEE, Linköping, Sweden, 99–100.
- [97] Justus Rischke, Peter Sossalla, Sebastian Itting, Frank HP Fitzek, and Martin Reisslein. 2021. 5G campus networks: A first measurement study. *IEEE Access* 9 (2021), 121786–121803.
- [98] John Paul Rula. 2017. *Adopting a Gateway Centric View for Cellular Network Content Delivery*. PhD Thesis. Northwestern University.
- [99] Donald G Saari. 2000. Mathematical structure of voting paradoxes: II. Positional voting. *Economic Theory* 15 (2000), 55–102. Publisher: Springer.
- [100] Howard Schuman and Stanley Presser. 1996. *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Sage, London, UK.
- [101] Lucy Simko, Harshini Sri Ramulu, Tadayoshi Kohno, and Yasemin Acar. 2023. The Use and Non-Use of Technology During Hurricanes. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2 (Oct. 2023), 1–54. <https://doi.org/10.1145/3610215> Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [102] Rishabh Singla, Shreyas Srinivasa, Narasimha Reddy, Jens Myrup Pedersen, Emmanouil Vasilomanolakis, and Riccardo Bettati. 2023. An Analysis of War Impact on Ukrainian Critical Infrastructure Through Network Measurements. In *2023 7th Network Traffic Measurement and Analysis Conference (TMA)*. IEEE, Napoli, Italy, 1–10. <https://doi.org/10.23919/TMA58422.2023.10199005>
- [103] Robert Soden and Leysia Palen. 2018. Informing crisis: Expanding critical perspectives in crisis informatics. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–22. Publisher: ACM New York, NY, USA.
- [104] Kate Starbird and Leysia Palen. 2011. “Voluntweeters” self-organizing by digital volunteers in times of crisis. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, Vancouver, Canada, 1071–1080.

- [105] Statista. 2024. Number of smartphone users worldwide 2016-2028. On-line. <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> [Accessed 2024-11-07].
- [106] Artur Sterz, Robin Klose, Markus Sommer, Jonas Höchst, Jakob Link, Bernd Simon, Anja Klein, Matthias Hollick, and Bernd Freisleben. 2023. Energy-Efficient Decentralized Broadcasting in Wireless Multi-Hop Networks. *Sensors* 23, 17 (2023), 7419. Publisher: MDPI.
- [107] Milan Stute, Max Maass, Tom Schons, and Matthias Hollick. 2017. Reverse Engineering Human Mobility in Large-scale Natural Disasters. In *Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems* (Miami, Florida, USA) (MSWiM '17). Association for Computing Machinery, New York, NY, USA, 219–226. <https://doi.org/10.1145/3127540.3127542>
- [108] Marion Lara Tan, Raj Prasanna, Kristin Stock, Emma EH Doyle, Graham Leonard, and David Johnston. 2020. Usability factors influencing the continuance intention of disaster apps: A mixed-methods study. *International Journal of Disaster Risk Reduction* 50 (2020), 101874. Publisher: Elsevier.
- [109] Marion Lara Tan, Raj Prasanna, Kristin Stock, Emma Hudson-Doyle, Graham Leonard, and David Johnston. 2017. Mobile applications in crisis informatics literature: A systematic review. *International journal of disaster risk reduction* 24 (2017), 297–311. Publisher: Elsevier.
- [110] George R Terrell and David W Scott. 1992. Variable kernel density estimation. *The Annals of Statistics* 20, 3 (1992), 1236–1265.
- [111] Annegret H Thieken, Philip Bubeck, Anna Heidenreich, Jennifer Von Keyserlingk, Lisa Dillenardt, and Antje Otto. 2023. Performance of the flood warning system in Germany in July 2021—insights from affected residents. *Natural hazards and earth system sciences* 23, 2 (2023), 973–990. Publisher: Copernicus GmbH.
- [112] Amos Tversky and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science* 185, 4157 (1974), 1124–1131.
- [113] Houcai Wang, Li Xiong, Chengwen Wang, and Nan Chen. 2022. Understanding Chinese mobile social media users' communication behaviors during public health emergencies. *Journal of Risk Research* 25, 7 (2022), 874–891. Publisher: Taylor & Francis.
- [114] Valentin Weber, Maria Pericàs Riera, and Emma Laumann. 2023. *Mapping the World's Critical Infrastructure Sectors*. Technical Report. Forschungsinstitut der Deutschen Gesellschaft für Auswärtige Politik e.V. 9 pages. <https://www.ssoar.info/ssoar/handle/document/91328> Publisher: DEU.
- [115] Leon Würsching and Matthias Hollick. 2023. Maintaining App Services in Disrupted Cities: A Crisis and Resilience Evaluation Tool. In *NetMob '23 Book of Abstracts*. IMDEA, Madrid, Spain, Madrid, Spain, 141–142. <http://tubiblio.ulb.tu-darmstadt.de/140700/> event-place: Madrid, Spain.
- [116] Kerstin K Zander, Duy Nguyen, Milad Mirbabaie, and Stephen T Garnett. 2023. Aware but not prepared: understanding situational awareness during the century flood in Germany in 2021. *International journal of disaster risk reduction* 96 (2023), 103936. Publisher: Elsevier.
- [117] Kateryna Zarembo, Michèle Knodt, and Jannis Kachel. 2024. Smartphone resilience: ICT in Ukrainian civic response to the Russian full-scale invasion. *Media, War & Conflict* 0, 0 (2024), 17506352241236449. Publisher: SAGE Publications Sage UK: London, England.
- [118] Kyriakos Zarifis, Tobias Flach, Srikanth Nori, David Choffnes, Ramesh Govindan, Ethan Katz-Bassett, Z Morley Mao, and Matt Welsh. 2014. Diagnosing path inflation of mobile client traffic. In *Passive and Active Measurement: 15th International Conference, PAM 2014, Los Angeles, CA, USA, March 10-11, 2014, Proceedings 15*. Springer, Los Angeles, US, 23–33.
- [119] Haijun Zhang, Chunxiao Jiang, Rose Qingyang Hu, and Yi Qian. 2016. Self-organization in disaster-resilient heterogeneous small cell networks. *IEEE Network* 30, 2 (2016), 116–121. <https://doi.org/10.1109/MNET.2016.7437033>
- [120] Zesen Zhang, Alexander Marder, Ricky Mok, Bradley Huffaker, Matthew Luckie, Kimberly C Claffy, and Aaron Schulman. 2021. Inferring regional access network topologies: Methods and applications. In *Proceedings of the 21st ACM Internet Measurement Conference*. ACM, Virtual Event, USA, 720–738.
- [121] Shiyao Zhu, Dezhi Li, Haibo Feng, Tiantian Gu, Kasun Hewage, and Rehan Sadiq. 2020. Smart city and resilient city: Differences and connections. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10, 6 (2020), e1388. Publisher: Wiley Online Library.
- [122] Naomi Ziv and Revital Hollander-Shabtai. 2022. Music and COVID-19: Changes in uses and emotional reaction to music under stay-at-home restrictions. *Psychology of Music* 50, 2 (2022), 475–491.

A QUESTIONNAIRE

We translated the transcript of the questionnaire from German to English. We did not provide the headlines and part descriptions during the survey.

A.1 Part 1: Consent and Screening

A.1.1 **Age (Q1)** . "How old are you?" ("Under 18", "18-29", "30-39", "40-49", "50-59", "60-69", "70+")

A.1.2 **Gender (Q2)** . "Which gender do you identify with?" ("Female", "Male", "Non-Binary", "Own description" with a text field, "Prefer not to disclose")

A.1.3 **Education (Q3)** . "What is your highest educational qualification?" ("None (yet)", "Hauptschulabschluss", "Polytechnische Oberschule", "Mittlere Reife, Realschulabschluss", "(Fach-)Hochschulreife", "Bachelor", "Master", "Diplom", "Promotion", "Other qualification" with a textfield)

A.1.4 **Income (Q4)** . "If you add up all the incomes of your household, Which of the following income groups does your monthly net household income fall into (i.e., household income after deducting taxes)?" ("under 1000 €", "1000 € to 1999 €", "2000 € to 2999 €", "3000 € to 3999 €", "4000 € to 4999 €", "5000 € and more", "No answer")

A.1.5 **Federal State (Q5)** . "In which federal state do you currently live?" ("Baden-Württemberg", "Bavaria", "Berlin", "Brandenburg", "Bremen", "Hamburg", "Hesse", "Mecklenburg-Western Pomerania", "Lower Saxony", "North Rhine-Westphalia", "Rhineland-Palatinate", "Saarland", "Saxony", "Saxony-Anhalt", "Schleswig-Holstein", "Thuringia")

A.1.6 **Citizen and Smartphone Ownership (Q6)** . "Have you lived in a major German city (population at least 100.000) for at least 1 year since 2008?" ("Yes", "No")

A.1.7 **City Name (Q7)** . "What is the name of the German city where you have lived for at least one year since 2008?" ("Aachen", "Augsburg", "Bergisch Gladbach", "Berlin", "Bielefeld", "Bochum", "Bonn", "Bottrop", "Braunschweig", "Bremen", "Bremerhaven", "Chemnitz", "Darmstadt", "Dortmund", "Dresden", "Duisburg", "Düsseldorf", "Erfurt", "Erlangen", "Essen", "Frankfurt am Main", "Freiburg im Breisgau", "Fürth", "Gelsenkirchen", "Göttingen", "Gütersloh", "Hagen", "Halle (Saale)", "Hamburg", "Hamm", "Hanau", "Hannover", "Heidelberg", "Heilbronn", "Herne", "Hildesheim", "Ingolstadt", "Jena", "Kaiserslautern", "Karlsruhe", "Kassel", "Kiel", "Koblenz", "Köln", "Krefeld", "Leipzig", "Leverkusen", "Lübeck", "Ludwigshafen am Rhein", "Magdeburg", "Mainz", "Mannheim", "Moers", "Mönchen-Gladbach", "Mülheim an der Ruhr", "München", "Münster", "Neuss", "Nürnberg", "Oberhausen", "Offenbach am Main", "Oldenburg (Oldb)", "Osnabrück", "Paderborn", "Pforzheim", "Potsdam", "Recklinghausen", "Regensburg", "Remscheid", "Reutlingen", "Rostock", "Saarbrücken", "Salzgitter", "Siegen", "Solingen", "Stuttgart", "Trier", "Ulm", "Wiesbaden", "Wolfsburg", "Wuppertal", "Würzburg")

A.1.8 **City Duration (Q8)** . "How many years have you lived in this city since 2008?" (Numerical input field)

A.1.9 **City Smartphone Usage Confirmation (Q9)** . "Have you owned and used a smartphone during this time?" ("Yes", "No")

A.2 Part 2: Everyday Smartphone Usage

A.2.1 *Daily Smartphone Screentime (Q10)* . "Please estimate your average daily smartphone screen time in hours." (Hint: By screen time, we mean the time you actively use your smartphone. For example, Apple smartphones offer these values via the 'Screen Time' function. On newer Android smartphones, this function is called 'Digital Wellbeing'.) *(Numerical input field)*

A.2.2 *App Usage in Everyday Life (Q11)* . "Please sort the following app categories according to the frequency with which you use apps from these categories. Sort in descending order, i.e., the most frequently used app category should come first, and the least used app category should be last. If you never use certain app categories, you do not have to include them and can ignore them. However, you must specify at least one category." (Hint: The apps in brackets only suggest the respective categories. The mouse can move the elements between the left and right lists. A double-click moves an element to the other list.) *(positional voting of app_categories)*

A.3 Part 3: Crisis Smartphone Usage

A.3.1 *Crisis Introduction (Comic)* . "Below is a short comic board that introduces you to the scenario of this survey. When we mention major cities, please think of the_city." *The comic board as depicted in Figure 1.*

A.3.2 *App Demand for Crisis-Specific Use Cases (Q12)* . "In the previous comic, an earthquake was introduced as an example of a crisis. For the following crises, please indicate which apps you would use. The categories are the same as in Q11. Multiple answers are possible." (U1) "You want to receive warnings about the crisis. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U2) "You want to receive tips on how to stay out of danger in the event of a crisis. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U3) "You want to contact an emergency service (but calling 112 is not possible). Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U4) "You want to search for information about the crisis. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U5) "You want to contact and communicate with other citizens. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U6) "You want to navigate in the_city. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")* (U7) "You want to be distracted from the crisis. Which apps would you use?" *(app_categories, "Other" with a text field, "No app," "No answer")*

A.4 Part 4: Crisis Smartphone Interaction with Island Connectivity

A.4.1 *City Isolation Introduction (Comic)* . "The following comic builds on the previous scenario and expands on it, continuing the earthquake scenario in the_city." *The comicboard as depicted in Figure 3.*

A.4.2 *City Isolation Introduction (Video)* . "Please start the video if it does not start automatically. If this does not work either, there is a description of the video below the video."

A.4.3 *Video Transcript* . "First, you want to send a message to a friend outside the city using a messenger app. The message is usually sent via the Internet. However, the earthquake disconnected the city from the Internet, so the message could not reach its destination. After the recovery measures taken by the mobile network operator, all apps on your smartphone are available for local use. Therefore, you can continue using your already installed apps to a limited extent, e.g., you can only navigate inside the city with your navigation apps, only read the local newspaper app but not global newspaper apps, or only communicate with citizens inside the city with your messenger apps, but not with people outside the city."

A.4.4 *Video Description* . "You can see how a chat message, symbolized by a letter, is to be sent to another city. However, the message cannot be sent because your city is disconnected from the Internet. You can then see that the mobile operator enables the use of your smartphone apps locally. This means, for example, that you can only use your navigation apps to navigate within the city, that you can only read your local newspaper app but not a global news app, or that you can only use your messenger apps to communicate with people within the city but not with people outside it."

A.4.5 *App Usage in Isolated Cities (Q13)* . "The consequences of the earthquake have now lasted for several days. For this question, please assume that your apps work locally in the city, as the video explains. Apps from all categories are available in this way. Such local recovery measures are visions for the future that have not yet been implemented. To answer this question, please imagine that apps on your smartphone work locally, as shown in the comic and video. The categories are the same as for Q11 and Q12. Please sort the following app categories according to the frequency you use apps from these categories. Sort in descending order, i.e., the most frequently used app category should come first, and the least used app category should be last. If you never use certain app categories, you do not have to include them and can ignore them. However, you must specify at least one category." (Hint: The apps in brackets only suggest the respective categories. The mouse can move the elements between the left and right lists. A double-click moves an element to the other list.) *(positional voting of app_categories)*

A.5 Part 5: Control Variables and Data Quality

"The following questions are independent of the previous scenario and relate only to you as a person."

A.5.1 *Crisis Experience (Q14)* . "Please indicate which of the following crises you have experienced." (Hint: We are not counting the COVID-19 pandemic as a crisis here.) *("Earthquake," "Flooding," "Storm," "Wildfire," "War," "None," "Not sure" with a textfield)*

A.5.2 *Internet Outage Experience (Q15)* . "Please indicate which of the following Internet outages you have experienced." (Hint: Fixed network refers to your Internet access at home, e.g., via Wi-Fi.) "No Internet via fixed network but via cellular network." *("Never," "up to 1h," "up to 6h," "up to 12h," "up to 24h," "more than")*

24h", "Not sure") "No Internet via cellular network but via fixed network." <"Never," "up to 1h", "up to 6h", "up to 12h", "up to 24h", "more than 24h", "Not sure"> "No Internet via fixed network nor cellular network." <"Never," "up to 1h", "up to 6h", "up to 12h", "up to 24h", "more than 24h", "Not sure">

A.5.3 Emergency Forces Experience (Q16) . "Are you currently working as an emergency worker, e.g., for the fire department, rescue service, police, or THW?" <"Yes," "No," "No answer," "Other" with a textfield>

A.5.4 ATI Scale with IRI (Q17) . "The following is about your interaction with technical systems. By 'technical systems,' we mean apps and other software applications as well as complete digital devices (e.g., cell phone, computer, TV, car navigation). Please indicate your level of agreement with the following statements." "I like to occupy myself in greater detail with technical systems." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "I like testing the functions of new technical systems." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "I predominantly deal with technical systems because I have to." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "When I have a new technical system in front of me, I try it extensively." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "I enjoy spending time becoming acquainted with a new technical

system." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "It is enough for me that a technical system works; I don't care how and why." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "Please indicate 'Completely disagree' here." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "I try to understand how a technical system exactly works." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "It is enough for me to know the basic functions of a technical system." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree"> "I try to make full use of the capabilities of a technical system." <"Completely disagree," "Largely disagree," "Slightly disagree," "Slightly agree," "Largely agree," "Completely agree">

A.5.5 Technical Issues Textfield (Q18) . "Did you have any technical problems answering the questions?" <Text field>

A.5.6 Additional Comments Textfield (Q19) . "If you have anything else you want to tell us, please use the text field below." <Text field>

A.5.7 SRSI UseMe (Q20) . "Please be honest: This survey is important to our research. Should we use your answers for our studies?" (Hint: Do not worry; the answer to this question is only for our data analysis. We will not pass them on to GapFish or third parties; they will not affect your payment.) <"Yes", "No">

Table 3: Rank Distributions by Variable and Group. For each variable, this table compares the median rank $M_g(a)$ that app category a has scored in group g . There is a row per group g , e.g., younger participants or participants with medium technology affinity, and a column per app category a , e.g., messengers or news apps. Our participants did not have to rank all app categories, and those that they left unranked are consequently not included in the median calculation. Note that our overall rank calculation in Table 2 uses the Borda count [76], which accounts for unranked apps during the calculation (e.g., for crisis apps, which many participants did not include).

Variable	Category	Messengers	News	Crisis	Social Media	Weather	Web Browsers	Navigation	Entertainment	Music	Utilities	Reference	Finance	Food & Drink	Games	Productivity	Books	Health & Fitness	Medical	Shopping	Education	Travel	Creativity	Sports
Age	Younger	2	3	2	4	6	6	6	7	5	6	7	7	7	10	10	6	8	11	11	4	11	12	12.5
	Middle-Aged	2	3	1	4	6	5	5	8	7	6	7	8	8	9	10	7.5	9.5	8	11	8	14	12	14
	Older	2	3	1	4	5	5	4	8	12	6	10	9	10	11	9.5	5.5	13	7	14	9.5	12	16.5	12
Gender	Men	2	3	2	4	6	5	5	8	6	5	8	7	8	10	10	8	8	8	12	6.5	13	13	12
	Women	2	3	2	4	6	6	6	7	6	7	6	8	8	10	9.5	6	10	10	10	7	11	12	16
	Non-Binary	3	4	4	4	9	3	8	4.5	8		5	4	6	3.5		2	1	2	9.5				
Education	Low	2	3	1	4	5	5.5	4	6	6	5	5	14	9	8.5	10.5	6	10	5.5	8	11	11.5	16	13
	Medium	2	3	2	4	6	5	5	8	6	6	7	8	8	10	11	5	9	10	12	6	13	13	13
	High	2	3	2	4	6	6	5	7	5	7	8	7	8	10	9	8	9	10	11	7	11	12.5	14
Income	Low	2	3	1	4	5	5	5	7	6	6	6	11	9.5	8.5	8	6	10	7	11	5	11	13	14.5
	Medium	2	3	2	4	6	5	5	7.5	6	6	7	7	8	9.5	10	6	10	10.5	10	7.5	11	12	12
	High	2	3	2	4	6	6	6	8	7	7	8	7	7	11	10	7	8	9	13.5	6	14	14	14.5
City Size	Medium	2	3	1.5	4	5.5	5	5	6	5.5	6	7	7	6	9	9	9	11	10.5	10	8.5	10.5	13.5	9
	Smaller	2	3	1	4	6.5	6	5	10	7	5.5	5.5	9.5	8	11	11.5	5	7.5	8.5	12	5	15	14	16
	Larger	2	3	2	4	6	5.5	6	8	8	6	9	8	9	9.5	10	12	9	9	12	9	12.5	10	12
	Metropolis	2	3	1	4	5	5	5	7	5	6	6	7.5	7.5	9	8	6	10	10	8.5	5	11	15	15
Residence Time	Short	2	2	2	5	7	4	5	7	5.5	8	7	10	8.5	10	8	7.5	16.5	15	12	4	10	14	12
	Medium	2	3	2	5	7.5	6.5	5	8	5	6	7	9	8	11	7.5	7	8	9	11.5	7	11.5	13	13
	Long	2	3	1	4	6	5	5	7	7	6	7	7	8	9	10	6	9	9	11	7	13	13	13
Screen Time	Medium	2	3	1	4	5	5	5	7	7	6	7	8	9	11.5	10	5	9.5	8	14	6.5	14.5	13	15
	High	2	3	2	4	6	5	5	8	5	6	7	7	7	10	9	7	9.5	10	10	8	11	13.5	12
	Low	2	3	2	4	6	6	6	8	6	5.5	7	8	8	9	10.5	7	9	10	10	5	12	11.5	13
Crisis Experience	Yes	2	3	3	4	6	5	5	11	6	5.5	13	7	8	19.5	11	7.5	3	15	12	5.5	15.5	16.5	18
	No	2	3	1	4	6	5	5	7	6	6	7	8	8	10	10	6	10	10	11	7	12	13	13
Outage Experience	L. Out.: Yes	2	3	2	4	5	5	5	8	4	4	7.5	6.5	6	8	10	4.5	10	9	11	6	11	11	13
	L. Out.: No	2	3	1.5	4	6	6	5	7	6	6	7	8	8	10	10	7	9	10	11	7	12	13	13
	C. Out.: Yes	2	3	2	5	6	6	6	11	6	5.5	8	7	8.5	10.5	12	4	8	9	11	3.5	14	11	13
	C. Out.: No	2	3	1	4	6	5	5	7	6	6	7	8	8	10	9	7	10	10	11	7	12	13	13
	L.+C. Out.: Yes	2	3	2	5	6	6	6	11	6	5	9	6	7.5	9	10	5	8.5	9	10	4	13	10	11
	L.+C. Out.: No	2	3	1	4	6	5	5	7	6	6	7	8	8	10	9	6	9.5	10	12	7	12	14	14
First Responder	Yes	2	3	3	4	6	5	5	11	6	5.5	13	7	8	19.5	11	7.5	3	15	12	5.5	15.5	16.5	18
	No	2	3	1	4	6	5	5	7	6	6	7	8	8	10	10	6	10	10	11	7	12	13	13
Technology Affinity	Low	1	3	2	4	5	5	4	7	5	5	8	6	7.5	10	10	5	8	13	8	4	14	12	13
	Medium	2	3	2	4	6	5	5	7	6	6	7	8	8	10	10	6	9	10	11	6	12	12.5	13
	High	2	3	1	5	6.5	6	5	8	7	6	6.5	7.5	9	10	9	9	10	8	12	8	12	14.5	15

Table 4: Study Sample Description. This table describes our study sample in terms of the variables derived from the 20 questions of our questionnaire. We depict demographic variables from Q1 - Q4 and control variables from the remaining questionnaire. Q6 and Q9 were simple confirmations, dependent variables Q11 - Q13 answered our research questions (Section 5), and we used the open-ended questions Q18 - Q20 to discuss data quality. We introduce categories for each variable to simplify discussions about individual groups. For numerical variables, we report the mean μ , median M , and the inter-quartile range (IQR).

Variable	Category	Value	Goal	Sample ($N = 857$)	Description
Age (Q1)	Younger	18 yr to 29 yr	21 %	12 % (103)	
		30 yr to 39 yr	19 %	20 % (174)	
	Middle-Aged	40 yr to 49 yr	18 %	21 % (180)	
		50 yr to 59 yr	24 %	26 % (225)	
	Older	60 yr to 69 yr ≥ 70 yr	19 % 0 %	20 % (175) (0)	
Gender (Q2)	Men	Male	50 %	49 % (416)	
	Women	Female	50 %	51 % (437)	
	Non-Binary	Non-Binary		0 % (4)	
Education ¹ (Q3)	Low	Low	26 %	13 % (109)	
	Medium	Medium	33 %	59 % (502)	
	High	High	40 %	29 % (246)	
Income (Q4)	Low	<1000 €	6 %	6 % (48)	
		1000 € to 1999 €	16 %	18 % (152)	
	Medium	2000 € to 2999 €	22 %	24 % (202)	
		3000 € to 3999 €	21 %	22 % (186)	
	High	4000 € to 4999 € ≥ 5000 €	22 % 14 %	16 % (135) 14 % (122)	
Federal State (Q5)					Our sample covers all 16 federal states ² .
City Size (Q7)	Smaller	100 000 to 199 999		13 % (138)	Our sample covers 79 of the 82 German major cities
	Medium	200 000 to 499 999		22 % (201)	
	Larger	500 000 to 999 999		29 % (244)	
	Metropolis	≥ 1 000 000		32 % (274)	
Residence Time (Q8)	Short	<4 years		13 % (109)	$\mu = 13.48$; $M = 16$; IQR = 3
	Medium	4-9 years		59 % (502)	
	Long	≥ 10 years		29 % (246)	
Screen Time (Q10)	Low	<3 h		41 % (348)	$\mu = 3 \text{ h } 44 \text{ min}$; $M = 3 \text{ h}$; IQR = 3 h
	Medium	3 h to 5 h		31 % (266)	
	High	≥ 5 h		28 % (243)	
Crisis Experience (Q14)	Earthquake			14 % (122)	
	Flooding			19 % (163)	
	Storm			49 % (424)	
	Wildfire			5 % (46)	
	War			3 % (24)	
Outage Experience ³ (Q15)	L. Outage	≥ 6 h		25 % (213)	
	C. Outage	≥ 6 h		14 % (116)	
	L.+C. Outage	≥ 6 h		13 % (113)	
First Responder (Q16)				3 % (28)	
Technology Affinity (Q17)	Low	1 to 2.5		12 % (100)	$\mu = 3.68$; $M = 3.67$; IQR = 1.33; $\alpha = .9$
	Medium	2.5 to 4.5		67 % (578)	
	High	4.5 to 6		21 % (179)	

¹ "Low" is ⟨ "None (yet)," "Hauptschulabschluss" ⟩, "Medium" is ⟨ "Polytechnische Oberschule," "Mittlere Reife, Realschulabschluss," "(Fach-)Hochschulreife" ⟩, "High" is ⟨ "Bachelor," "Master," "Diplom," "Promotion" ⟩.

² 6 % Baden-Württemberg (BW), 11 % Bavaria (BY), 16 % Berlin (BE), 1 % Brandenburg (BB), 2 % Bremen (HB), 8 % Hamburg (HH), 6 % Hesse (HE), 1 % Mecklenburg-Western Pomerania (MV), 5 % Lower Saxony (NI), 28 % North Rhine-Westphalia (NW), 3 % Rhineland-Palatinate (RP), 0 % Saarland (SL), 7 % Saxony (SN), 3 % Saxony-Anhalt (ST), 3 % Schleswig-Holstein (SH), and 2 % Thuringia (TH).

³ "L." and "C." are abbreviations for "Landline" and "Cellular," respectively.

Sounds Good? Fast and Secure Contact Exchange in Groups

Citation	F. Putz, S. Haesler, and M. Hollick. "Sounds Good? Fast and Secure Contact Exchange in Groups." In: <i>Proceedings of the ACM on Human-Computer Interaction</i> 8. CSCW '24, Nov. 2024. DOI: 10.1145/3686964
Contribution	This publication is discussed in Chapter 4. My contribution is described in the section " <i>Collaborations and My Contribution</i> " on page xi.
Rights notice	© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is included in this dissertation in accordance with ACM author rights. The definitive Version of Record was published in <i>Proceedings of the ACM on Human-Computer Interaction</i> , DOI: https://doi.org/10.1145/3686964 .

Sounds Good? Fast and Secure Contact Exchange in Groups

FLORENTIN PUTZ, Technical University of Darmstadt, Germany

STEFFEN HAESLER, Technical University of Darmstadt, Germany

MATTHIAS HOLLICK, Technical University of Darmstadt, Germany



Fig. 1. Group of users exchanging their contact information using our acoustic group pairing protocol.

Trustworthy digital communication requires the secure exchange of contact information, but current approaches lack usability and scalability for larger groups of users. We evaluate the usability of two secure contact exchange systems: the current state of the art, SafeSlinger, and our newly designed protocol, *PairSonic*, which extends trust from physical encounters to spontaneous online communication. Our lab study ($N = 45$) demonstrates PairSonic's superior usability, automating the tedious verification tasks from previous approaches via an acoustic out-of-band channel. Although participants significantly preferred our system, minimizing user effort surprisingly decreased the perceived security for some users, who associated security with complexity. We discuss user perceptions of the different protocol components and identify remaining usability barriers for CSCW application scenarios.

CCS Concepts: • **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; **Empirical studies in collaborative and social computing**; **Computer supported cooperative**

Authors' Contact Information: [Florentin Putz](mailto:fputz@seemoo.de), Technical University of Darmstadt, Darmstadt, Germany, fputz@seemoo.de; [Steffen Haesler](mailto:haesler@peasec.tu-darmstadt.de), Technical University of Darmstadt, Darmstadt, Germany, haesler@peasec.tu-darmstadt.de; [Matthias Hollick](mailto:mhollick@seemoo.de), Technical University of Darmstadt, Darmstadt, Germany, mhollick@seemoo.de.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of the ACM on Human-Computer Interaction*, <https://doi.org/10.1145/3686964>.

work; Collaborative interaction; Smartphones; • Security and privacy → Usability in security and privacy; Authentication.

Additional Key Words and Phrases: usable security; group pairing; secure device pairing; trust establishment; acoustic communication; data-over-sound; authentication ceremony; key verification; peer-to-peer-authentication; public-key cryptography; security; privacy; social cybersecurity; end-to-end-encrypted messaging; MitM attacks; instant-messaging; online collaboration; contact management; contact sharing; device association; binding; bonding; coupling; decentral; ad-hoc; interoperability; decentral; spontaneous; nearby; proximity; smartphone; user interaction; user perception; prototype; open-source; WiFi; NFC

ACM Reference Format:

Florentin Putz, Steffen Haesler, and Matthias Hollick. 2024. Sounds Good? Fast and Secure Contact Exchange in Groups. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 425 (November 2024), 44 pages. <https://doi.org/10.1145/3686964>

1 Introduction

Imagine a group of seven researchers who just met for the first time at a conference. During the event, they notice that they share a lot of common interests – they might even have some ideas for future collaborations – so they decide to stay in contact. They all have a smartphone, but to securely communicate online, they need a way to *securely exchange their contact information*. In this paper, we study *group pairing* methods for the fast and secure exchange of contact information that fulfill the following two goals: first, they allow the participants to contact each other online, to establish new conversations or to start new digital collaborations. Second, they authenticate the participants by leveraging their physical encounter to verify their contact information.¹

There are countless similar use cases for online collaboration besides the aforementioned group of researchers, ranging from professional activities (team communication, remote work, project management), to educational settings (interactive learning platforms, knowledge management), and leisure (event planning and coordination), affecting billions of users in today’s digital world [42, 57, 75, 82, 109]. For all these scenarios, secure exchange of contact information is required to associate known physical individuals with their online profiles and protect their privacy and security when interacting online. Without such a secure contact exchange, an adversary can eavesdrop on private communication or even impersonate one of the communication participants in order to modify or create fake messages that appear legitimate.

Effective communication and collaboration thrives in a trusted environment, where participants have clear knowledge of the identities of others involved and no concerns about unauthorized access to their conversations. Such an environment encourages open discussions, even on sensitive topics [104]. Previous research has shown that users desire authenticated conversations – wanting to know exactly *who* they are communicating with – when talking to friends and family members [26], but also when communicating with colleagues or business partners [25], and especially so for vulnerable groups such as journalists and activists [45]. Secure contact exchange is therefore an important prerequisite for fruitful and trustworthy online communication and collaboration [86].

1.1 Current Group Pairing Is Insufficient

Throughout the last decade, users have increasingly adopted end-to-end encrypted (E2EE) tools such as Signal [106], Whatsapp [80], Telegram [112], and Threema [113], facilitating group collaboration with features like real-time messaging, file sharing, and video calls. These E2EE tools also offer the verification of contact information through an *authentication ceremony*, which requires users to perform a manual verification action, either by scanning quick-response (QR) codes or by manually

¹The exchanged contact information usually contains additional metadata and public-key material, e.g., for secure end-to-end encrypted (E2EE) instant messaging.

comparing public key fingerprints on their devices. This authentication ceremony allows users to transfer an existing trust relationship from a physical encounter to the corresponding digital identity of their communication partners, enabling authenticated conversations. There are two problems with the design of authentication ceremonies in state-of-the-art tools for secure communication and collaboration:

- (1) **Low usability.** Current authentication ceremonies are *hard to use*, even for professionals.
- (2) **Bad scalability.** Current authentication ceremonies *do not support larger groups* of users.

First, previous usability research has shown that users struggle to understand and perform authentication ceremonies [86, 122, 123, 129]. Most users expect that E2EE tools like Signal are secure by default without further verification actions, when in reality these tools cannot protect against active machine-in-the-middle (MitM) attacks and rogue service operators unless the users diligently perform the authentication ceremony [45, 121]. These authentication ceremonies, however, require considerable user interaction [44], take a long time [122], and are difficult to use even for security professionals [26, 100]. Recently, the risks of active MitM attacks in E2EE tools have gained more attention due to reports of law enforcement agencies compromising app providers [37, 69], demonstrating a practical and concerning risk for privacy infringements, especially if such vulnerabilities were to be exploited by malicious actors in the future.

Second, state-of-the-art tools like Signal and WhatsApp do not support efficient and secure contact exchange for *groups of multiple users*, but only between two users. This is a substantial shortcoming, as collaboration often involves more than two participants. While each pair of group members could separately perform the bilateral authentication ceremony one after another, this is infeasible as the number of manual verification actions scales quadratically ($\frac{n(n-1)}{2}$) with the group size (n). To illustrate this: if our exemplary group of seven researchers were to use Signal, they would need to perform 21 manual actions amongst each other for authentication, but each user can only perform one of these actions at a time. Such an authentication takes too much time and would discourage the spontaneous decision to collaborate.

The main challenge is how to authenticate the exchanged contact information and key material without having to rely on any trusted third parties or public-key infrastructures (PKIs) that facilitate the exchange. Farb et al. [23] achieved a breakthrough towards solving this problem when they proposed the more usable *SafeSlinger* protocol for the secure exchange of contact information in the absence of trusted third parties. While SafeSlinger offers significantly higher usability than alternatives [23], it still requires the participants to perform $\frac{n(n-1)}{2}$ manual text comparisons to achieve its security guarantees. Since the release of SafeSlinger, several studies on authentication methods suggested that users generally prefer simpler schemes requiring less involvement [78, 130] and that involving users for manual comparisons is prone to errors [53, 111, 122]. Furthermore, recent research has explored the intuitive actions users might take to associate multiple devices [14, 48], suggesting that the user experience in SafeSlinger may not align with user expectations. The required amount of user interaction is potentially discouraging, especially considering larger groups. Instead of facilitating trustworthy communication and collaboration, state-of-the-art methods rather inhibit it (see Section 3 for a more extensive discussion of related work).

1.2 Approach

In this work, we approach the challenge of securely exchanging contact information from a different angle, under the research hypothesis that *users would prefer a contact exchange method with minimal user interaction* [64, 78, 130]. Consequently, we design a new group pairing protocol named *PairSonic* that requires less effort and can efficiently scale to multiple participants. The illustrative group of researchers can exchange their contact information securely and quickly by starting PairSonic and

simply holding their devices close together for a few seconds, aligning with intuitive association behavior [14, 48, 50]. PairSonic automatically exchanges the participants' contact information using an ad-hoc WiFi channel for communication and a location-limited acoustic out-of-band (OOB) channel for verification. The effectiveness of the acoustic channel has already been demonstrated in previous research [32, 74, 76, 78] and in industry adoption, such as with the Sonos smart speaker [52], for the similar use case of pairwise key verification. Even though the acoustic channel seems promising, its applicability and usability between more than two devices remains unexplored.

1.3 Contributions

Our main contribution to CSCW is the usability evaluation of our novel acoustic approach PairSonic for the secure exchange of contact information and cryptographic public keys for groups of multiple users as a basic requirement for teams to work with sensitive content. We conduct a lab study with $N = 45$ participants – consisting of practical authentication tasks on smartphones, subsequent interviews, and a questionnaire – to compare the state-of-the-art system SafeSlinger with our novel protocol PairSonic, which we design and implement for smartphones. PairSonic requires only minimal user interaction by leveraging an acoustic OOB channel between the users' smartphones, automating the cumbersome verification tasks from previous commercial and academic approaches [19, 45, 117, 118, 122]. This paper is structured as follows:

- We define the problem and establish our requirements for secure contact exchange (Section 2), discussing why the state of the art fails to solve this problem (Section 3).
- We design PairSonic, an improved protocol for fast and secure contact exchange, and implement it for smartphones (Section 5).
- To answer our research questions (Section 4), aiming to understand how automating the cumbersome verification can improve on the current state of the art (SafeSlinger), we conduct a lab study with $N = 45$ participants to compare both systems (Section 6).
- Based on the quantitative insights from our questionnaire (Section 7) and qualitative interviews (Section 8), we discuss our observation that less user interaction seems to improve usability in our case, but can have unexpected negative effects on perceived security. We evaluate the suitability of the acoustic OOB channel for pairing and give recommendations how to improve the state of the art of secure contact exchange (Section 9).
- We provide a replication package with our evaluation scripts and the pseudonymized dataset from our study, which contains 69 variables for each of the 45 participants [90].

2 Problem Setting

This section defines the problem we are addressing (Section 2.1) and our assumptions (Section 2.2).

2.1 Problem Definition

Our goal is to let a group of two or more users that physically meet (Figure 1) spontaneously exchange their contact information [23]. The expected outcome of this *secure contact exchange*² is that each participant obtains the authentic and verified contact information of all other participants, including their cryptographic public keys, without having to rely on external key management infrastructure, prior associations, or shared secrets. We target a group of users who physically meet for the pairing process (Section 9.9) and want to be able to set up secure communication

²The process of exchanging and verifying the identities with a group of participants is sometimes also called *group key verification*, or *group authentication ceremony*, or *peer-to-peer authentication*, or *secure group contact sharing*, or *key establishment for groups*, or *spontaneous device association*. In this work, we call this process *secure contact exchange* or *group pairing*, as we always target groups of two or more users unless otherwise noted.

channels between all subsets of users within the group. After this pairing process, the users can subsequently communicate in a confidential and authenticated manner over untrusted channels such as the Internet. Our contact exchange protocol has the following requirements:

USABILITY REQUIREMENTS

- (1) **Minimal user interaction**, due to our research hypothesis that users would prefer such a contact exchange method [64, 78, 130]. The protocol should be easy to use for lay users, facilitating spontaneous ad-hoc collaboration plans [13, 16].
- (2) **Scalable to larger groups of users**. The pairing protocol should be fast and not noticeably depend on the number of users.
- (3) **Error-proof**, as users can make mistakes. In particular, our protocol should be robust to *rushing users* who prioritize speed and efficiency over adhering to instructions or prompts, potentially compromising the security of the protocol.

DEPLOYABILITY REQUIREMENTS

- (4) **Compatibility**. Our protocol should seamlessly run on existing commercial off-the-shelf (COTS) devices, enabling users to leverage its benefits immediately by simply installing our software. We focus on user-controlled wireless devices such as smartphones, tablets, laptops, and smart watches, without requiring further hardware or firmware modifications.
- (5) **Ad-hoc**. Our protocol should not require any existing security context in the form of pre-shared keys or a jointly trusted third party.
- (6) **Decentral**. Our protocol should operate independently of any central infrastructure and remain fully functional without an active Internet connection. This requirement is crucial to support people's spontaneous decision-making, facilitating seamless collaboration and communication regardless of their location. This means that no metadata can leak to third parties, as part of our privacy-by-design approach. If a security vulnerability arises in our protocol or implementation, the offline nature of our exchange makes exploitation by attackers significantly more challenging.

SECURITY REQUIREMENTS

- (7) **Confidentiality**. Only the intended participants should receive the contact information.
- (8) **Contact authentication**. Each received contact information should match the corresponding honest participant. We require mutual authentication, i.e., each participant authenticates all the other participants.
- (9) **Collective pairwise security**. This requirement was postulated by Farb et al. [23], stipulating that each pair of participants in our groupwise exchange should get the same security properties as if they had performed a direct pairwise exchange. This implies that adversarial participants cannot degrade the bindings between other honest participants.

2.2 Assumptions

We assume that all users are physically present at the same location and can utilize the concept of “*group demonstrative identification*” [70], which involves legitimate participants excluding unintended participants or imposters based on personal identification, such as appearance or voice. The denotation of intended and legitimate participants is an inherently social classification – the decision of whom *users* expect to communicate with cannot be achieved by technical measures alone and needs human assistance. Users are responsible for verifying the received contact information from their exchange to ensure it matches their intended participants [23]. They should also reject non-participants and detect any instances of impersonation within the group, such

as duplicate entries corresponding to honest participants. We also assume that the participants accurately determine their group size.

2.3 Adversary Model

Although a detailed security evaluation is beyond the scope of this work, we consider potential threats from an adversary with diverse capabilities. The attacker’s goal is to breach the security requirements stated in [Section 2.1](#). For example, the adversary might attempt to impersonate one of the legitimate participants by making another participant accept an adversarial cryptographic public key. Such an attack would allow the attacker to launch MitM attacks in the future, manipulating or intercepting confidential communication. However, ensuring the uninterrupted availability of the contact exchange process is not one of our security goals. Therefore, if an adversary disrupts the contact exchange (e.g., by jamming the radio channel), users would need to retry the exchange at another time or place.

Our adversary model is based on the model used in SafeSlinger [23], focusing on both nearby and remote adversaries. These adversaries possess full Dolev-Yao [20] control over WiFi network messages, meaning they can eavesdrop, intercept, modify, replay, or inject radio communication between devices. Likewise, we also assume that the adversary has not compromised the hardware and software of the participants’ smartphones, as such attacks are beyond this paper’s scope.

In addition to SafeSlinger’s adversary model, our research also considers the acoustic out-of-band channel. Based on previous research by Stajano and Anderson [108] and Balfanz et al. [5], we regard this channel as location-limited, characterized by its ability to support *demonstrative identification* and ensure *authenticity* due to the inherent physical constraints of acoustic sound pressure waves. Consequently, while adversaries can eavesdrop on this channel, they cannot transmit undetected. We explore this assumption and potential attacks on the acoustic channel in more detail in [Section 9.7](#).

3 Related Work

Our work focuses on the problem of securely exchanging contact information, including cryptographic key material, in groups of users. This problem relates to several branches of research, which we first summarize and then elaborate in further detail:

- **Device association**³ is the general act of establishing a communication channel between two or more devices, independent of the duration and the association’s security [16]. Our problem is a special case of device association with a specific focus on authenticity and security to facilitate trustworthy communication.
- **Secure device pairing** is closely related to our work, as it involves device association aimed at establishing an authentic communication channel [28]. Whereas previous work mostly focused on connecting two devices (“pairing”), our work additionally considers the more complex case of connecting multiple users, although it also functions with just two users. A key distinction is that we aim not only to establish a secure communication channel for the entire group, but potentially for every subset of group members as well: CSCW research has shown that larger groups sometimes like to split dynamically into smaller task forces for more effective collaboration [43, 116]. The group exists temporarily to facilitate the exchange of contact information among its members, but subsequent communication can happen independently of the group’s context. Naturally, the group members can also opt to collaborate as a whole group.

³Device association is sometimes also called binding, coupling, or bonding [16].

- **Acoustic communication** can be used as a physical OOB channel to transmit data (such as contact information) between nearby devices, which is compatible with all smart devices having a microphone and a speaker. Previous work showed promising results for connecting two devices [32, 74, 78]. We are the first, to the best of our knowledge, to study the suitability and usability of acoustic communication for connecting multiple users.
- **Authentication ceremonies** refer to the process by which users can verify their cryptographic keys in the context of E2EE tools [44]. Existing approaches only have limited usability and do not support more than two users. In contrast, our approach aims to be a more usable authentication ceremony supporting multiple participants.

3.1 Device Association

Device association methods establish a basic communication channel between two or more devices [16, 49, 50]. However, since these methods do not consider the security of the association, they do not directly aid secure contact exchange. Nonetheless, usability evaluations exist on preferred user actions or gestures for associating their device with others [13, 40, 61]. These evaluations can help us optimize the user interaction for the secure exchange of contact information, which also requires temporary device association with other participants. One such line of work is especially interesting: guessability studies with plastic surrogates to determine intuitive user actions for associating a group of devices [12, 14, 48, 51]. Instead of multiple pairwise interactions, users preferred a singular group-wise interaction, favorably by pointing their smartphones towards each other or bringing them into proximity. These results inform the design of our approach, where we use a single group-wise interaction, requiring all smartphones to be in close proximity during the exchange.

3.2 Secure Device Pairing

There is a large body of work dealing with the secure exchange of cryptographic public keys, called secure device pairing, which is very similar to our more general problem of the secure exchange of contact information. This line of work mostly focused on establishing keys between two devices [28], and previous usability studies focused mostly on this use case [13, 47, 53, 59, 62, 118, 119]. Pairing between two devices does not securely generalize to more than two parties due to new types of attacks that are possible when there are multiple participants [23, 64].

However, there is also some prior work on establishing keys between multiple devices, which roughly falls into one of four categories: (1) human-in-the-loop, (2) shared homogeneous context, (3) physical layer security, (4) OOB channels.

3.2.1 Humans-in-the-loop. There is previous research focusing on key establishment where each participant has to perform some manual action, such as entering a shared secret [1, 3, 8, 120], comparing text [65, 81, 84, 109, 120] or visual patterns [11, 56, 72, 73], or moving all devices in a correlated manner [58, 77]. Besides the obvious drawback that these schemes require substantial user interaction and do not scale well, they often cannot protect against malicious insiders [64].

3.2.2 Shared Homogeneous Context (Zero-interaction Pairing). Zero-interaction schemes try to establish a shared key based on the entropy of observable common events in the surroundings [29, 130]. While such approaches at first glance seem to fit well to our requirement of minimal user interaction, they are not suitable for user-controlled device association and take more time than is tolerable for an ad-hoc spontaneous exchange [24, 71]. These schemes were designed for a different use case, namely for pairing multiple devices owned by a single user, where the key establishment process is allowed to take its time in the background, without user involvement. Instead, group

pairing is a social use case, which not only involves devices but also multiple people, and thus a group pairing protocol should not neglect the social interaction in the group.

3.2.3 Physical Layer Security. Another approach makes use of the physical laws of signal propagation on the wireless radio channel to securely establish keys [34–36, 41, 46]. A similar line of work uses a Faraday cage to prevent outside attackers from interacting with the key establishment [63, 66]. These approaches are often hard to deploy, as they either require specific firmware or hardware, making them incompatible with COTS devices, or they require additional devices to assist with the pairing process.

3.2.4 Out-of-Band Channels. One line of work studies key establishment using OOB channels, which are low-bandwidth auxiliary channels that are usually location-limited to protect the integrity of its messages [108]. Previous research considered infrared channels [5], light channels [60, 88, 99], and near-field communication (NFC) [15]. The audio channel has also been explored as a promising OOB channel for pairing two devices [38, 39, 89, 107]. As we develop this concept further for pairing multiple devices, we will elaborate on the details of the audio channel in the following section.

3.3 Acoustic Communication

The acoustic channel is an attractive choice for connecting nearby devices, thanks to its ubiquitous compatibility with smartphones by repurposing their integrated audio hardware [32]. The physical layer can be entirely software-defined [74], facilitating easy adaptation of modem properties to fit specific communication scenarios. Audio communication is often employed as a location-limited OOB channel, as it has much shorter range than radio communication and is restricted by physical barriers like walls [107]. Several proposals in the literature have explored using audio hardware for communication, including both audible and ultrasonic inaudible frequencies [33, 67, 83, 89, 98, 125].

We are only aware of one previous study by Mehrabi et al. [78] evaluating the usability of acoustic data transmission. In a lab study involving 12 participants in a pairwise data exchange scenario, they compared audible acoustic communication with Bluetooth Low Energy (BLE) and QR codes. Transaction times were lowest for audio and considerably higher for BLE. QR codes and audio had significantly higher usability than BLE, highlighting the promising nature of the acoustic OOB channel. In our work, we build on these previous works to investigate the acoustic channel's suitability and usability for connecting multiple users simultaneously.

3.4 Authentication Ceremonies in Secure Messaging Applications

Modern E2EE tools such as Signal or WhatsApp use opportunistic E2EE (also known as trust on first use (TOFU)), requiring almost no user interaction. However, TOFU cannot defend against active attacks unless users perform the *authentication ceremony*, which is the process by which users can verify their cryptographic keys [45, 117]. Authentication ceremonies usually require a human in the loop to match specific representations of the cryptographic fingerprint, such as hexadecimal strings, word phrases, or visualizations [19, 111].

The authentication ceremonies in E2EE tools are not designed for multiple people and would require multiple pairwise instantiations, which is a prohibitive amount of effort in a collaborative scenario. In addition to that, previous user studies have shown that even for just two users, current authentication ceremonies are too hard to use [2, 4] and error-prone [102, 103], which means that, in practice, people almost never perform them [26, 86, 100, 122, 123, 129]. This is especially alarming considering that the default mode without performing the authentication ceremony fails to meet users' security expectations; most users are unaware of this issue [44].

3.5 SafeSlinger

The most promising approach to date in literature is the SafeSlinger protocol by Farb et al. [23], enabling the secure exchange of contact information between multiple people. Built upon the recommendations of a previous user study on group pairing protocols [85], SafeSlinger surpasses prior human-in-the-loop approaches [11, 73] in usability, efficiency, and security, making it the state of the art in group pairing protocols. It was designed specifically for groups of people to address the security and usability problems of using pairwise protocols in such scenarios. Participants perform the following steps while using SafeSlinger:

- (1) **Initialization.** Each participant must select the group size. Next, each participant sees a number on their smartphone. They must coordinate to identify and enter the lowest number among all participants on their smartphone.
- (2) **Verification.** Each participant gets displayed three word phrases. They must coordinate to identify and select the phrase common to all smartphones.
- (3) **Finalization.** The protocol then securely exchanges their contact information. Finally, each participant verifies that they received correct contact information from the others.

In their evaluation, SafeSlinger was significantly faster and more usable than the pairwise contact exchange method Bump,⁴ as shown by a within-subjects user study ($N = 24$). However, each step of the SafeSlinger protocol requires all participants to perform multiple comparisons with all other group members. This increases effort for larger group sizes and is error-prone [53, 111]. Furthermore, SafeSlinger requires a stable Internet connection during the pairing process and relies on a third-party server to facilitate the contact exchange. These limitations prevent SafeSlinger from meeting our usability and deployability requirements (Section 2).

4 Research Questions

Based on our analysis of related work in the previous section, the main problem of previous approaches is the lack of usability for ordinary users, which gets amplified for larger groups of people due to inadequate scalability. SafeSlinger is the most promising approach so far, but lacks in scalability and deployability. The authentication ceremonies in E2EE tools, as well as most current secure device pairing methods, are solely designed for two devices. Users expect a single associating interaction for the entire group [14, 48, 50], but pairwise schemes cannot be easily generalized to multiple participants [23].

In this work, we explore the problem of secure contact exchange from a new angle, under the following research hypothesis: *users prefer a contact exchange method with minimal user interaction*. Consequently, we design a new protocol named PairSonic that minimizes user interaction, implement it on smartphones, and conduct a lab study comparing it with the current state-of-the-art approach, SafeSlinger. The research questions we address in our work are:

- **RQ1.** Which initialization and verification steps do users prefer?
- **RQ2.** Which method has better usability?
- **RQ3.** How do users perceive the security of both methods?
- **RQ4.** How do users like the audible acoustic OOB channel for pairing?

5 PairSonic: Acoustic Group Pairing Protocol

This section describes the design and implementation of our novel acoustic group pairing protocol, called *PairSonic*. We developed it to meet the usability, deployability, and security requirements outlined in Section 2. Our starting point was the design of SafeSlinger, which we identified as the

⁴Bump was discontinued in 2014.

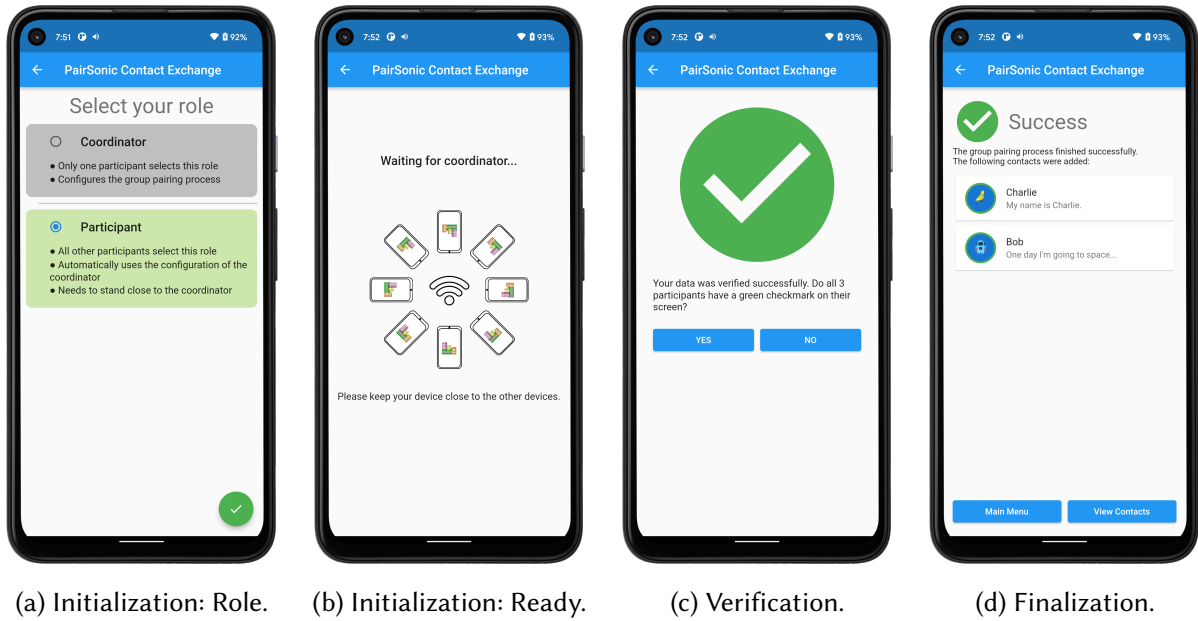


Fig. 2. This figure shows the PairSonic contact exchange from the perspective of the *participant* role. (a) The participants begin by selecting their role. (b) They then wait for the coordinator to initiate the exchange via the acoustic OOB channel, which is shown in Figure 10. (c) Next, they verify that each group member's screen displays a green checkmark. (d) After this process, the app shows the exchanged contacts.

most promising approach to date in Section 3, as it fulfills all our security requirements – but not all usability and deployability requirements.

5.1 Protocol Overview

PairSonic, our innovative acoustic group pairing protocol, involves several steps as shown in Figure 2. First, a group of users wishing to exchange contact data initiates PairSonic on their smartphones. This brings them to the first screen, where they nominate a temporary *coordinator* (Figure 2a). The rest of the group assumes the role of *participants*, which brings their devices into the ready state (Figure 2b). The coordinator enters the total group size and confirms that all participants' smartphones are ready and nearby.⁵

Following this, the coordinator's smartphone creates a temporary ad-hoc WiFi network, transmitting a short acoustic message containing the network details to other smartphones. This allows them to automatically join the network.⁶

Next, the devices use this WiFi network to perform a cryptographic protocol, which is based on the SafeSlinger protocol. Each participant's smartphone sends a nested commitment with their contact data⁷ to the coordinator's smartphone, which then disperses it to the other devices. The coordinator's smartphone generates a hash value from all participants' commitments and contact information, including their cryptographic public keys, and transmits this hash value over the acoustic OOB channel. Other smartphones are monitoring this channel and will abort if they receive any other message. Upon receipt of the correct hash value, a green checkmark appears on the smartphones, prompting users to confirm that all other smartphones display the same checkmark

⁵The coordinator's perspective of the PairSonic contact exchange is depicted in Figure 10 (appendix).

⁶PairSonic can function even if a user is already connected to a WiFi network; the user will temporarily switch to the ad-hoc network instead.

⁷The exchanged data includes the cryptographic public key and, optionally, further information such as a profile picture.

(Figure 2c). Upon confirmation, the smartphones use the WiFi network to publish success nonces. These allow others to verify the commitments and, if verification is successful, display the exchanged contacts (Figure 2d). Conversely, if the checkmark is not confirmed, they release abort nonces, terminating the protocol for all participants. This procedure ensures that all participants receive accurate and verified contact information, including authenticated cryptographic public keys.

5.2 Automating the User Interaction With Acoustic Communication

To reduce user effort during the protocol, we automated the majority of SafeSlinger’s user interaction. Instead of relying on users to act as data channels via error-prone shared device interactions, we utilize an OOB channel to verify the integrity of the exchanged contact information by directly connecting the devices, thereby minimizing the required user interaction. We use an acoustic location-limited channel where smartphones repurpose their integrated speakers and microphones to exchange data within a short radius of about one meter. The smartphone speaker, rather than playing music, emits sound waves that encode information by varying their frequency or amplitude. This process has been used to connect two devices in related work (Section 3.3), and we extend this concept to multiple devices. The other smartphones capture the sound waves, decode the embedded data, and use it as an additional security layer to verify the integrity of the exchanged contact information.

The acoustic channel uniquely meets all our deployability requirements, with the added advantage that its physical layer is entirely customizable in software. This means the security of the exchanged data can be directly protected on the physical layer [89], an advantage when implementing our ad-hoc requirement without having to rely on existing security contexts such as pre-shared keys.

Another reason for choosing an acoustic OOB channel is that the acoustic data transmission only works when the users are in proximity and bring their devices close to each other, as depicted in Figure 1. The physical interaction of moving all devices close to each other matches the intuitive behavior that users would naturally like to perform to connect their devices [12, 61], similar to a handshake [64]. In contrast, WiFi or Bluetooth also work over longer distances and even through walls, failing to ensure proximity.

Furthermore, we use the acoustic OOB channel not only for verification but also to initiate the protocol, automatically connecting all participants’ smartphones to the coordinator’s WiFi Direct network. Both WiFi and Bluetooth require a manual pairing process before communication can occur. We automate this process using acoustic communication, making the temporary network discovery and association as part of our group pairing protocol effortless and seamless.

5.3 PairSonic Implementation

We implemented PairSonic as an Android application (app) using the Flutter framework and utilized Android’s WiFi Direct API for local ad-hoc WiFi functionality. The acoustic OOB channel was built using the ggwave library.⁸ The physical layer uses multi-frequency Frequency-Shift Keying (FSK) modulation with 96 evenly spaced frequencies ranging from 1875 Hz to 6375 Hz, according to ggwave’s AUDIBLE_FAST profile. We selected an audible frequency range to assess user perception of this form of communication. Practical applications could choose between audible or inaudible frequency ranges, based on factors such as robustness, usability, and hardware compatibility.

As our implementation is a custom prototype using a novel communication stack combining acoustic communication with WiFi Direct, we conducted functionality and compatibility tests with smartphones from various manufacturers before our lab study. Our prototype worked on all tested devices supporting WiFi Direct, with a minimum required Android version 6.0 (API level 23,

⁸Data-over-sound library ggwave: <https://github.com/ggerganov/ggwave>

Table 1. Overview of the two group pairing protocols compared in our study.

	SafeSlinger	PairSonic
User Grouping	– peer-based	– leader-based
Device Communication	– Internet via central server	– decentral via WiFi ad-hoc – local acoustic communication
Phase 1: Initialization	– each participant counts + enters group size – each participant compares ID with others and enters lowest ID	– only leader counts + enters group size – participants bring devices into proximity
Phase 2: Verification	– each participant compares three 3-word phrases with others and selects matching phrase	– each participant confirms whether all devices show green checkmark
Phase 3: Finalization	– each participant verifies and selects which contact entries to import	– each participant verifies list of new contact entries

released in 2015): LG Nexus 5X, Huawei Nexus 6P, Google Pixel 4, Pixel 4a, Pixel 5, Pixel 6 Pro, Samsung Galaxy S20 Ultra, Galaxy S22, Oppo Reno 6, OnePlus 10 Pro, Xiaomi 11T Pro. Our design is principally also compatible with iOS, given the ubiquitous compatibility of acoustic communication. However, since iOS does not support WiFi Direct, this channel would need to be replaced with another communication layer supported on Apple devices, such as BLE or AWDL [110]. Our design does not inherently depend on any specific radio channel, making such a modification feasible.

5.4 Comparison to SafeSlinger

Our design’s key differentiator is the automation of as many manual steps as possible, resulting in a group pairing protocol that is effortless for the user. Such a design allows us to verify our research hypothesis by comparing SafeSlinger with PairSonic in a user study. Table 1 lists the main differences between PairSonic and SafeSlinger, which we now discuss in more detail.

5.4.1 Usability. Whereas in SafeSlinger all participants must enter the group size in the initialization phase, in PairSonic only one participant needs to enter the group size. This smartphone then starts broadcasting initialization information to the other participants’ smartphones via the acoustic OOB channel. The other participants only need to bring their devices close together. The second part of SafeSlinger’s initialization phase, which requires all participants to coordinate to find and enter the lowest number shown on their devices, is completely automated in PairSonic using the acoustic OOB channel.

After exchanging their contact information, the participants need to verify that no adversary interfered with the protocol. Whereas SafeSlinger requires all participants to find a matching word phrase among three options, we automate this using an acoustic verification message in PairSonic. This reduces the information that participants need to verify manually to a single bit, which is represented by a clear green checkmark or a red warning symbol. The participants must abort the process in case not all devices show the green checkmark. In PairSonic, the devices automatically verify a second acoustic message sent by the initiating smartphone to determine whether it matches the data received over the WiFi channel. PairSonic minimizes the required user interaction compared to SafeSlinger, without reducing the security guarantees.

In the final phase of the protocol, the participants must verify that the exchanged contact entries correspond to the involved participants. This step cannot be automated as it is a social decision to

determine if a contact entry matches the person standing in front of them, rather than a technical one. In SafeSlinger, each participant also has to select which contact entries to import. In PairSonic, we defer this decision by importing all contact entries, considering that users can manually delete contact entries whenever they want.

5.4.2 Deployability. In addition to usability improvements, our design also addresses the deployability limitations of SafeSlinger by decentralizing the protocol. Whereas SafeSlinger requires a stable Internet connection during the pairing process, PairSonic operates completely offline. Thus, it can function even in areas without WiFi or cellular connectivity, such as rural zones, emergency situations, or developing countries without Internet access. Our protocol does not require a third-party server because it operates in a decentralized manner, i. e., directly between the participants. This results in greater availability and resilience since our pairing process does not depend on external infrastructure. It also offers a privacy advantage, as no metadata reaches any third party.

5.4.3 Security. While PairSonic fundamentally changes both the user interaction and physical layer of SafeSlinger, it retains the same cryptographic protocol, thereby preserving the strong security benefits from SafeSlinger. PairSonic thus fulfills all security requirements of confidentiality, contact authentication, and collective pairwise security. We rely on SafeSlinger’s cryptographic primitives, namely hierarchical multi-value commitments and Group Diffie-Hellman key agreement (for further information, see Farb et al. [23]). Additionally, we employ an acoustic OOB channel for location-limited verification as an extra security layer.

6 Study Methods

To answer our research questions on the usability and security of the group pairing protocols SafeSlinger and PairSonic, we conducted a comparative lab study featuring practical contact exchange tasks (Section 6.1), followed by a quantitative questionnaire and qualitative interviews (Section 6.2). We also discuss our study sample (Section 6.3), ethical considerations (Section 6.4), and our analysis methodology (Section 6.5).

6.1 Material

Our goal for the study setting was to resemble a typical contact exchange scenario. We built a dummy smartphone app of a fictional social networking platform consisting of user profiles, direct messaging functionality, and public message boards. This app was built using Flutter targeting Android smartphones and supports E2EE messaging using public-key cryptography. Each user has a profile consisting of a username and additional profile information such as an avatar and a profile description. Additionally, the app automatically generates public-key credentials for each user, which is used to secure the communication. The app’s main menu has a button to initiate the contact exchange process, triggering either SafeSlinger or PairSonic.

We integrated SafeSlinger into our app using the official Android library⁹ by Farb et al. [23], including their user interface. We only modified the look and feel to match the rest of our app. Figure 3 shows the different steps that participants have to perform using SafeSlinger. We used Google’s cloud platform AppEngine to host SafeSlinger’s server component¹⁰ written in Python.

6.2 Study Design

We followed a mixed methods approach, combining usability testing within a lab study, open-ended oral interviews, and a post-test questionnaire. We performed our lab study with different groups

⁹SafeSlinger Android client library: <https://github.com/SafeSlingerProject/exchange-android>

¹⁰SafeSlinger Python server: <https://github.com/SafeSlingerProject/SafeSlinger-AppEngine>

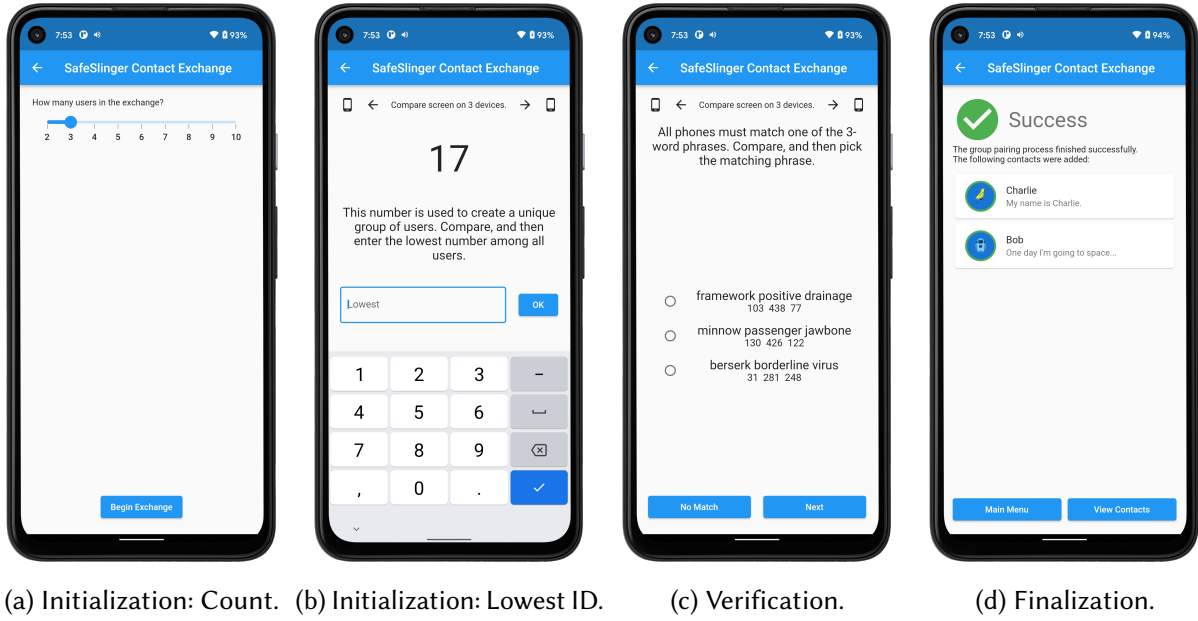


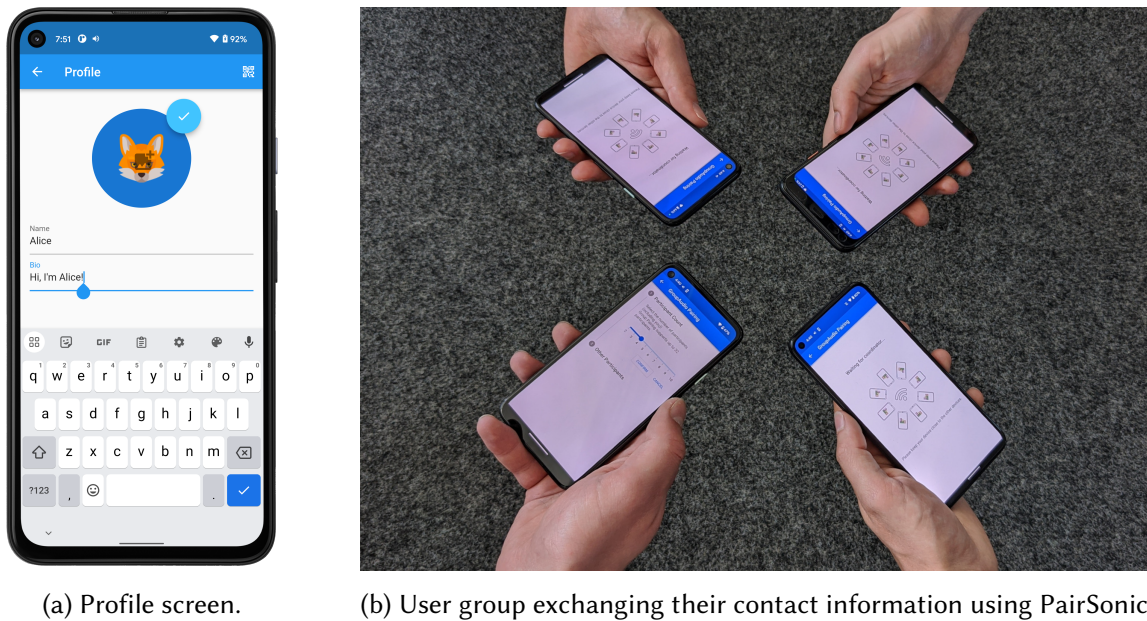
Fig. 3. This figure shows the SafeSlinger contact exchange, which we integrated into our app using the official Android library, only adjusting its aesthetics to align with our app’s design. (a) The SafeSlinger contact exchange process begins with each participant selecting the total number of group members. (b) Next, all members compare their IDs to identify and input the smallest ID. (c) During the subsequent verification phase, members compare their word phrases to find and choose the phrase that appears on all devices. (d) After this process, the app displays the exchanged contacts.

of two to six participants, who got to exchange their contact information amongst each other in an actual group scenario. We used a within-subjects design to identify the differences between PairSonic and the state of the art (SafeSlinger), meaning that each participant performed two rounds of tasks within the same group of participants to gain hands-on experience with both systems. We used counterbalancing to vary the order in which the participants encountered these systems, which reduced practice and boredom effects. We decided for a within-subjects design due to greater statistical power than a between-subjects design. We successfully tested our study setup using a pilot study with six participants.

We conducted our lab study in a neutral meeting room, where each group of participants met the study conductor, who was always present during each trial to verify that the participants performed the tasks and to conduct the interviews. The lab setting in a neutral meeting room with the same smartphones¹¹ ensured consistent conditions for all participants, minimizing confounding factors. During the study, the participants filled in our questionnaire section by section after each experiment round. The study conductor additionally interviewed them about their impressions after the experiment.

6.2.1 Briefing. Participants first reviewed and signed our consent form, which outlined the study’s aim, procedure, and privacy policy. We then explained the study’s agenda, briefly introducing the problem of securely exchanging contact information and how current approaches fail for larger groups (approximately four minutes total). We used the same slide set for all groups to increase the internal validity of our study by ensuring that all participants received the same information.

¹¹We always used the same smartphone models for the experiment, in the following order depending on the group size: 3×Google Pixel 4a, Google Pixel 5, Samsung Galaxy S22, Oppo Reno 6.



(a) Profile screen.

(b) User group exchanging their contact information using PairSonic.

Fig. 4. Our participants first created their profiles in the app, then collaboratively tested both contact exchange systems with the other group members. During PairSonic, the participants bring their devices close together.

The goal of the briefing was to ensure that all participants had the same understanding of the systems they were about to try, reducing potential confounding effects from misunderstandings. Importantly, we did not disclose that we had designed one of the systems ourselves.

We then handed each participant the questionnaire (Appendix B) and a smartphone, and asked them to fill in their profile information within the app (name, avatar from a gallery of options, freely chosen text description; shown in Figure 4a). We gave each participant a choice between using their real name or a pseudonym within the app.

6.2.2 Experiment. Upon distributing the smartphones, we conducted our experiment in two rounds (A and B) corresponding to both group pairing systems. Each round was structured identically, but we alternated the order of SafeSlinger and PairSonic between each group to neutralize order effects.

Each round started with a warm-up task, where we asked the group to try the contact exchange system on their own without further guidance except the prompts within the app. Figure 4b illustrates how PairSonic's context exchange looked like for a group of four users. The study conductor noted down potential problems in the study protocol and clarified the process with the group after they successfully exchanged their contact information.

We then asked the group to use the system a second time, but remotely and unbeknownst to the participants we manipulated the experiment in such a way as to simulate an active attack on the pairing process. This resulted in the protocol failing on one of the smartphones. The study conductor observed the group reacting to this situation and asked them why they thought the process failed. We opted to simulate an attack because this provides a controllable and reproducible method to ensure all participants have a consistent experience. By simulating rather than actively disrupting the wireless communication, we also avoided technical risks to nearby networks. Our simulated attack mirrors the maximum capabilities of our theoretical adversary, who can remotely control smartphone radio communication undetected (Section 2.3). Incorporating this attack into our study helped us evaluate the usability of PairSonic, especially in adversarial scenarios our system aims to counter.

After explaining that this protocol failure could indicate a security problem, we asked the group to use the system a third time. We remotely disabled the attack simulation and told them that this time there would be no attack. At the same time, we measured the time it took the group to perform this third task from the beginning till the end, to get an assessment of how fast the system works after some practice.

After the first round, we collected all smartphones, switched to the other group pairing protocol, and conducted the second round likewise. Before concluding the study, we thanked the participants for their time. In total, each of our participants spent 60 minutes to take part in our study.

6.2.3 Questionnaire. After each of the two rounds, participants filled in the respective sections of the questionnaire to reflect on their experiences during the experiment tasks. Once both rounds were completed, they filled in sections about their preference of both system, about their usage patterns of different social or collaborative tools, and demographic questions at the end. Our corresponding questionnaire, presented in [Appendix B](#), was used to collect the following variables:

- **Dependent variables.** The dependent variables captured our participants post-test assessment of both PairSonic and SafeSlinger. For each system, we used the System Usability Scale (SUS) [7] to answer RQ2 on usability. Because efficiency is one aspect of usability [95], the study conductor also measured the time it took the groups to complete the protocol during the third task. To answer RQ3, we asked an additional question about the perceived security of the systems, using a five-level Likert item.
- **Preference.** To answer RQ1, this binary variable captures the participants' preference: either they preferred PairSonic or SafeSlinger.
- **Collaborative/social tool usage.** To measure our participants' use of collaborative platforms and tools, we provided a list of nine distinct types of digital communication groups. For each type, we asked the participants to state whether they actively took part in them, the average number of participants in such groups, and if security was important to them in these contexts. Participants could also provide information about additional group types using free text fields. Additionally, we requested them to estimate the total number of digital communication groups they overall participate in.
- **Control variables.** We used two control variables: participants' prior experience with smartphones and their Affinity for Technology Interaction (ATI) [31]. Both variables help to assess the background of our sample and to explore potential correlations between technology affinity and the usability or security of both systems [14]. As a standardized measure, the ATI also improves the utility of our dataset for future research on usable security.
- **Demographic variables.** We collected the gender, age, education, field of study/work, and student status of our participants as demographic variables.

6.2.4 Interview. At four stages during the study, we interviewed our participants with open-ended qualitative questions to better understand their behavior and reasoning. All participants agreed to be recorded in our consent form. The study conductor also asked follow-up questions to learn more about the intricacies of their reasoning, but always asked the same set of starting questions for each group, in accordance with our research questions.

After filling in each round's questionnaire section about the prior used system, we asked participants:

- What did you like or dislike about this system?
- How secure do you think this system is?

Table 2. Our participants' demographic data and control variables. We report percentages (and frequencies).

	Variable	Our Sample		Variable	Our Sample
Gender	Female	40% (18)	School Education	Intermediary secondary	4% (2)
	Male	49% (22)		University entrance qualification	93% (42)
	Diverse	4% (2)		No answer	2% (1)
	No answer	7% (3)		Vocational training	11% (5)
Age	18–19	2% (1)	Prof. Education	Technical college	4% (2)
	20–24	49% (22)		Bachelor	22% (10)
	25–29	36% (16)		Master	22% (10)
	30–34	11% (5)		PhD	2% (1)
	40–44	2% (1)		No degree	36% (16)
Student	Yes	82% (37)			
	No	18% (8)			
Smartphone Exp.	> 2 years	98% (44)			
	≤ 2 years	2% (1)			
ATI	Median	4.6			
	IQR	1			

Once participants gave feedback to each approach individually and completed the questionnaire about their general preference, we interviewed participants to evaluate the two approaches in direct comparison:

- Which system do you prefer, and why?
- *⟨Study conductor clarifies which part of each system constitutes the initialization step.⟩* Which initialization step do you prefer, and why?
- *⟨Study conductor clarifies which part of each system constitutes the verification step.⟩* Which verification step do you prefer, and why?

After the participants filled in the questionnaire section about their usage patterns for various types of collaborative tools, we asked about security:

- How important is secure communication to you in group or collaborative scenarios?

6.3 Participants

We conducted our study with 47 participants between January and March 2023. For recruitment, we used mailing lists, social media groups, word-of-mouth, and snowball sampling, both within and outside our university. Participants had to be over 18 years old to be eligible. Interested participants self-registered on an online form to choose all suitable time slots for voluntary participation. We randomly distributed our participants into groups of two to six participants based on their availability and invited them to our lab study. This resulted in five groups with four people, two groups with each of two, three, and six people, and one group with five people, for a total of 12 unique groups. Each participant was compensated 15 EUR (in cash) for taking part in our study.

Two participants did not fully complete the questionnaire, so we removed them from our final sample ($N = 45$). Of these, 22 identified as male, 18 as female, 2 as diverse, and 3 preferred not to answer their gender. While 16 participants did not have any professional or university degree, the education level of our participants was generally high, with 10 participants having a Master's

Table 3. Comparison of the dependent variables.

Dependent Variable	SafeSlinger (<i>N</i> = 45)	PairSonic (<i>N</i> = 45)	Statistic	ES
System Usability Scale (SUS)	75 (25)	85 (17.5)	$V = 597.5$ $p = .004$	$r = -.43$
Security	4 (1)	4 (1)	$P = .5$ $p = 1$	$g = 0$
Preference	31% (14)	69% (31)	$P = .69$ $p = .016$	$g = .19$
Completion Time (Groupwise, <i>N</i> = 12)	35.5 s (8.5 s)	33.5 s (10 s)	$V = 49.5$ $p = .432$	$r = -.23$

Note: for the binary preference rating, we report percentages (and frequencies), the binomial test, and Cohen's g [17] as the effect size (ES). For the other variables, we report the median (and IQR). For the SUS and completion times, we additionally report the Wilcoxon signed-rank test, and the ES estimate based on Rosenthal's method [94]; for the security ratings, we report the sign test and Cohen's g .

degree. Most participants (37) were students. Our participants were aged 18–44; most participants (38) were between 20 and 29 years old (Table 2).

6.4 Ethical Concerns

Our university's institutional review board (IRB) reviewed and approved this study. We gathered written consent from all participants after informing them about the study's purpose and data collection in compliance with the General Data Protection Regulation (GDPR).

We did not collect any sensitive data, and we made answering demographic questions in our questionnaire optional. Participants had the voluntary choice to permit interview recordings; all agreed, so we recorded all interviews and deleted the files after transcription to safeguard participants' privacy. We stored each participant's responses pseudonymously, using sequential numbers without including any identifying information. Participants voluntarily engaged in our study and received compensation higher than our country's minimum wage as recognition for their time and effort.

6.5 Data Analysis

We used R 4.2.2 for all quantitative data analysis [92] and MaxQDA 2022 for the qualitative data coding [124]. We calculated the central tendencies and correlations to answer our research questions. We compared SafeSlinger and PairSonic using the Wilcoxon signed-rank test [128] for interval variables, the sign test [105] for ordinal variables, Pearson's chi-squared test [87] for nominal variables, and the binomial test [105] for proportions. We report descriptive statistics for our quantitative variables using the median (Mdn) and the interquartile range (IQR). We used Kendall's rank correlation coefficient [55] to determine the relationship between our control and dependent variables. For all statistical tests, we used an alpha level of .05, following the common practice in our field to balance the risk of type I and type II errors.

Furthermore, we conducted a qualitative analysis of the interview transcripts, performing the following qualitative coding steps: (1) We first deductively created an initial codebook covering all our research questions (Section 4). (2) Two researchers then independently coded the transcripts from the study session of a single group, to find out how well the codebook fits the participants' statements. During this separate coding process, they inductively extended the codebook with new codes to match interesting observations outside the research questions [79]. (3) The researchers

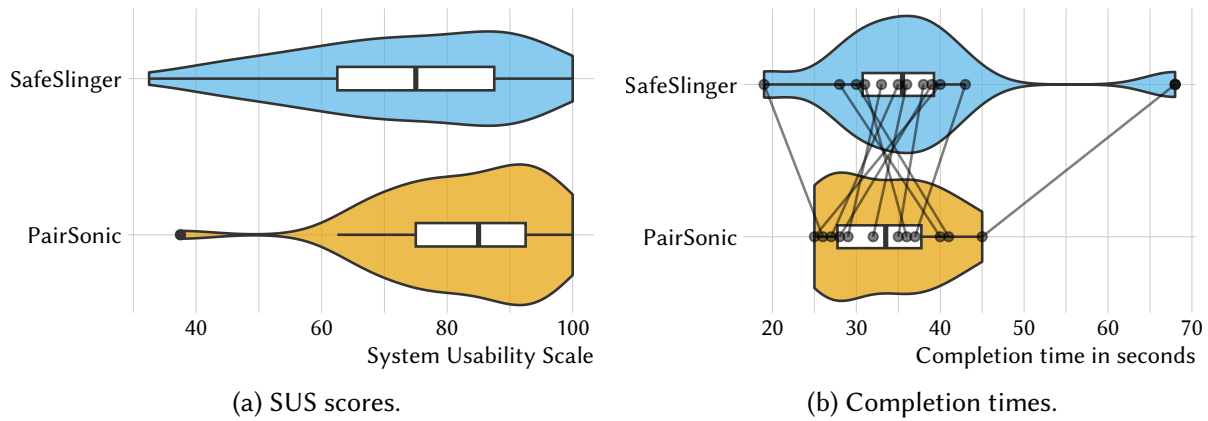


Fig. 5. Comparison of our participants' SUS scores and completion times for SafeSlinger and PairSonic. The violin plots show a density estimation of the distributions. The boxplots show quartiles, median, and outliers.

then discussed and merged their two codebooks into the final codebook: preference, mental model, scenarios, suggestions/improvements, security, relevance of security in groups, initialization step, verification step, Internet requirement, third party involved, user interface, and user roles. (4) The two researchers then split the remaining transcripts and deductively coded them according to the final codebook. (5) Finally, the researchers swapped the transcripts to review and extend the other researcher's codes, and to discuss and resolve coding differences.

7 Quantitative Results

This section presents our quantitative results and the analysis of participant task performance (Table 3). The questionnaire includes quantitative scales, numeric fields, and control variables (Section 6.2.3). We studied the participants' perceptions of secure contact exchange methods regarding their usability (Section 7.1), security (Section 7.2), and preference (Section 7.3). We further examined their task completion times (Section 7.4) and general usage patterns for social and collaborative tools (Section 7.5). An analysis of the control variables is also provided (Section 7.6).

7.1 Usability

After getting hands-on experience with both group pairing protocols, the participants completed our questionnaire, including the SUS [7]. We use the Shapiro-Wilk test [101] to determine whether this interval variable is normally distributed. The SUS scores for SafeSlinger ($W = 0.96$, $p = .095$) are approximately normal, but the scores for PairSonic ($W = 0.92$, $p = .003$) are significantly non-normal distributed. Figure 5a depicts the statistics of the SUS scores per protocol. Both distributions have visible negative skew, indicating that most participants selected values on the higher end of the scale. Additionally, there is a single outlier for PairSonic. Parametric methods such as the t-test assume normal sampling distributions and can give inaccurate results in the presence of outliers [127]. We, therefore, use non-parametric statistical methods for data analysis (Section 6.5).

SUS scores for SafeSlinger (Mdn = 75) differed significantly from PairSonic (Mdn = 85) according to the Wilcoxon signed-rank test, $V = 597.5$, $p = .004$. The effect size is $r = -.43$, which corresponds to a large difference according to Cohen [17]. Thus, **PairSonic showed significantly better usability than SafeSlinger.**

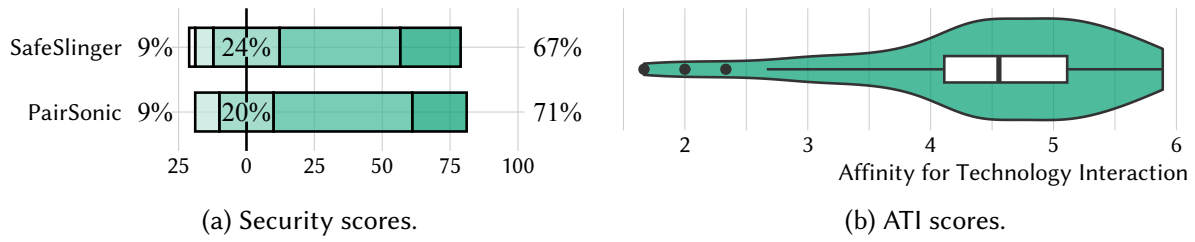


Fig. 6. Comparison of our participants' security and ATI scores for SafeSlinger and PairSonic. The stacked bars for the security scores correspond to the five levels of agreement, ranging from "strongly disagree" (left) to "strongly agree" (right), centered at the neutral response. The percentages (left, middle, right) represent the share of negative, neutral, and positive responses, respectively. The violin plot shows a density estimation of the ATI distribution. The boxplot shows quartiles, median, and outliers.

7.2 Security

Based on their experience during the lab tasks, we asked our participants how secure they think both systems are on a five-level Likert item. Figure 6a shows the security scores, which are ordinal assessments to the statement "I think that this system is secure", ranging from "strongly disagree" to "strongly agree". According to the sign test, which is appropriate for ordinal values [105], the security scores for SafeSlinger (Mdn = "agree") did not differ significantly from PairSonic (Mdn = "agree"), $p = 1$, $g = 0$. Hence, **users have a similar security impression of both methods.**

7.3 Preference

After our participants tried both systems, we asked them which one they liked better. Figure 9 in the appendix shows how many participants preferred PairSonic in each study group. Most participants (69%; 31) preferred PairSonic compared to the state of the art (SafeSlinger), which is a significant difference according to the binomial test, $p = .016$. The effect size is $g = .19$, which corresponds to a medium difference according to Cohen [17]. Overall, **participants prefer PairSonic compared to SafeSlinger.**

7.4 Completion Time

As an objective measure, we determined the time it took our participants to complete each contact exchange protocol, after a brief learning period during the initial two tasks. Figure 5b shows the completion times of each individual group, connected for both protocols, to show the change within each group. The completion times for PairSonic are quite consistent and approximately normal distributed ($W = .94$, $p = .528$). There is an outlier for SafeSlinger as it took one group 68 s to complete the protocol, making the SafeSlinger completion times significantly non-normal ($W = .85$, $p = .035$). This reaffirms our choice for the robust non-parametric Wilcoxon signed-rank test. Most groups (8; 67%) were faster with PairSonic, but the completion times between SafeSlinger (Mdn = 35.5 s) and PairSonic (Mdn = 33.5 s) did not differ significantly ($V = 49.5$, $p = .432$, $r = -.23$). The completion times for both schemes showed no significant correlation with the group size ($\tau = .46$, $p = .052$ in both cases). In summary, **the completion times are similar for PairSonic and SafeSlinger.**

7.5 Usage of Social and Collaborative Tools

We were also interested in understanding which types of social and collaborative tools our participants use, as this would reveal the most prominent use cases. Figure 7 summarizes our participants' responses. All participants use private chat groups, such as WhatsApp or Signal, as well as audio

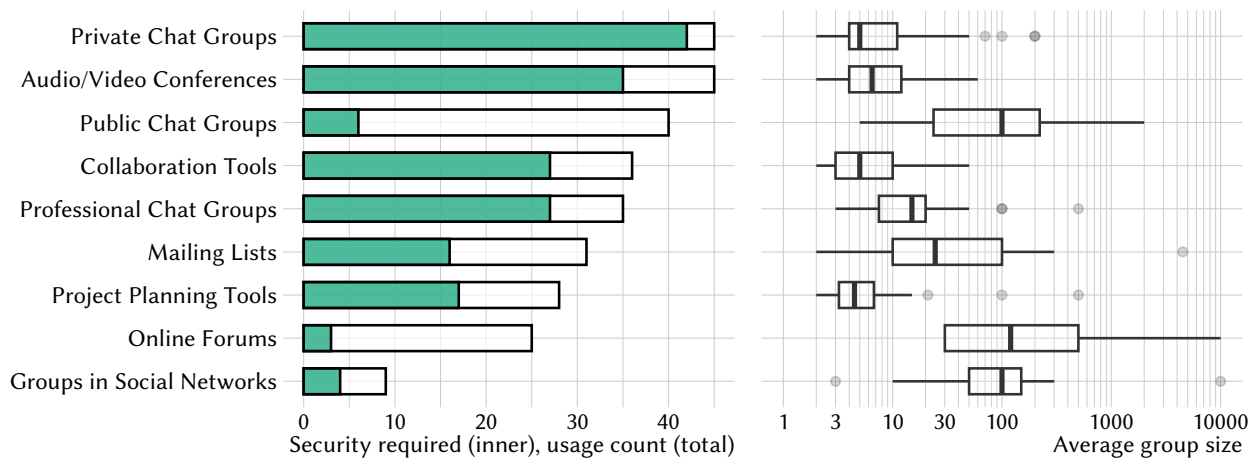


Fig. 7. The left-hand bar chart illustrates the number of participants using the given social or collaborative tools (total bar length), with the inner green bar representing those who deem security and authentication important. The right-hand boxplots depict the distribution of the average number of group members indicated by our participants for each tool type. They show quartiles, median, and outliers on a logarithmic axis. One outlier is not included: one participant gave an average online forum group size of 1.6 million.

or video conferencing tools like Zoom. Additionally, a substantial portion partakes in public chat groups (89%), online collaboration tools like Google Docs (80%), and professional chat groups such as Microsoft Teams (78%), showing the relevance of digital collaboration applications.

We also queried our participants about their security desires for these tools, i. e., ensuring no unauthorized access. Tools requiring such security measures could benefit from group pairing protocols, such as PairSonic, indicating potential use cases.

Figure 7 demonstrates that participants seek authentication for most tools, especially private chat groups, audio/video conferencing, online collaboration tools, and professional chat groups.

Futhermore, we asked our participants to estimate the average number of group members for each type of collaboration. Groups demanding security requirements were typically smaller: the group sizes for private chat groups (Mdn = 5), online collaboration tools (Mdn = 5), and audio or video conferences (Mdn = 6) were much smaller compared to public chat groups (Mdn = 100) and online forums (Mdn = 135). Our participants are active in a significant number of groups, with a median of 70 groups (IQR = 70).

7.6 Control Variables

As a control variable, we measured our participant's ATI (Mdn = 4.6, IQR = 1) and smartphone familiarity. All 45 participants except one have been using a smartphone for more than two years, indicating general familiarity with smartphone usage.

We calculate bivariate Kendall's rank correlation coefficients in Table 4 to determine the relationship between the control and dependent variables. As shown, our participants' SUS scores, security ratings, and preferences do not significantly depend on their technology affinity or smartphone familiarity. The SUS scores for SafeSlinger significantly correlated with our participant's preference, indicating that participants who found SafeSlinger less usable preferred PairSonic instead. There was also a significant negative correlation between SafeSlinger SUS scores and completion times. Additionally, we found that a method's SUS score significantly correlated with that method's security assessment.

We noticed that the order in which our participants encountered the systems significantly correlates with the SUS scores for SafeSlinger (Figure 11), but not with the SUS scores for PairSonic or with the other variables. We discuss this potential order effect in Section 10.2.

8 Qualitative Results

Our previous section’s quantitative evaluation indicated that participants found PairSonic to be more usable and preferred it over SafeSlinger. In this section, we complement the quantitative analysis with a qualitative analysis: we conducted post-test interviews with the participants to understand the reasons behind their preference. We begin by determining the factors contributing to their overall preference (Section 8.1). Subsequently, we closely compare user perceptions regarding the two main phases of group pairing protocols: the initialization phase (Section 8.2) and the verification phase (Section 8.3). Lastly, we capture our participants’ perspectives on acoustic communication (Section 8.4).

8.1 Which Method Do Users Prefer Overall?

During the interviews, our participants’ comments aligned with their preference for PairSonic overall. They specifically highlighted two key advantages: better *usability* and *scalability*:

“If you have used [PairSonic] once or twice, then it’s also just much, much more convenient than [SafeSlinger]. I mean, you open the app, say »here, you’re the person who sets this up«, hold the phones together, done.” (Group 7)

However, some participants argued in favor of SafeSlinger, especially for *smaller groups* where the manual comparison tasks require less effort.

Regarding *security*, the qualitative results were mixed, supporting the results of our quantitative analysis (Section 7.2). Some participants expressed that SafeSlinger’s additional user interaction and complexity instilled a sense of enhanced security compared to the simpler PairSonic. Conversely, other participants favored the security of PairSonic, as it did not solely rely on potentially error-prone user interaction but also incorporated acoustic verification as an additional layer of security.

Most participants raised concerns about the *Internet requirement* of SafeSlinger, citing instances where they lacked cellular connectivity or access to public WiFi. They worried that this could limit the situations in which they can exchange their contact data. Additionally, a few participants expressed privacy concerns since the contact data was transmitted over the Internet. Despite SafeSlinger being designed to safeguard the confidentiality of exchanged information from third parties [23], the fact that the protocol relied on Internet connectivity was enough to make some participants worry about their data. In contrast, the decentralized nature of PairSonic emerged as a distinct privacy and availability advantage, as no third parties are required to facilitate the protocol.

8.2 Which Initialization Step Do Users Prefer?

The initialization step of SafeSlinger and PairSonic temporarily associates the group members’ devices to exchange contact information and cryptographic public keys. However, there are distinct approaches employed by each protocol. In SafeSlinger, a peer-based approach is used, requiring each group member to input the group size and determine the lowest identification number among all members. On the other hand, PairSonic utilizes a coordinator-based approach, where only one group member needs to input the group size, and the rest of the initialization is automated through acoustic communication. It is important to note that the role of the group leader in PairSonic does not confer any special privileges and is only relevant to the contact exchange without implications for subsequent communication.

During the interviews, we asked participants to compare the two initialization steps. The majority expressed a preference for our coordinator-based initialization, although some participants argued in favor of the peer-based initialization. This observation is intriguing, particularly when considering a previous user study by Nithyanand et al. [85], which suggested a greater overall acceptance of peer-based methods. This discrepancy invites a deeper analysis to understand the reasoning behind our participants' preferences:

8.2.1 Effort. Some participants did not like the additional effort in SafeSlinger to determine the smallest identification number:

"It is annoying that you have to find out the smallest numbers first. Sure, you can do that relatively efficiently, but if you are in a group with a bunch of people, then maybe people just call in their numbers and then nobody has an overview anymore." (Group 3)

8.2.2 Scalability. Other participants highlighted a scalability issue regarding the effort required, particularly as group size increases:

"If it is a larger group, it is more practical to have a coordinator." (Group 12)

Participants noted that in SafeSlinger, all group members are required to enter the group size, which can potentially lead to failures if one member miscounts.

8.2.3 Determining the Coordinator. Although no group in our study encountered difficulties in appointing a coordinator during the protocol, some participants expressed concerns about this aspect during the interviews. Even though SafeSlinger's peer-based approach requires more user interaction for each participant, it does not require the explicit social decision of determining the coordinator [49]. Interestingly, some participants preferred this trade-off:

"I think [SafeSlinger's initialization step] would probably be better, because it leads to fewer disputes about who is now the coordinator. Although, of course, that doesn't matter at all in the end. But in different social groups this can lead to disputes." (Group 11)

Other participants argued that there often already is someone who proposes to exchange contact information in the first place, so there would likely be few disputes about the coordinator in practice:

"But if I think about it now, typically someone comes up with the idea of exchanging the contact data. That means that in the group there will already be some spokesperson who has crystallized, even if only for this specific situation. That means that, okay, I'll say here, I'm the coordinator, who else wants to participate and wants to exchange contacts with me, let's do that. And then I know that if four or five people approach me, okay, we are six people, I can set it up and off we go. That is, it actually fits relatively organically into the group dynamic, as I would expect in such a situation." (Group 4)

8.3 Which Verification Step Do Users Prefer?

In SafeSlinger, all group members are required to compare and select a matching three-word phrase, whereas PairSonic mostly automates the verification, with group members only needing to verify that all devices display a green checkmark symbol. We now compare participants' perceptions of the verification step in the group pairing protocols.

8.3.1 Acoustic Verification Is Effortless. Most participants preferred PairSonic, arguing that it is easier to use:

“For the verification at the end, I think I prefer [PairSonic], because all you have to do is just make sure there is the green check. So even at that, there’s no need to communicate and just raise your phone and show the green check to everyone around. And it’s faster that way.” (Group 10)

8.3.2 Word Phrases Are Prone to Errors. Some users felt nervous comparing the word phrases and worried about making mistakes. Others liked the thorough verification in SafeSlinger, as it made the process feel more secure:

“So I think, because I have to look again carefully, that [SafeSlinger] is really secure.” (Group 9)

8.3.3 Rushing Users. However, many participants mentioned that they did not really check whether all participants have the same word phrases, and only roughly compared the words:

“We had the situation where only one person didn’t match, but the others did. It’s tempting to say right after the first match, okay, yes, it matches, I’ll tick the box, great, done. And the other person is then either passed over or I don’t know what happens.” (Group 4)

Some rushed the comparison due to the pressure of not wanting to keep the group waiting, particularly in the group scenario:

“The reason why you don’t take the time to compare accurately is that you just hurry. In a larger constellation, you don’t want to be the person because of whom the procedure takes longer, that you need more time to go through the individual words.” (Group 11)

Some participants also rushed with PairSonic, mistakenly assuming that the green checkmark symbol indicated a successful completion of the protocol, and thus confirmed the dialog without reading the instructions. In the subsequent discussion, the group proposed using a different symbol or adding a confirmation dialog to mitigate this issue.

8.3.4 Accessibility. Some participants mentioned accessibility problems with the word phrases due to non-native speakers, foreign accents, or use of uncommon words.

“The word phrases are basically nonsense in English. That is, it depends on how well it is pronounced, how well it is understood, which again becomes difficult in larger groups. And it also becomes more and more difficult depending on the countries of origin. Red and green as signal colors are actually quite international and easy to distinguish.” (Group 4)

8.4 What Do Users Think of Acoustic Communication?

As most users lack prior experience with acoustic communication, our study aimed to gauge their reactions to this new technology. Our interviews revealed a generally positive sentiment; many participants found its novelty appealing and enjoyed their devices producing melodic sounds. Past research suggests that introducing fun elements, such as gamification, can encourage users to adopt secure behaviors [44].

“That you can actually hear the sound is very funny. And it also somehow gives a feeling of, yes, I can hear what’s happening. It feels nice to use.” (Group 6)

There are mainly two concerns that our participants had with acoustic communication, specifically related to its security and robustness.

8.4.1 Security. Some participants felt the audible acoustic verification enhanced safety, while others doubted its security, fearing potential eavesdroppers. It is important to note, however, that we designed PairSonic to maintain security even if acoustic messages are intercepted.

Some technically minded participants questioned the system’s security, noting that the acoustic messages always sounded identical:

“I think intuitively I would find [PairSonic] a little bit more insecure because you see how they communicate. Or you hear it in this case. If you don’t know how it works, what they’re doing exactly, I might think, yes, if the smartphones are talking there, then maybe someone can listen in. Or record it.” (Group 12)

Despite the acoustic messages sounding similar, they always encode varying data based on the group participants. Better in-app explanations can address both of these security concerns. Alternatively, a different modulation scheme can be used to produce audibly distinct melodies.

8.4.2 Robustness and Noise Interference. During the study, unexpected device interactions led to occasional protocol restarts due to failed acoustic transmissions – something we hadn’t encountered in our pre-study testing. These failures typically occurred when users placed their devices on the table during acoustic transmission, causing impulsive noise that complicated message reception.

Participants also voiced concerns about the system’s performance in loud environments, such as parties or festivals. Our proof-of-concept implementation is not yet optimized to handle these challenging acoustic environments. Some participants criticized the inconsistent reliability of audible acoustic communication, expressing a preference for the more reliable SafeSlinger system. During our study, SafeSlinger’s performance remained mostly consistent, but this requires a stable Internet connection, which is not available in many scenarios (Section 5.4.2). We derive recommendations for future work to improve PairSonic’s usability by addressing this criticism (Section 9.10).

8.5 Alignment with Questionnaire Responses

We cross-validated whether the interview statements of our participants align with their questionnaire responses in our mixed-methods study and found general agreement between the two data sources. In the questionnaire, two groups favored SafeSlinger, two were equally split, whereas the majority (8 of 12) preferred PairSonic (Figure 9), aligning with their interview statements. For instance, group 11 unanimously preferred PairSonic in the questionnaire, also rating it higher in SUS scores. During the interviews, every member of group 11 consistently expressed a preference for PairSonic, arguing for its better usability and scalability.

9 Discussion

We now discuss the main findings of our study. First, we found users favoring PairSonic for its simplicity (Section 9.1), particularly valuing the reduced effort by automating SafeSlinger’s manual verification (Section 9.2). Interestingly, though, some users perceived the more complex system as more secure (Section 9.3). We address the associated issue that complex systems tend to be more error-prone (Section 9.4) and investigate how increased transparency and education may enhance

perceived security, even for simpler systems (Section 9.5). We discuss the applicability of PairSonic for different usage scenarios within the CSCW community based on our participants' comments (Section 9.6), alongside a brief security analysis (Section 9.7). We then discuss the scalability of group pairing protocols and compare which actions users have to perform in PairSonic and SafeSlinger (Section 9.8). This section concludes by addressing PairSonic's requirement that the users have to physically meet (Section 9.9), and proposing directions for future research (Section 9.10).

9.1 Users Generally Prefer Simple and Effortless Group Pairing Systems

We designed PairSonic to require less user effort than SafeSlinger, in line with our research hypothesis that users would prefer a simpler, more effortless system (Section 4). Our lab study supports this hypothesis, as participants significantly favored PairSonic over SafeSlinger, giving it higher SUS scores (Section 7). When we interviewed our participants, most explained that they prefer PairSonic because it is easier to use and requires less effort. They especially liked the automated verification step, which only requires a simple binary comparison. Furthermore, the group-wise user interaction of moving the devices closer for pairing worked very well, confirming previous research on intuitive device association gestures [14, 48, 50, 61].

9.2 Automating Verification Tasks Using Acoustic Communication

Previous user studies have shown that users find it difficult to perform authentication ceremonies with current E2EE tools like Signal or WhatsApp, because they are time-consuming and require much effort [44, 122]. Our system, PairSonic, improves the state of the art because it automates manual verification tasks such as comparing phrases, digits, or pictures. Users responded favorably to this streamlined approach in our interviews, with most participants being able to exchange contact information in less than 35 s.

Our system utilizes an acoustic OOB channel for automated verification, which offers several unique advantages over other wireless channels that modern smartphones support: unlike WiFi or Bluetooth, it is limited to close proximity. This enhances usability by reinforcing the intuitive gesture of bringing devices closer together to share contacts, and also improves security by making distant attacks more challenging. In comparison to NFC or QR codes, the acoustic OOB channel can easily accommodate multiple devices – a necessary feature for group scenarios.

Moreover, the acoustic OOB channel is a software-defined physical layer, unlike WiFi or Bluetooth whose physical layers cannot be customized in smartphones. This adaptability enables advanced physical layer security techniques, which can protect the group pairing protocol against sophisticated attacks, adding an extra layer of security [89]. This reduces the need for meticulous user behavior during verification. By safeguarding the acoustic channel at the physical layer, we can further enhance both security and usability of our system, while meeting all our deployability requirements (Section 2).

9.3 Complex and Effortful Systems Can Feel More Secure

While most participants favored PairSonic, our quantitative analysis revealed no significant difference in how users perceive the security of PairSonic and SafeSlinger (Section 7). In fact, the interviews even indicate that some users have less trust in PairSonic due to its simplicity. They felt that the added complexity of SafeSlinger made the system seem more secure:

"I think the perceived security with [SafeSlinger] was higher because, well it sounds stupid, but because you enter so many things. [...] But I thought, because it took much longer, that it's probably more secure." (Group 4)

Our research hypothesis emphasizes minimal user interaction, but our interviews revealed a potential issue for security systems like PairSonic. Its effortless nature can make it seem too simple to be secure, almost like magic, leading to distrust. When a system automatically handles security, users may feel a loss of control, viewing the system as a black box that they can't understand [68, 96, 97]. They might also equate complexity with security due to everyday experiences like complex password requirements or cumbersome multi-factor authentication [54].

This unexpected finding emphasizes a dilemma: a system needs to be simple and effortless for usability, but if it's too simple and requires minimal effort, users may doubt its security. Our results highlight the challenging trade-off that security engineers must balance – not just optimizing usability, deployability, and security, but also the *perceived* security.

9.4 Error-Prone Collaborative User Interaction

User interaction requiring substantial effort often leads to errors. This poses a particular problem for systems like SafeSlinger, which depend on meticulous user behavior for security, such as the verification step involving comparison of word phrases. We noticed in our study that many participants did not carefully compare SafeSlinger's word phrases on each device – often checking only the first word, a behavior referred to as “*rushing users*”. This tendency has also been observed in a previous user study by Tan et al. [111].

As another example, in some larger groups, our participants only compared SafeSlinger's words with adjacent group members and accepted the matching word phrase without confirming this with the rest of the group. In other groups, a single member took charge of coordinating the verification step. However, once a few members confirmed a match, the group often selected that phrase based on majority agreement, potentially neglecting quieter members who did not share the same phrase. Differently to SafeSlinger, the security of PairSonic does not heavily depend on user diligence, thus offering systematic security advantages against active adversaries.

In contrast to the traditional pairwise authentication schemes described in Section 3, in our collaborative situation the whole group works together to achieve a common security goal. We made the interesting observation that some users checked and helped each other, leading to fewer overall failures. This phenomenon was also observed in previous user studies in collaborative scenarios [48, 54].

However, Kainda et al. found that the primary source of failure was poor communication among group members [54]. We noted similar issues in some groups, particularly where participants were more reserved and hesitant to voice problems or discrepancies in word phrases. These findings highlight that a system requiring less user interaction and communication results in a more reliable and less error-prone solution.

9.5 Transparency and Comprehensibility Can Improve Perceived Security

Technical users typically like to understand how and why a system works, whereas non-technical users need convincing evidence that a system is secure, without necessarily needing to understand all technical details [25]. Our interviews indicated that most users were not familiar with acoustic communication, especially if it is audible, and would benefit from better explanations within the app. Systems like SafeSlinger, which depend on third parties over the Internet, could also enhance user experience with increased transparency and education regarding data access, addressing users' privacy concerns. CSCW research during the COVID-19 pandemic has shown that privacy-related trust in the application provider plays an important role in the user's willingness to use online collaborative tools [82].

9.6 Usage Scenarios

There are a number of collaboration scenarios that have been proposed by CSCW researchers, where we argue that PairSonic could improve the usability of the group formation process. Previous research focused mainly on remote collaboration using tools like groupwork platforms [82], crowd-sourcing [43], and mobile devices [131]. Additionally, there are separate lines of work studying computer-supported collaboration for creative group work and education [10], for networking and meeting new collaborators [93, 116, 126], and within the context of spontaneous ad-hoc interaction [18, 22]. A common challenge in these studies is connecting people and their devices, often using technologies like QR codes that struggle with usability and scalability in larger groups (Section 1.1). PairSonic offers a scalable, user-friendly drop-in replacement to these legacy pairing technologies, providing a more efficient solution for associating devices in group settings.

There are also usage scenarios potentially involving larger groups of participants. For example, Tolmie et al. [114] and Fosh et al. [30] explored collaborative interactions in cultural sites such as museums, suggesting that connecting visitors, possibly in ad-hoc groups with a tour guide, enhances engagement with both the group and the exhibits. PairSonic can facilitate this by connecting visitors through their own devices, avoiding the need for museum rental devices. Our testing confirms PairSonic’s effectiveness for larger groups, including those with more than 10 participants, when their devices are close and in low-noise environments. Museums are ideal for PairSonic, especially using the inaudible frequency range instead of audible acoustic signals (Section 5.3). The quiet setting ensures reliable acoustic communication, and its inaudibility ensures it does not disturb other visitors. Additionally, like QR codes, the acoustic channel can transmit small data packets, which could further improve engagement with exhibitions. For example, exhibitions could emit an audio beacon offering additional information to nearby visitors, benefiting from the deployability advantages of the acoustic channel compared to alternatives such as Bluetooth or WiFi (Section 5.4.2).

Our study participants identified several scenarios where PairSonic could serve as a secure method for exchanging contacts, such as private Signal chat groups, Zoom video conferences, and online collaboration tools (Section 7.5). These situations usually have higher security requirements, often relating to personal, familial, or professional collaborations. Many interviewees shared their preferences for online security in social contexts:

“For me, it depends on how many people are in the group and who is in the group. If it’s a group where I’m only with friends, where maybe more sensitive data is discussed, then [security] is the most important requirement of all, I would say.” (Group 12)

Others felt knowing the other participants is essential for effective online collaboration:

“I find it very awkward to write in groups where I don’t know who else can read it. That’s why I usually just don’t participate at all.” (Group 6)

Our results align with previous CSCW research on the security and privacy needs of users, which found that especially vulnerable types of users and users discussing sensitive topics have a substantial need for interpersonal trust and knowing the other participants [104]. Similarly, our participants also named specific scenarios where they saw the potential for securely exchanging contact information using PairSonic. These included festivals, exhibitions, conventions, conferences, but also everyday situations when meeting new people, like meeting new people through friends. Some expressed interest in using PairSonic with non-technical users, anticipating potential difficulties these users might face with SafeSlinger:

“You can also use [PairSonic] with people who normally would need further assistance. Instead it’s enough if one person is the coordinator, and then the others just have to click yes or no. I think that’s actually quite good, because it takes the hurdle out for people who are not so tech-savvy, if you tell them that they don’t have to do anything except click yes or no once.” (Group 6)

9.7 PairSonic Inherits SafeSlinger’s Strong Security

While our primary focus is on the usability of PairSonic, and a comprehensive security analysis is not within this work’s scope, we do provide a brief discussion on its security aspects. We adhered to the security best practice of not unnecessarily creating an entirely new cryptographic protocol from scratch, opting instead to rely on the established security guarantees of the SafeSlinger protocol whenever possible. Consequently, PairSonic protects against all types of attacks mentioned in the SafeSlinger paper, such as threats from malicious bystanders, impersonation, sybil/hidden node attacks [21, 54], and Group-in-the-Middle attacks [64].

Unlike the SafeSlinger protocol, PairSonic automates information exchange stages that previously required manual effort, using the acoustic out-of-band channel for this purpose. The acoustic channel must have the same security properties as the human-mediated channel in order to not degrade PairSonic’s security. Our adversary model assumes the acoustic channel’s authenticity, implying that adversaries cannot tamper with it. This assumption is supported by previous studies on wireless authentication, which highlight the physical constraints of sound waves, such as their significantly shorter range compared to radio communication and their inability to penetrate physical barriers like walls [5, 107, 108]. Furthermore, the authenticity of the audio channel can be explicitly ensured using physical layer security techniques [89], whose assumptions and system model align with PairSonic’s protocol. Future work could incorporate them to enhance or replace the current gwave physical layer used in our protocol.

In our adversary model, we assume that an adversary can eavesdrop on the audio channel, enabling them to gain information about the ad-hoc WiFi network and potentially interfere with it. However, any WiFi interference attempt by the adversary to alter the group’s contact information would be futile, as PairSonic is designed to detect such modifications and terminate the protocol during the acoustic verification phase. The use of the acoustic out-of-band channel in PairSonic also enhances security by supporting longer hash values than SafeSlinger, reducing the likelihood of hash collisions. The SafeSlinger protocol employs 24-bit hash values as Short Authentication Strings. Increasing the entropy of these hashes would negatively impact usability, as it would require users to manually compare a greater number of words or more complex phrases. In contrast, PairSonic streamlines this process by automatically transmitting the hash value via the acoustic channel. This method allows for longer hash values without user involvement and is less susceptible to security-critical user errors, a concern highlighted in our study (Section 9.4).

9.8 Scalability

Besides usability, one motivation for PairSonic was improved scalability. PairSonic requires less interaction between participants by design, making it better suited for larger groups of users than SafeSlinger. To illustrate this, Figure 8 shows timelines of PairSonic compared to SafeSlinger for a typical user group during our study. From a user’s perspective, each protocol consists of phases where they have to wait, but also phases requiring them to read, choose an option, compare information, or enter a number: these are phases requiring manual actions. The total cumulative duration that users have to diligently perform these actions (i.e., excluding waiting times) is lowest

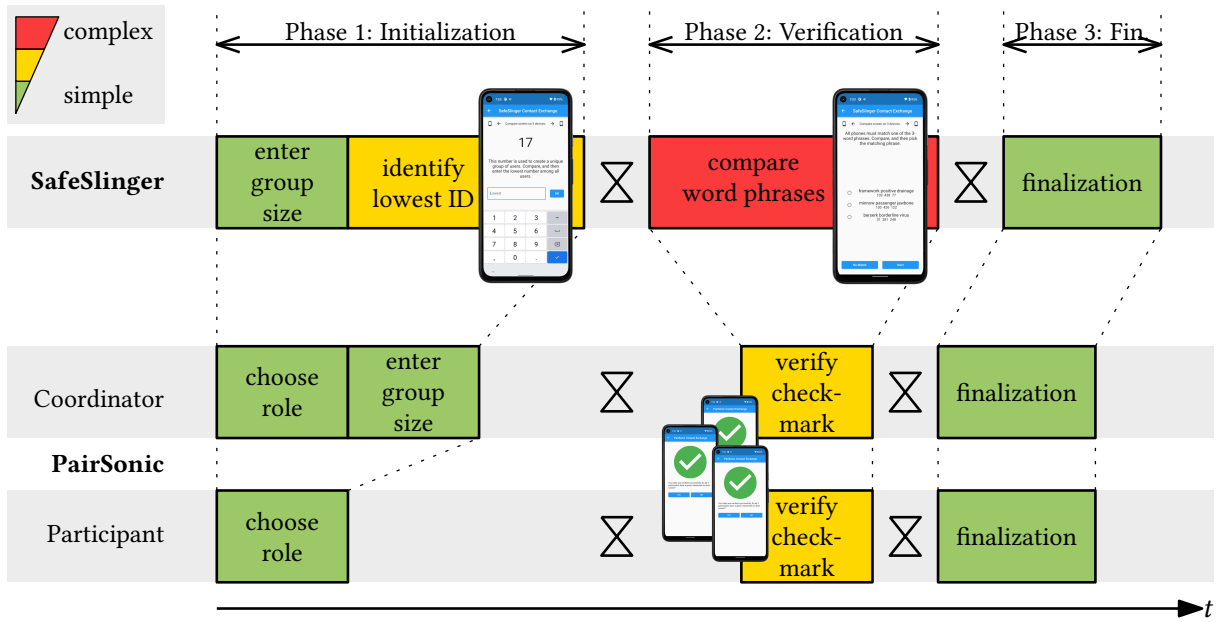


Fig. 8. Qualitative timelines show which actions users have to perform during all three phases of the contact exchange process (Table 1). The symbol $\otimes$ indicates waiting time. The yellow and particularly the red user actions are more complex than the green actions and take more time for larger groups, showing that PairSonic scales better than SafeSlinger.

for PairSonic with the participant role, followed by the coordinator role, with SafeSlinger requiring the longest attentive time for all users.

The figure also highlights the protocol phases that take noticeably more time for larger groups. SafeSlinger contains two such phases: identifying and entering the lowest identification number amongst all participants, and the word phrase comparison. PairSonic, however, only has a single phase scaling with the group size, namely determining that every device shows a large checkmark symbol, a task that is comparatively much simpler than the tasks in SafeSlinger, since it only requires a quick glance at the smartphones. In summary, PairSonic requires less attentive user duration, making it well-suited also for larger groups.

Larger groups further reinforce the issue of error-prone collaborative user interaction that we discussed previously (Section 9.4). With SafeSlinger, larger group sizes involve more manual text comparisons, which not only raises the likelihood of mistakes, but also reinforces the problem of *rushing users*. We observed that some users felt pressured by larger groups and wanted to avoid keeping the group waiting (Section 8.3.3). PairSonic was less affected by group size, but our proof-of-concept implementation struggled with acoustic noise interference (Section 8.4), which might occur more often with more participants.

Contrary to our expectation, we did not find a significant difference in completion times between both protocols, even though most groups in our study were faster with PairSonic. The completion times of SafeSlinger were faster than expected and can be attributed to many participants not thoroughly comparing the word phrases on each device, as explained earlier (Section 9.4). While this explains the comparable completion times, it degrades SafeSlinger's security in practice, as partial comparisons involve less entropy [111]. We also note that the scalability advantage of PairSonic is likely more pronounced for participants who take longer with SafeSlinger's word comparison, such as older users, users with bad eyesight, or shy participants.

9.9 Physical Meetings

PairSonic assumes that all users are co-located, making it suitable for offline and online collaboration, provided the users meet in-person first (Section 9.6). This limitation is by design and arises from our requirement that the system functions without relying on any central infrastructure and without necessitating any pre-established security context (Section 2.1). While SafeSlinger can operate remotely over the Internet, it violates these requirements by assuming the existence of another trusted communication channel to authenticate each other. However, such a channel typically does not exist yet, which is precisely why the group pairing process is needed to begin with.

This requirement for physical meetings means that to securely collaborate with a new contact using PairSonic, you must first meet them in person. In larger collaborative groups, it can often be impractical to meet with every new member as they join the group. However, PairSonic should be seen as just one component of a broader key exchange process, which cannot address the whole key management lifecycle by itself [6]. To address this limitation, PairSonic could be integrated into a more comprehensive key management system like the decentralized web of trust (e.g., OpenPGP [27]). This integration could replace the traditional manual verification of PGP keys in person or as part of keysigning parties [9, 115].

In such a system, transitive trust can be utilized, meaning that you might choose to trust the contact list of someone you have already verified, rather than verifying every new contact yourself. As a result, a new group member would only need to meet and exchange keys with one existing group member. This member then signs and forwards the authenticated contact information to the rest of the group. In this scenario, the initial group members use PairSonic to establish an initial security context, which then serves as a base for other key management protocols that do not require physical interaction.¹²

9.10 Future Work

Our study identified the primary weakness of PairSonic as the limited robustness of the acoustic OOB channel in noisy environments, reducing the potential use cases for contact exchange. Whereas our paper focused on the usability evaluation of PairSonic's prototype, the interviews have shown that future work is needed to improve the OOB channel's reliability concerning realistic noise sources, exploring both audible and inaudible frequency ranges. Additional future work should include a comprehensive evaluation of the acoustic channel's functionality and performance for various environments, to determine the practical limits of this novel technology. The acoustic physical layer in the user study by Mehrabi et al. [78] was more reliable, suggesting that a better design can address this issue.

Additionally, enhancements to the user interface could further increase PairSonic's usability, reflecting findings from previous studies on other authentication ceremonies [123, 129]. During our study, we observed that some participants rushed the verification step in PairSonic. This was not to bypass a tedious task, as seen in previous approaches, but because they misconstrued the green checkmark symbol as a sign of successful completion (Section 8.3). Using a different symbol, such as a question mark, could signal the need for user involvement in security verification, potentially resolving this issue. Alternatively, a confirmation dialog could be introduced following the initial prompt to avoid unintentional acceptance. Future research could validate these enhancements by

¹²Similar established strategies could be employed to deal with other aspects of the key management lifecycle, such as the loss or change of keys. For example, OpenPGP typically manages key replacement by generating a new keypair, signing it with the old key, and then issuing a revocation certificate for the old key. This process ensures a seamless transition from the old to the new key while maintaining security and trust within the system.

complementing our lab study with an additional field study analyzing the long-term usage of group pairing protocols in real-world conditions.

Lastly, tech-savvy users expressed curiosity about PairSonic's operation, its security properties, and how the acoustic channel protects security. Including an in-app help menu with detailed explanations could assist users with questions about the system's functionality or security. Similarly, Herzberg et al. [44] suggested that for an authentication ceremony to be effective, it must not only be usable, but users also need to understand its necessity.

10 Limitations

This section addresses the limitations of our work due to our recruitment process and study design.

10.1 Recruitment

Our sample was relatively young, mainly comprising students and participants aged 20-29. Rather than selectively sampling for a representative distribution, we controlled for factors potentially influencing usability (Section 7.6). Neither our control variables nor age significantly correlated with usability, security, or preference scores.

However, a previous study by Kobsa et al. [59] reported significant differences in completion times between younger and older participants in pairing protocols. Given this, our younger sample may underestimate completion times for the broader population. Particularly, the scalability difference between SafeSlinger's word comparison and our automated verification could be more notable among older users.

10.2 Order Effect in Our Within-Subjects Design

In our correlational analysis (Section 7.6), we found some evidence of a potential order effect impacting the SUS scores. Participants who first tried PairSonic generally rated SafeSlinger lower (Mdn = 67.5), possibly due to the extra effort SafeSlinger requires (contrast effect). However, we found no evidence of a contrast effect in reverse: participants who started with SafeSlinger typically assigned high initial SUS scores (Mdn = 86.25), likely because they had to fill in the SUS questionnaire before trying the second system. Although these participants generally rated PairSonic higher after using it, the high initial score for SafeSlinger limited the relative improvement. We counterbalanced the sequence of systems in our study design to minimize potential order effects. Regardless of the order of exposure to the systems, participants generally gave PairSonic higher SUS scores.

10.3 Anchor Effect in Groups

We conducted our study with groups of multiple participants. During the group interviews, we sequentially asked the same question to each participant, which may have triggered an anchoring effect, as participants' responses could have influenced each other. We noticed that some participants picked up arguments from previous respondents, but we also noticed instances of disagreement; most groups did not have a unanimous opinion (Figure 9). Our study design partially mitigates the anchoring effect by having participants complete questionnaire sections before the corresponding interviews. This approach allowed them to reflect on the tasks and form their own opinions prior to the discussion. Given the cooperative nature of the group pairing process, conducting the experiment within a group of participants appears justified.

10.4 Lab Study

Our participants tested PairSonic and SafeSlinger three times each in our lab. Although a lab setting does not fully replicate real-world conditions [16], it offers a consistent and reproducible environment, which minimizes confounding factors. While a field study might have more accurately

represented real-world usage, the presence of uncontrollable environmental and external variables could have adversely impacted the quality of the data. Given that PairSonic is still a proof-of-concept implementation, the lab environment was particularly advantageous for observing our participants' interactions and reactions to the apps in real time and to discover problems in the user experience. While observing participants might influence their behavior, this setup enabled us to identify an unexpected user interaction that resulted in protocol failures (Section 8.4). Such issues, which were not evident during our testing, would have been challenging to detect in a field study. In addition, our study design strategically interleaved practical tasks with post-test questionnaires and interviews. This allowed us to promptly capture our participants' thoughts on each system before they experienced the other, a process more feasibly managed in the lab setting. We also wanted to specifically evaluate how our participants would handle potential active attacks on the pairing process. Simulating such attacks in a controlled manner is much easier in a lab environment.

11 Conclusion

We introduced a novel, practical method for secure contact information exchange, known as *PairSonic*, and compared it to the state-of-the-art system *SafeSlinger* [23] in a within-subjects lab study ($N = 45$). Our system's primary advantage is that it reduces user effort by automating laborious manual tasks using acoustic communication. Our results reveal that while users value ease in authentication systems, they often associate more complex systems with higher security.

RQ1. *"Which initialization and verification steps do users prefer?"* Our participants showed a significant preference for PairSonic over SafeSlinger. In the interviews, most preferred our leader-based initialization method and the acoustic verification step.

RQ2. *"Which method has better usability?"* PairSonic is significantly more usable than SafeSlinger. Our participants appreciated the seamless contact exchange and the effortless scalability to larger groups. However, there's a need to enhance the reliability of the OOB channel.

RQ3. *"How do users perceive the security of both methods?"* We observed no significant difference in the security ratings for both methods. Although minimizing user interaction improved usability, it surprisingly decreased perceived security for some participants, according to our interviews. Many users, drawing from their experiences, associate security with complexity, underscoring a tricky trade-off between usability and perceived security.

RQ4. *"How do users like the audible acoustic OOB channel for pairing?"* Our study indicates that acoustic communication is a promising, user-friendly method for data exchange between nearby devices, reinforcing the intuitive action of bringing devices close for association. However, future work should focus on enhancing the robustness of the physical layer in everyday scenarios.

Availability

Together with this paper, we provide an overview of the PairSonic project online at <https://seemoo.de/s/pairsonic>. PairSonic is available as open-source software on GitHub: <https://github.com/seemoo-lab/pairsonic>. The PairSonic app will also be presented as a demo at CSCW 2024 [91].

Alongside this paper, we release a replication package that includes our evaluation scripts and the pseudonymized dataset from our study [90]. This dataset contains usability, security, and preference scores, completion times, reported usage of nine types of social and collaborative tools, and seven demographic and control variables, for each of our 45 participants.

Acknowledgments

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72]. We thank the anonymous reviewers for their helpful suggestions. Furthermore, we thank Maximilian Gehring for his contributions to the initial PairSonic prototype, and Lea Halaus for drawing the illustration in [Figure 1](#). We also acknowledge SafeSlinger's substantial innovation over previous pairwise systems, which greatly inspired us in developing PairSonic.

References

- [1] Michel Abdalla, Emmanuel Bresson, Olivier Chevassut, and David Pointcheval. 2006. Password-Based Group Key Exchange in a Constant Number of Rounds. In *Public Key Cryptography - PKC 2006 (Lecture Notes in Computer Science)*, Moti Yung, Yevgeniy Dodis, Aggelos Kiayias, and Tal Malkin (Eds.). Springer, Berlin, Heidelberg, 427–442. https://doi.org/10.1007/11745853_28
- [2] Ruba Abu-Salma, M. Angela Sasse, Joseph Bonneau, Anastasia Danilova, Alena Naiakshina, and Matthew Smith. 2017. Obstacles to the Adoption of Secure Communication Tools. In *2017 IEEE Symposium on Security and Privacy (SP)*. 137–153. <https://doi.org/10.1109/SP.2017.65>
- [3] N Asokan and Philip Ginzboorg. 2000. Key Agreement in Ad Hoc Networks. *Computer Communications* 23, 17 (Nov. 2000), 1627–1637. [https://doi.org/10.1016/S0140-3664\(00\)00249-8](https://doi.org/10.1016/S0140-3664(00)00249-8)
- [4] Hala Assal, Stephanie Hurtado, Ahsan Imran, and Sonia Chiasson. 2015. What's the Deal with Privacy Apps? A Comprehensive Exploration of User Perception and Usability. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM '15)*. Association for Computing Machinery, New York, NY, USA, 25–36. <https://doi.org/10.1145/2836041.2836044>
- [5] Dirk Balfanz, Diana K. Smetters, Paul Stewart, and H. Chi Wong. 2002. Talking to Strangers: Authentication in Ad-Hoc Wireless Networks. In *NDSS Symposium*.
- [6] Matt Blaze, Joan Feigenbaum, and Jack Lacy. 1996. Decentralized Trust Management. In *Proceedings 1996 IEEE Symposium on Security and Privacy*. 164–173. <https://doi.org/10.1109/SECPRI.1996.502679>
- [7] John Brooke. 1996. SUS: A Quick and Dirty Usability Scale. *Usability Evaluation in Industry* (1996).
- [8] Mike Burmester and Yvo G. Desmedt. 1997. Efficient and Secure Conference-Key Distribution. In *Security Protocols (Lecture Notes in Computer Science)*, Mark Lomas (Ed.). Springer, Berlin, Heidelberg, 119–129. https://doi.org/10.1007/3-540-62494-5_12
- [9] Germano Caronni. 2000. Walking the Web of Trust. In *Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2000)*. 153–158. <https://doi.org/10.1109/ENABL.2000.883720>
- [10] Bin Chen, Koki Hatada, Keiju Okabayashi, Hiroyuki Kuromiya, Ichiro Hidaka, Yoshiharu Yamamoto, and Kazumasa Togami. 2019. Group Activity Recognition to Support Collaboration in Creative Digital Space. In *Companion Publication of the 2019 Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 175–179. <https://doi.org/10.1145/3311957.3359471>
- [11] Chia-Hsin Owen Chen, Chung-Wei Chen, Cynthia Kuo, Yan-Hao Lai, Jonathan M. McCune, Ahren Studer, Adrian Perrig, Bo-Yin Yang, and Tzong-Chen Wu. 2008. GAnGS: Gather, Authenticate 'n Group Securely. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom '08)*. Association for Computing Machinery, New York, NY, USA, 92–103. <https://doi.org/10.1145/1409944.1409957>
- [12] Ming Ki Chong and Hans Gellersen. 2011. How Users Associate Wireless Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1909–1918. <https://doi.org/10.1145/1978942.1979219>
- [13] Ming Ki Chong and Hans Gellersen. 2012. Usability Classification for Spontaneous Device Association. *Personal and Ubiquitous Computing* 16, 1 (Jan. 2012), 77–89. <https://doi.org/10.1007/s00779-011-0421-1>
- [14] Ming Ki Chong and Hans W. Gellersen. 2013. How Groups of Users Associate Wireless Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1559–1568. <https://doi.org/10.1145/2470654.2466207>
- [15] Ming Ki Chong, Fahim Kawsar, and Hans Gellersen. 2011. Spatial Co-Location for Device Association: The Connected Object Way. In *Proceedings of the 2011 International Workshop on Networking and Object Memories for the Internet of Things (NoME-IoT '11)*. Association for Computing Machinery, New York, NY, USA, 21–26. <https://doi.org/10.1145/2029932.2029941>
- [16] Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A Survey of User Interaction for Spontaneous Device Association. *Comput. Surveys* 47, 1 (May 2014), 8:1–8:40. <https://doi.org/10.1145/2597768>

- [17] Jacob Cohen. 1992. A power primer. *Psychological bulletin* 112(1) (1992), 155–159.
- [18] Adrian A. de Freitas and Anind K. Dey. 2015. The Group Context Framework: An Extensible Toolkit for Opportunistic Grouping and Collaboration. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1602–1611. <https://doi.org/10.1145/2675133.2675205>
- [19] Sergej Dechand, Yasemin Acar, Dominik Schurmann, Sascha Fahl, Karoline Busse, and Matthew Smith. 2016. An Empirical Study of Textual Key-Fingerprint Representations. *USENIX Security* (2016).
- [20] Danny Dolev and Andrew Yao. 1983. On the Security of Public Key Protocols. *IEEE Transactions on Information Theory* 29, 2 (March 1983), 198–208. <https://doi.org/10.1109/TIT.1983.1056650>
- [21] John R. Douceur. 2002. The Sybil Attack. In *Peer-to-Peer Systems (Lecture Notes in Computer Science)*, Peter Druschel, Frans Kaashoek, and Antony Rowstron (Eds.). Springer, Berlin, Heidelberg, 251–260. https://doi.org/10.1007/3-540-45748-8_24
- [22] W. Keith Edwards, Mark W. Newman, Jana Z. Sedivy, Trevor F. Smith, Dirk Balfanz, D. K. Smetters, H. Chi Wong, and Shahram Izadi. 2002. Using Speakeasy for Ad Hoc Peer-to-Peer Collaboration. In *Proceedings of the 2002 ACM Conference on Computer-Supported Cooperative Work (CSCW '02)*. Association for Computing Machinery, New York, NY, USA, 256–265. <https://doi.org/10.1145/587078.587114>
- [23] Michael Farb, Yue-Hsun Lin, Tiffany Hyun-Jin Kim, Jonathan McCune, and Adrian Perrig. 2013. SafeSlinger: Easy-to-Use and Secure Public-Key Exchange. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom '13)*. Association for Computing Machinery, New York, NY, USA, 417–428. <https://doi.org/10.1145/2500423.2500428>
- [24] Habiba Farrukh, Muslum Ozmen, Faik Ors, and Berkay Celik. 2023. One Key to Rule Them All: Secure Group Pairing for Heterogeneous IoT Devices. In *IEEE S&P 2023*. <https://www.computer.org/csdl/proceedings-article/sp/2023/933600b693/1Js0EfOaua1>
- [25] Matthias Fassl, Lea Theresa Gröber, and Katharina Krombholz. 2021. Exploring User-Centered Security Design for Usable Authentication Ceremonies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3411764.3445164>
- [26] Matthias Fassl and Katharina Krombholz. 2023. Why I Can't Authenticate — Understanding the Low Adoption of Authentication Ceremonies with Autoethnography. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Hamburg, Germany. <https://publications.cispa.saarland/3895/>
- [27] Hal Finney, Lutz Donnerhacke, Jon Callas, Rodney L. Thayer, and Daphne Shaw. 2007. *OpenPGP Message Format*. Request for Comments RFC 4880. Internet Engineering Task Force. <https://doi.org/10.17487/RFC4880>
- [28] Mikhail Fomichev, Flor Alvarez, Daniel Steinmetzer, Paul Gardner-Stephen, and Matthias Hollick. 2018. Survey and Systematization of Secure Device Pairing. *IEEE Communications Surveys & Tutorials* 20, 1 (2018), 517–550. <https://doi.org/10.1109/COMST.2017.2748278>
- [29] Mikhail Fomichev, Max Maass, Lars Almon, Alejandro Molina, and Matthias Hollick. 2019. Perils of Zero-Interaction Security in the Internet of Things. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 10 (mar 2019), 38 pages. <https://doi.org/10.1145/3314397>
- [30] Lesley Fosh, Steve Benford, and Boriana Koleva. 2016. Supporting Group Coherence in a Museum Visit. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/2818048.2819970>
- [31] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human-Computer Interaction* (2019).
- [32] Vadim Gerasimov and Walter Bender. 2000. Things That Talk: Using Sound for Device-to-Device and Device-to-Human Communication. *IBM Systems Journal* 39, 3.4 (2000), 530–546–530–546. <https://doi.org/10.1147/sj.393.0530>
- [33] Pascal Getreuer, Chet Gnedy, Richard F. Lyon, and Rif A. Saurous. 2018/june. Ultrasonic Communication Using Consumer Hardware. *IEEE Transactions on Multimedia* 20, 6 (2018/june), 1277–1290. <https://doi.org/10.1109/TMM.2017.2766049>
- [34] Nirnimesh Ghose, Loukas Lazos, and Ming Li. 2017. HELP: Helper-Enabled in-Band Device Pairing Resistant against Signal Cancellation. In *Proceedings of the 26th USENIX Conference on Security Symposium (SEC'17)*. USENIX Association, USA, 433–450.
- [35] Nirnimesh Ghose, Loukas Lazos, and Ming Li. 2018. Secure Device Bootstrapping Without Secrets Resistant to Signal Manipulation Attacks. In *2018 IEEE Symposium on Security and Privacy (SP)*. 819–835. <https://doi.org/10.1109/SP.2018.00055>
- [36] Nirnimesh Ghose, Loukas Lazos, and Ming Li. 2022. In-Band Secret-Free Pairing for COTS Wireless Devices. *IEEE Transactions on Mobile Computing* 21, 2 (Feb. 2022), 612–628. <https://doi.org/10.1109/TMC.2020.3015010>

- [37] Dan Goodin. 2018. *Police Decrypt 258,000 Messages after Breaking Pricey IronChat Crypto App*. <https://arstechnica.com/information-technology/2018/11/police-decrypt-258000-messages-after-breaking-pricey-ironchat-crypto-app/> [Retrieved 2024-01-07].
- [38] Michael T. Goodrich, Michael Sirivianos, John Solis, Claudio Soriente, Gene Tsudik, and Ersin Uzun. 2009/february. Using Audio in Secure Device Pairing. *Int. J. Secur. Netw.* 4, 1/2 (2009/february), 57–68–57–68. <https://doi.org/10.1504/IJSN.2009.023426>
- [39] Michael T. Goodrich, Michael Sirivianos, John Solis, Gene Tsudik, and Ersin Uzun. 2006. Loud and Clear: Human-Verifiable Authentication Based on Audio. In *26th IEEE International Conference on Distributed Computing Systems (ICDCS. IEEE)*. <https://doi.org/10.1109/icdcs.2006.52>
- [40] Jens Emil Grønbaek, Mille Skovhus Knudsen, Kenton O'Hara, Peter Gall Krogh, Jo Vermeulen, and Marianne Graves Petersen. 2020. Proxemics Beyond Proximity: Designing for Flexible Social Interaction Through Cross-Device Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376379>
- [41] Dianqi Han, Ang Li, Jiawei Li, Yan Zhang, Tao Li, and Yanchao Zhang. 2021. DroneKey: A Drone-Aided Group-Key Generation Scheme for Large-Scale IoT Networks. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS '21)*. Association for Computing Machinery, New York, NY, USA, 1306–1319. <https://doi.org/10.1145/3460120.3484789>
- [42] Mark Handel and James D. Herbsleb. 2002. What Is Chat Doing in the Workplace?. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW '02)*. Association for Computing Machinery, New York, NY, USA, 1–10. <https://doi.org/10.1145/587078.587080>
- [43] Mahboobeh Harandi. 2019. Supporting Occasional Groups in Crowdsourcing Platforms. In *Companion Publication of the 2019 Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 52–55. <https://doi.org/10.1145/3311957.3361856>
- [44] Amir Herzberg and Hemi Leibowitz. 2016. Can Johnny Finally Encrypt? Evaluating E2E-encryption in Popular IM Applications. In *Proceedings of the 6th Workshop on Socio-Technical Aspects in Security and Trust (STAST '16)*. Association for Computing Machinery, New York, NY, USA, 17–28. <https://doi.org/10.1145/3046055.3046059>
- [45] Amir Herzberg, Hemi Leibowitz, Kent Seamons, Elham Vaziripour, Justin Wu, and Daniel Zappala. 2021. Secure Messaging Authentication Ceremonies Are Broken. *IEEE Security & Privacy* 19, 2 (March 2021), 29–37. <https://doi.org/10.1109/MSEC.2020.3039727>
- [46] Yantian Hou, Ming Li, and Joshua D. Guttman. 2013. Chorus: Scalable In-band Trust Establishment for Multiple Constrained Devices over the Insecure Wireless Channel. In *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*. ACM, 167–178–167–178. <https://doi.org/10.1145/2462096.2462124>
- [47] Iulia Ion, Marc Langheinrich, Ponnurangam Kumaraguru, and Srdjan Čapkun. 2010. Influence of User Perception, Security Needs, and Social Factors on Device Pairing Method Choices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/1837110.1837118>
- [48] Tero Jokela, Ming Ki Chong, Andrés Lucero, and Hans Gellersen. 2015. Connecting Devices for Collaborative Interactions. *Interactions* 22, 4 (June 2015), 39–43. <https://doi.org/10.1145/2776887>
- [49] Tero Jokela and Andrés Lucero. 2013. A Comparative Evaluation of Touch-Based Methods to Bind Mobile Devices for Collaborative Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3355–3364. <https://doi.org/10.1145/2470654.2466459>
- [50] Tero Jokela and Andrés Lucero. 2014. FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 369–378. <https://doi.org/10.1145/2628363.2628376>
- [51] Tero Jokela, Parisa Pour Rezaei, and Kaisa Väänänen. 2016. Natural Group Binding and Cross-Display Object Movement Methods for Wearable Devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 206–216. <https://doi.org/10.1145/2935334.2935346>
- [52] Daniel Jones. 2022. How Near-Ultrasonic Audio Adds Spatial Awareness to the Sonos System. <https://tech-blog.sonos.com/posts/how-near-ultrasonic-audio-adds-spatial-awareness-to-the-sonos-system/> [Retrieved 2024-01-07].
- [53] Ronald Kainda, Ivan Flechais, and A. W. Roscoe. 2009. Usability and Security of Out-of-Band Channels in Secure Device Pairing Protocols. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/1572532.1572547>
- [54] Ronald Kainda, Ivan Flechais, and A. W. Roscoe. 2010. Two Heads Are Better than One: Security and Usability of Device Associations in Group Scenarios. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS*

- '10). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/1837110.1837117>
- [55] Maurice G. Kendall. 1938. A New Measure of Rank Correlation. *Biometrika* 30 (1938).
- [56] Sye Loong Keoh, Emil Lupu, and Morris Sloman. 2009. Securing Body Sensor Networks: Sensor Association and Key Management. In *2009 IEEE International Conference on Pervasive Computing and Communications*. 1–6. <https://doi.org/10.1109/PERCOM.2009.4912756>
- [57] Tim Kindberg and Armando Fox. 2002. System Software for Ubiquitous Computing. *IEEE Pervasive Computing* 1, 1 (Jan. 2002), 70–81. <https://doi.org/10.1109/MPRV.2002.993146>
- [58] Darko Kirovski, Michael Sinclair, and David Wilson. 2007. The Martini Synch: Device Pairing via Joint Quantization. In *2007 IEEE International Symposium on Information Theory*. 466–470. <https://doi.org/10.1109/ISIT.2007.4557269>
- [59] Alfred Kobsa, Rahim Sonawalla, Gene Tsudik, Ersin Uzun, and Yang Wang. 2009. Serial Hook-Ups: A Comparative Usability Study of Secure Device Pairing Methods. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/1572532.1572546>
- [60] Tonko Kovačević, Toni Perković, and Mario Čagalj. 2016. Flashing Displays: User-Friendly Solution for Bootstrapping Secure Associations between Multiple Constrained Wireless Devices. *Security and Communication Networks* 9, 10 (2016), 1050–1071. <https://doi.org/10.1002/sec.1400>
- [61] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. 2010. User-Defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 239–248. <https://doi.org/10.1145/1851600.1851640>
- [62] Arun Kumar, Nitesh Saxena, Gene Tsudik, and Ersin Uzun. 2009. A Comparative Study of Secure Device Pairing Methods. *Pervasive and Mobile Computing* 5, 6 (Dec. 2009), 734–749. <https://doi.org/10.1016/j.pmcj.2009.07.008>
- [63] Cynthia Kuo, Mark Luk, Rohit Negi, and Adrian Perrig. 2007. Message-in-a-Bottle: User-Friendly and Secure Key Deployment for Sensor Nodes. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys '07)*. Association for Computing Machinery, New York, NY, USA, 233–246. <https://doi.org/10.1145/1322263.1322286>
- [64] Cynthia Kuo, Ahren Studer, and Adrian Perrig. 2008. Mind Your Manners: Socially Appropriate Wireless Key Establishment for Groups. In *Proceedings of the First ACM Conference on Wireless Network Security (WiSec '08)*. Association for Computing Machinery, New York, NY, USA, 125–130. <https://doi.org/10.1145/1352533.1352553>
- [65] Sven Laur and Sylvain Pasini. 2008. SAS-Based Group Authentication and Key Agreement Protocols. In *Public Key Cryptography – PKC 2008 (Lecture Notes in Computer Science)*, Ronald Cramer (Ed.). Springer, Berlin, Heidelberg, 197–213. https://doi.org/10.1007/978-3-540-78440-1_12
- [66] Yee Wei Law, Giorgi Moniava, Zheng Gong, Pieter Hartel, and Marimuthu Palaniswami. 2011. KALwEN: A New Practical and Interoperable Key Management Scheme for Body Sensor Networks. *Security and Communication Networks* 4, 11 (Nov. 2011), 1309–1329. <https://doi.org/10.1002/sec.256>
- [67] Hyewon Lee, Tae H. Kim, Jun W. Choi, and Sunghyun Choi. 2015/april. Chirp Signal-Based Aerial Acoustic Communication for Smart Devices. In *INFOCOM.2015*. 2407–2415. <https://doi.org/10.1109/INFOCOM.2015.7218629>
- [68] Ada Lerner, Eric Zeng, and Franziska Roesner. 2017. Confidante: Usable Encrypted Email: A Case Study with Lawyers and Journalists. In *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*. 385–400. <https://doi.org/10.1109/EuroSP.2017.41>
- [69] Ian Levy and Crispin Robinson. 2018. *Principles for a More Informed Exceptional Access Debate*. <https://www.lawfaremedia.org/article/principles-more-informed-exceptional-access-debate> [Retrieved 2024-01-07].
- [70] Ming Li, Shucheng Yu, Wenjing Lou, and Kui Ren. 2010. Group Device Pairing Based Secure Sensor Association and Key Management for Body Area Networks. In *2010 Proceedings IEEE INFOCOM*. 1–9. <https://doi.org/10.1109/INFOCOM.2010.5462095>
- [71] Xiaopeng Li, Qiang Zeng, Lannan Luo, and Tongbo Luo. 2020. T2Pair: Secure and Usable Pairing for Heterogeneous IoT Devices. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. Association for Computing Machinery, New York, NY, USA, 309–323. <https://doi.org/10.1145/3372297.3417286>
- [72] Yue-Hsun Lin, Ahren Studer, Yao-Hsin Chen, Hsu-Chun Hsiao, Li-Hsiang Kuo, Jonathan M. McCune, King-Hang Wang, Maxwell Krohn, Adrian Perrig, Bo-Yin Yang, Hung-Min Sun, Phen-Lan Lin, and Jason Lee. 2010. SPATE: Small-Group PKI-Less Authenticated Trust Establishment. *IEEE Transactions on Mobile Computing* 9, 12 (Dec. 2010), 1666–1681. <https://doi.org/10.1109/TMC.2010.150>
- [73] Yue-Hsun Lin, Ahren Studer, Hsu-Chin Hsiao, Jonathan M. McCune, King-Hang Wang, Maxwell Krohn, Phen-Lan Lin, Adrian Perrig, Hung-Min Sun, and Bo-Yin Yang. 2009. SPATE: Small-Group PKI-less Authenticated Trust Establishment. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/1555816.1555818>

- [74] Cristina V. Lopes and Pedro M. Q. Aguiar. 2003/july. Acoustic Modems for Ubiquitous Computing. *IEEE Pervasive Computing* 2, 3 (2003/july), 62–71. <https://doi.org/10.1109/MPRV.2003.1228528>
- [75] Sus Lundgren, Joel E. Fischer, Stuart Reeves, and Olof Torgersson. 2015. Designing Mobile Experiences for Collocated Interaction. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 496–507. <https://doi.org/10.1145/2675133.2675171>
- [76] Anil Madhavapeddy, David Scott, and Richard Sharp. 2003. Context-Aware Computing with Sound. In *UbiComp 2003: Ubiquitous Computing (Lecture Notes in Computer Science)*, Anind K. Dey, Albrecht Schmidt, and Joseph F. McCarthy (Eds.). Springer, Berlin, Heidelberg, 315–332. https://doi.org/10.1007/978-3-540-39653-6_25
- [77] Rene Mayrhofer and Hans Gellersen. 2009. Shake Well Before Use: Intuitive and Secure Pairing of Mobile Devices. *IEEE Transactions on Mobile Computing* 8, 6 (June 2009), 792–806. <https://doi.org/10.1109/TMC.2009.51>
- [78] Adib Mehrabi, Antonella Mazzoni, Daniel Jones, and Anthony Steed. 2019. Evaluating the User Experience of Acoustic Data Transmission. *Personal and Ubiquitous Computing* (Dec. 2019), 1–14–1–14. <https://doi.org/10.1007/s00779-019-01345-7>
- [79] Sharan B. Merriam and Elizabeth J. Tisdell. 2015. *Qualitative Research: A Guide to Design and Implementation*. John Wiley & Sons.
- [80] Meta. 2009. *WhatsApp*. <https://whatsapp.com> [Software].
- [81] Ghita Mezzour, Ahren Studer, Michael Farb, Jason Lee, Jonathan McCune, Hsu-Chun Hsiao, and Adrian Perrig. 2010. *Ho-Po Key: Leveraging Physical Constraints on Human Motion to Authentically Exchange Information in a Group*. Technical Report CMU-CyLab-11-004. Carnegie Mellon University.
- [82] Moses Namara and Bart P. Knijnenburg. 2021. The Differential Effect of Privacy-Related Trust on Groupware Application Adoption and Use during the COVID-19 Pandemic. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 405:1–405:34. <https://doi.org/10.1145/3479549>
- [83] Rajalakshmi Nandakumar, Krishna K. Chintalapudi, Venkat Padmanabhan, and Ramarathnam Venkatesan. 2013. Dhvani: Secure Peer-to-Peer Acoustic NFC. *ACM SIGCOMM Computer Communication Review* 43, 4 (Aug. 2013), 63–74–63–74. <https://doi.org/10.1145/2534169.2486037>
- [84] Long H. Nguyen and Andrew W. Roscoe. 2008. Authenticating Ad Hoc Networks by Comparison of Short Digests. *Information and Computation* 206, 2 (Feb. 2008), 250–271. <https://doi.org/10.1016/j.ic.2007.07.010>
- [85] Rishab Nithyanand, Nitesh Saxena, Gene Tsudik, and Ersin Uzun. 2010. Groupthink: Usability of Secure Group Association for Wireless Devices. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing (UbiComp '10)*. Association for Computing Machinery, New York, NY, USA, 331–340. <https://doi.org/10.1145/1864349.1864399>
- [86] Sean Oesch, Ruba Abu-Salma, Oumar Diallo, Juliane Krämer, James Simmons, Justin Wu, and Scott Ruoti. 2022. User Perceptions of Security and Privacy for Group Chat. *Digital Threats: Research and Practice* 3, 2 (Feb. 2022), 15:1–15:29. <https://doi.org/10.1145/3491265>
- [87] Karl Pearson. 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50 (1900).
- [88] Toni Perković, Mario Čagalj, Toni Mastelić, Nitesh Saxena, and Dinko Begušić. 2012. Secure Initialization of Multiple Constrained Wireless Devices for an Unaided User. *IEEE Transactions on Mobile Computing* 11, 2 (2012), 337–351. <https://doi.org/10.1109/TMC.2011.35>
- [89] Florentin Putz, Flor Álvarez, and Jiska Classen. 2020. Acoustic Integrity Codes: Secure Device Pairing Using Short-Range Acoustic Communication. In *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 31–41. <https://doi.org/10.1145/3395351.3399420>
- [90] Florentin Putz, Steffen Haesler, and Matthias Hollick. 2024. *Lab Study Dataset: Fast and Secure Contact Exchange in Groups*. <https://doi.org/10.5281/zenodo.13324112> [Dataset].
- [91] Florentin Putz, Steffen Haesler, Thomas Völkl, Maximilian Gehring, Nils Rollshausen, and Matthias Hollick. 2024. PairSonic: Helping Groups Securely Exchange Contact Information. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '24 Companion)*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3678884.3681818>
- [92] R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- [93] David A. Robb, Thomas S. Methven, Stefano Padilla, and Mike J. Chantler. 2016. Well-Connected: Promoting Collaboration by Effective Networking. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing Companion (CSCW '16 Companion)*. Association for Computing Machinery, New York, NY, USA, 90–93. <https://doi.org/10.1145/2818052.2874333>
- [94] Robert Rosenthal. 1991. *Meta-Analytic Procedures for Social Research*. Sage Publications, Inc.
- [95] Jeffrey Rubin and Dana Chrisnell. 2008. *Handbook of Usability Testing* (second edition ed.). Wiley Publishing, Inc.

- [96] Scott Ruoti, Jeff Andersen, Scott Heidbrink, Mark O'Neill, Elham Vaziripour, Justin Wu, Daniel Zappala, and Kent Seamons. 2016. "We're on the Same Page": A Usability Study of Secure Email Using Pairs of Novice Users. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4298–4308. <https://doi.org/10.1145/2858036.2858400>
- [97] Scott Ruoti, Nathan Kim, Ben Burgon, Timothy van der Horst, and Kent Seamons. 2013. Confused Johnny: When Automatic Encryption Leads to Confusion and Mistakes. In *Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS '13)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/2501604.2501609>
- [98] G. Enrico Santagati and Tommaso Melodia. 2015. U-Wear: Software-Defined Ultrasonic Networking for Wearable Devices. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 241–256–241–256. <https://doi.org/10.1145/2742647.2742655>
- [99] Nitesh Saxena and Md. Borhan Uddin. 2009. Blink 'Em All: Scalable, User-Friendly and Secure Initialization of Wireless Sensor Nodes. In *Cryptology and Network Security (Lecture Notes in Computer Science)*, Juan A. Garay, Atsuko Miyaji, and Akira Otsuka (Eds.). Springer, Berlin, Heidelberg, 154–173. https://doi.org/10.1007/978-3-642-10433-6_11
- [100] Svenja Schröder, Markus Huber, David Wind, and Christoph Rottermann. 2016. When SIGNAL Hits the Fan: On the Usability and Security of State-of-the-Art Secure Mobile Messaging. In *EuroUSEC*.
- [101] Samuel S. Shapiro and Martin B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (1965).
- [102] Maliheh Shirvanian and Nitesh Saxena. 2015. On the Security and Usability of Crypto Phones. In *Proceedings of the 31st Annual Computer Security Applications Conference (ACSAC '15)*. Association for Computing Machinery, New York, NY, USA, 21–30. <https://doi.org/10.1145/2818000.2818007>
- [103] Maliheh Shirvanian, Nitesh Saxena, and Jesvin James George. 2017. On the Pitfalls of End-to-End Encrypted Communications: A Study of Remote Key-Fingerprint Verification. In *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC '17)*. Association for Computing Machinery, New York, NY, USA, 499–511. <https://doi.org/10.1145/3134600.3134610>
- [104] Erica Shusas, Patrick Skeba, Eric P. S. Baumer, and Andrea Forte. 2023. Accounting for Privacy Pluralism: Lessons and Strategies from Community-Based Privacy Groups. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3544548.3581331>
- [105] Sidney Siegel. 1956. *Nonparametric statistics for the behavioral sciences*. McGraw-Hill, New York.
- [106] Signal Foundation. 2014. *Signal*. <https://signal.org> [Software].
- [107] Claudio Soriente, Gene Tsudik, and Ersin Uzun. 2008. HAPADEP: Human-Assisted Pure Audio Device Pairing. In *Information Security*, Tzong-Chen Wu, Chin-Laung Lei, Vincent Rijmen, and Der-Tsai Lee (Eds.). Springer Berlin Heidelberg, 385–400–385–400. https://doi.org/10.1007/978-3-540-85886-7_27
- [108] Frank Stajano and Ross Anderson. 2000. The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks. In *Security Protocols (Lecture Notes in Computer Science)*, Bruce Christianson, Bruno Crispo, James A. Malcolm, and Michael Roe (Eds.). Springer, Berlin, Heidelberg, 172–182. https://doi.org/10.1007/10720107_24
- [109] Ahren Studer, Christina Johns, Jaanus Kase, Kyle O'Meara, and Lorrie Cranor. 2008. A Survey to Guide Group Key Protocol Development. In *2008 Annual Computer Security Applications Conference (ACSAC)*. 475–484. <https://doi.org/10.1109/ACSAC.2008.28>
- [110] Milan Stute, David Kreitschmann, and Matthias Hollick. 2018. One Billion Apples' Secret Sauce: Recipe for the Apple Wireless Direct Link Ad Hoc Protocol. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking (New Delhi, India) (MobiCom '18)*. Association for Computing Machinery, New York, NY, USA, 529–543. <https://doi.org/10.1145/3241539.3241566>
- [111] Joshua Tan, Lujo Bauer, Joseph Bonneau, Lorrie Faith Cranor, Jeremy Thomas, and Blase Ur. 2017. Can Unicorns Help Users Compare Crypto Key Fingerprints?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3787–3798. <https://doi.org/10.1145/3025453.3025733>
- [112] Telegram FZ LLC. 2013. *Telegram*. <https://telegram.org/> [Software].
- [113] Threema GmbH. 2012. *Threema*. <https://threema.ch> [Software].
- [114] Peter Tolmie, Steve Benford, Chris Greenhalgh, Tom Rodden, and Stuart Reeves. 2014. Supporting Group Interactions in Museum Visiting. In *Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 1049–1059. <https://doi.org/10.1145/2531602.2531619>
- [115] Alexander Ulrich, Ralph Holz, Peter Hauck, and Georg Carle. 2011. Investigating the OpenPGP Web of Trust. In *Computer Security – ESORICS 2011 (Lecture Notes in Computer Science)*, Vijay Atluri and Claudia Diaz (Eds.). Springer, Berlin, Heidelberg, 489–507. https://doi.org/10.1007/978-3-642-23822-2_27

- [116] Gustavo Umbelino, Vivian Ta, Samuel Blake, Eric Truong, Amy Luo, and Steven Dow. 2019. ProtoTeams: Supporting Small Group Interactions in Co-Located Crowds. In *Companion Publication of the 2019 Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 392–397. <https://doi.org/10.1145/3311957.3359505>
- [117] Nik Unger, Sergej Dechand, Joseph Bonneau, Sascha Fahl, Henning Perl, Ian Goldberg, and Matthew Smith. 2015. SoK: Secure Messaging. In *2015 IEEE Symposium on Security and Privacy*. 232–249. <https://doi.org/10.1109/SP.2015.22>
- [118] Ersin Uzun, Kristiina Karvonen, and N. Asokan. 2007. Usability Analysis of Secure Pairing Methods. In *Financial Cryptography and Data Security (Lecture Notes in Computer Science)*, Sven Dietrich and Rachna Dhamija (Eds.). Springer, Berlin, Heidelberg, 307–324. https://doi.org/10.1007/978-3-540-77366-5_29
- [119] Ersin Uzun, Nitesh Saxena, and Arun Kumar. 2011. Pairing Devices for Social Interactions: A Comparative Usability Evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2315–2324. <https://doi.org/10.1145/1978942.1979282>
- [120] Jukka Valkonen, N. Asokan, and Kaisa Nyberg. 2006. Ad Hoc Security Associations for Groups. In *Security and Privacy in Ad-Hoc and Sensor Networks (Lecture Notes in Computer Science)*, Levente Buttyán, Virgil D. Gligor, and Dirk Westhoff (Eds.). Springer, Berlin, Heidelberg, 150–164. https://doi.org/10.1007/11964254_14
- [121] Diana A. Vasile, Martin Kleppmann, Daniel R. Thomas, and Alastair R. Beresford. 2020. Ghost Trace on the Wire? Using Key Evidence for Informed Decisions. In *Security Protocols XXVII (Lecture Notes in Computer Science)*, Jonathan Anderson, Frank Stajano, Bruce Christianson, and Vashek Matyáš (Eds.). Springer International Publishing, Cham, 245–257. https://doi.org/10.1007/978-3-030-57043-9_23
- [122] Elham Vaziripour, Justin Wu, Mark O'Neill, Ray Clinton, Jordan Whitehead, Scott Heidbrink, Kent Seamons, and Daniel Zappala. 2017. Is That You, Alice? A Usability Study of the Authentication Ceremony of Secure Messaging Applications. In *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security (SOUPS '17)*. USENIX Association, USA, 29–47.
- [123] Elham Vaziripour, Justin Wu, Mark O'Neill, Daniel Metro, Josh Cockrell, Timothy Moffett, Jordan Whitehead, Nick Bonner, Kent Seamons, and Daniel Zappala. 2018. Action Needed! Helping Users Find and Complete the Authentication Ceremony in Signal. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security (SOUPS '18)*. USENIX Association, USA, 47–62.
- [124] VERBI Software. 2021. MAXQDA 2022. VERBI Software, Berlin, Germany. <https://www.maxqda.com> [Software].
- [125] Quian Wang, Kui Ren, Man Zhou, Tao Lei, Dimitrios Koutsonikolas, and Lu Su. 2016. Messages behind the Sound. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking - MobiCom*. ACM Press. <https://doi.org/10.1145/2973750.2973765>
- [126] Honghao Wei, Cheng-Kang Hsieh, Longqi Yang, and Deborah Estrin. 2016. GroupLink: Group Event Recommendations Using Personal Digital Traces. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing Companion (CSCW '16 Companion)*. Association for Computing Machinery, New York, NY, USA, 110–113. <https://doi.org/10.1145/2818052.2874338>
- [127] Rand Wilcox. 2017. *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier, Academic Press.
- [128] Frank Wilcoxon. 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* 1, 6 (1945), 80–83. <https://doi.org/10.2307/3001968>
- [129] Justin Wu, Cyrus Gatrell, Devon Howard, Jake Tyler, Elham Vaziripour, Kent Seamons, and Daniel Zappala. 2019. Something Isn't Secure, but i'm Not Sure How That Translates into a Problem: Promoting Autonomy by Designing for Understanding in Signal. In *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security (SOUPS'19)*. USENIX Association, USA, 137–153.
- [130] Weitao Xu, Junqing Zhang, Shunqi Huang, Chengwen Luo, and Wei Li. 2021. Key Generation for Internet of Things: A Contemporary Survey. *Comput. Surveys* 54, 1 (Jan. 2021), 14:1–14:37. <https://doi.org/10.1145/3429740>
- [131] Jiahuan Zheng, Xin Peng, Jiacheng Yang, Huaqian Cai, Gang Huang, Ying Zhang, and Wenyun Zhao. 2017. CollaDroid: Automatic Augmentation of Android Application with Lightweight Interactive Collaboration. In *Proceedings of the 2017 ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 2462–2474. <https://doi.org/10.1145/2998181.2998278>

A Supplementary Figures

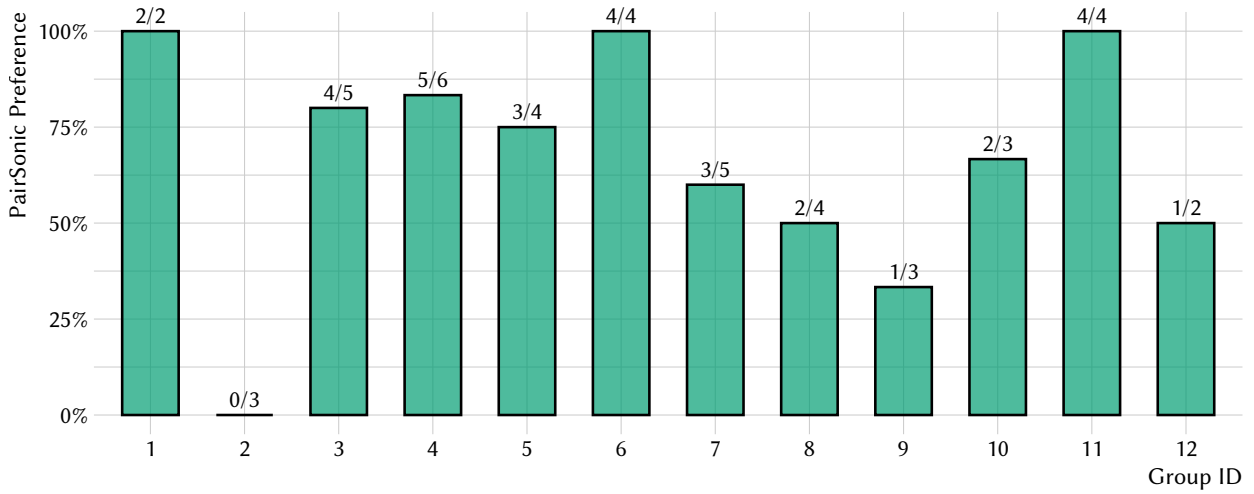


Fig. 9. This bar chart shows the proportion of group members preferring PairSonic in each study group. The label on top of each bar shows the corresponding number of group members who preferred PairSonic and the total group size. Only four groups had a unanimous preference (three favoring PairSonic, one favoring SafeSlinger). This variation suggests the possible anchor effect discussed in [Section 10.2](#) might not be very strong.

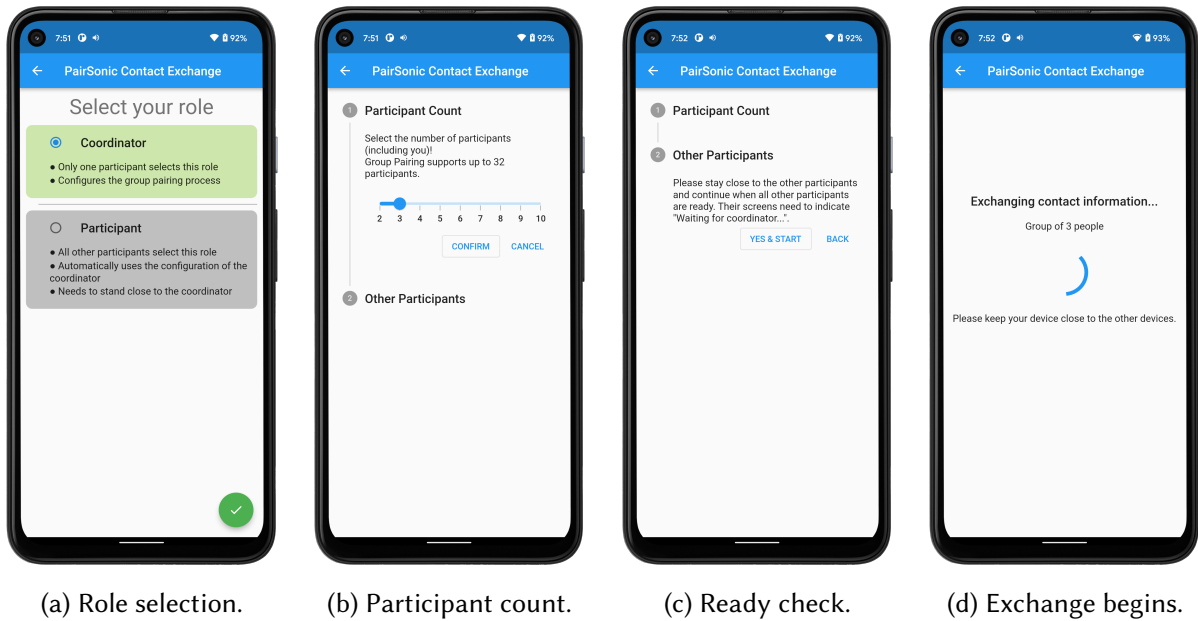
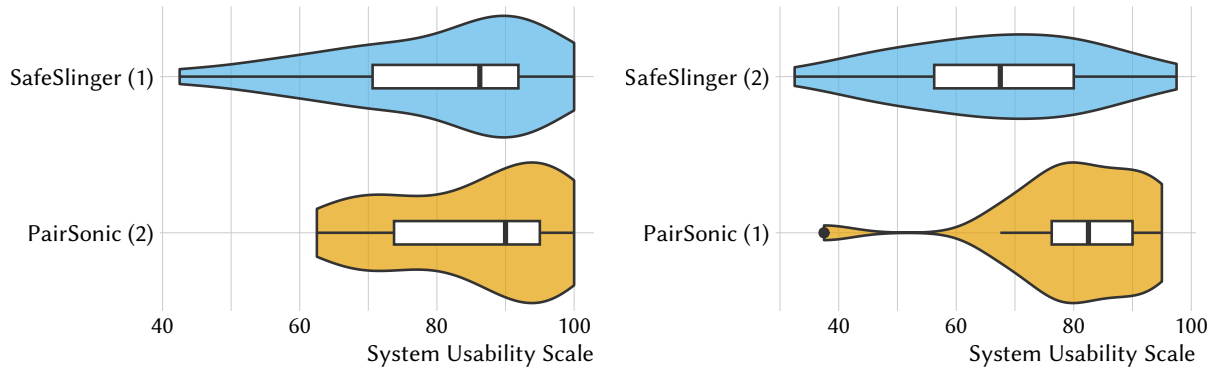


Fig. 10. This figure depicts PairSonic’s initialization phase from the perspective of the *coordinator* role. (a) The coordinator begins by selecting their role. (b) They proceed to select the total number of participants, counting themselves. (c) Before continuing, they ensure that the remaining participants are in the ready state ([Figure 2b](#)). (d) The coordinator’s smartphone then initiates the exchange via the acoustic OOB channel. The verification and finalization phases that follow are identical for both the *coordinator* and *participant* roles, as shown in [Figure 2c](#) and [Figure 2d](#).



(a) Participants who first encountered SafeSlinger. (b) Participants who first encountered PairSonic.

Fig. 11. Comparison of our participants' SUS scores for SafeSlinger and PairSonic, depending on the order they encountered both systems. The violin plots show a density estimation of the distributions. The boxplots show quartiles, median, and outliers.

Table 4. Kendall's correlation between control and dependent variables.

		2	3	4	5	6	7	8	9	ATI
1	Smartphone Familiarity	-.15	-.1	.19	.11	.15	.16	-.09	.2	-.14
2	Order (PairSonic First)		.11	-.33**	-.16	-.15	-.11	.11	-.24	-.18
3	Preference (PairSonic)			-.33*	.19	-.12	0	.14	.12	.18
4	SUS SafeSlinger				.17	.29*	.24*	-.34**	0	.07
5	SUS PairSonic					.31**	.28*	-.12	-.06	.16
6	Security SafeSlinger						.3*	0	-.05	.09
7	Security PairSonic							-.15	.04	.14
8	Time SafeSlinger								.09	.02
9	Time PairSonic									.14

Note: $N = 45$ * $p < .05$ ** $p < .01$

B Questionnaire

We provide a translation of our questionnaire. We gave the questionnaire to the participants in the native language of the country where we ran the study.

B.1 Experiment A

Please indicate the degree to which you agree/disagree with the following statements.

⟨Five responses ranging from “Strongly disagree” to “Strongly agree”⟩

- (1) I think that I would like to use this system frequently.
- (2) I found the system unnecessarily complex.
- (3) I thought the system was easy to use.
- (4) I think that I would need the support of a technical person to be able to use this system.
- (5) I found the various functions in this system were well integrated.
- (6) I thought there was too much inconsistency in this system.
- (7) I would imagine that most people would learn to use this system very quickly.
- (8) I found the system very cumbersome to use.
- (9) I felt very confident using the system.
- (10) I needed to learn a lot of things before i could get going with this system.

(11) I think that this system is secure.

B.2 Experiment B

⟨Same content as experiment A⟩

B.3 Preference

(1) Which system did you like better? ⟨Experiment A / Experiment B⟩

B.4 Groups

Which of the types of digital communication groups listed below do you participate in? What is the average number of participants in these groups? For which of these groups do you want to know exactly who is part of the group or that no unauthorized person has access?

⟨For each of the following categories: a checkbox “Participation?”, a field “Size?”, a checkbox “Security?”⟩

- (1) Public chat group (e. g., Discord, Telegram, IRC)
- (2) Private chat group (e. g., WhatsApp, Signal)
- (3) Professional chat group (e. g., Microsoft Teams)
- (4) Online forum
- (5) Group in social network (e. g., Facebook group)
- (6) Audio/video conference (e. g., TeamSpeak, Zoom)
- (7) Mailing list
- (8) Collaboration tools (e. g., Google Docs, Etherpad, Miro)
- (9) Project planning tools (e. g., Jira, Trello, GitHub)
- (10) ⟨Additional free text fields⟩

B.4.1 Total Number of Groups.

(1) Please estimate the total number of digital communication groups you participate in. ⟨Number field⟩

B.5 Demographic Information

- (1) Please state your gender. ⟨Male / Female / Diverse / Free text / No answer⟩
- (2) Please state your field of activity or field of studies. ⟨Free text response⟩
- (3) How old are you? ⟨18–19 / 20–24 / 25–29 / 30–34 / 35–39 / 40–44 / 45–49 / 50–54 / 55–59 / 60–64 / 65–69 / 70 or older / No answer⟩
- (4) Please state your general education. ⟨Still in school / Lower secondary education / Polytechnic high school / Intermediary secondary education / University entrance qualification / No general school leaving certificate / Free text / No answer⟩
- (5) Please state your professional/vocational education. ⟨ Vocational training or training in dual system / Technical college diploma / Technical college diploma in the former GDR / Bachelor / Master / Diploma / PhD / No professional or vocational degree / Free text / No answer⟩
- (6) Are you currently a student enrolled in a degree program (Bachelor, Master, Diploma, State examination, Magister)? ⟨Yes / No / No answer⟩
- (7) Have you been using a smartphone for more than two years? ⟨Yes / No / No answer⟩

B.6 Technology

In the following questionnaire, we will ask you about your interaction with technical systems. The term “technical systems” refers to apps and other software applications, as well as entire digital

devices (e. g., mobile phone, computer, TV, car navigation). Please indicate the degree to which you agree/disagree with the following statements.

⟨Six responses from “Completely disagree” to “Completely agree”⟩

- (1) I like to occupy myself in greater detail with technical systems.
- (2) I like testing the functions of new technical systems.
- (3) I predominantly deal with technical systems because I have to.
- (4) When I have a new technical system in front of me I try it out intensively.
- (5) I enjoy spending time becoming acquainted with a new technical system.
- (6) It is enough for me that a technical system works; I don’t care how or why.
- (7) I try to understand how a technical system exactly works.
- (8) It is enough for me to know the basic functions of a technical system.
- (9) I try to make full use of the capabilities of a technical system.

Received July 2023; revised January 2024; accepted March 2024

PairSonic: Helping Groups Securely Exchange Contact Information

Citation	F. Putz, S. Haesler, T. Völkl, M. Gehring, N. Rollshausen, and M. Hollick. "PairSonic: Helping Groups Securely Exchange Contact Information." In: <i>Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work & Social Computing</i> (Nov. 9–13, 2024). CSCW '24. San José, Costa Rica: ACM, Nov. 2024, pp. 69–71. doi: 10.1145/3678884.3681818 .
Contribution	This publication is discussed in Chapter 4. My contribution is described in the section " <i>Collaborations and My Contribution</i> " on page xi.
Rights notice	© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is included in this dissertation in accordance with ACM author rights. The definitive Version of Record was published in <i>Companion of the 2024 Computer-Supported Cooperative Work and Social Computing</i> (CSCW Companion '24), DOI: https://doi.org/10.1145/3678884.3681818 .

PairSonic: Helping Groups Securely Exchange Contact Information

Florentin Putz
Technical University of Darmstadt
Darmstadt, Germany
fputz@seemoo.de

Steffen Haesler
Technical University of Darmstadt
Darmstadt, Germany
haesler@peasec.tu-darmstadt.de

Thomas Völkl
Technical University of Darmstadt
Darmstadt, Germany
tvoelkl@seemoo.de

Maximilian Gehring
Technical University of Darmstadt
Darmstadt, Germany
gehring@cs.tu-darmstadt.de

Nils Rollshausen
Technical University of Darmstadt
Darmstadt, Germany
nrollshausen@seemoo.de

Matthias Hollick
Technical University of Darmstadt
Darmstadt, Germany
mhollick@seemoo.de

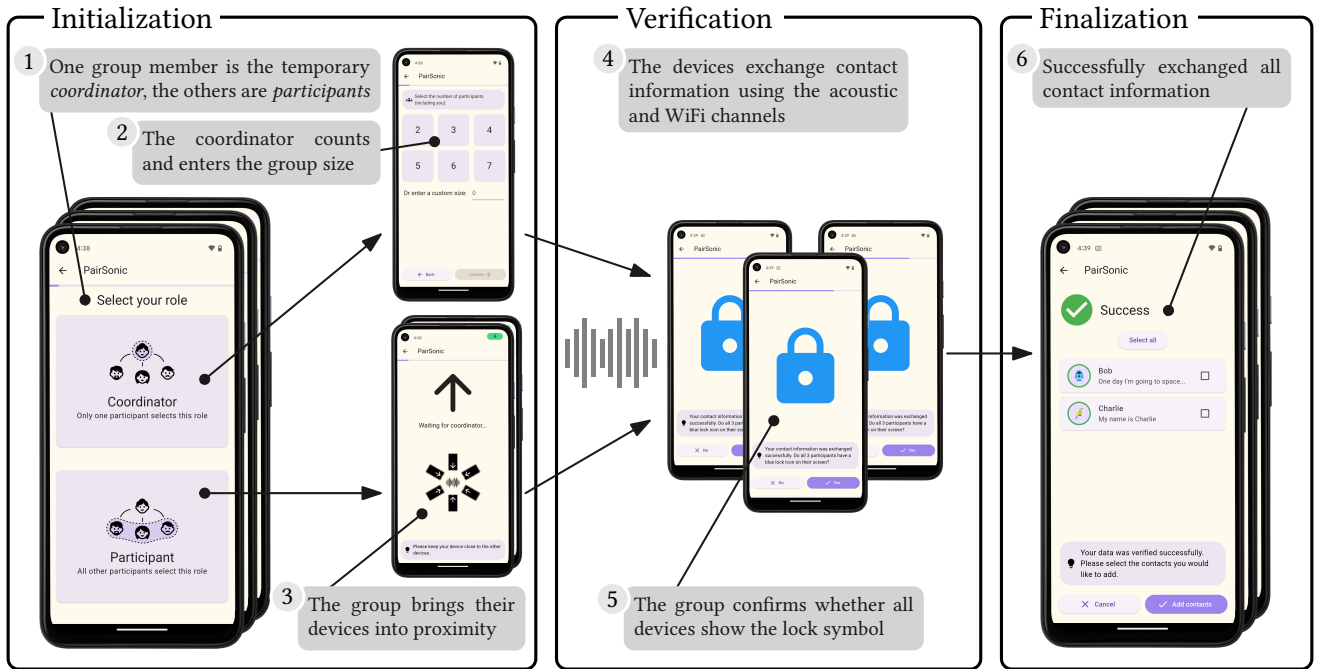


Figure 1: The PairSonic user interface, demonstrated for a group of three users wanting to exchange their contact information.

Abstract

Securely exchanging contact information is essential for establishing trustworthy communication channels that facilitate effective online collaboration. However, current methods are neither user-friendly nor scalable for large groups of users. In response, we introduce PairSonic, a novel group pairing protocol that extends trust from physical encounters to online communication. PairSonic simplifies the pairing process by automating the tedious verification tasks of previous methods through an acoustic out-of-band

channel using smartphones' built-in hardware. Our protocol not only facilitates connecting users for computer-supported collaboration, but also provides a more user-friendly and scalable solution to the authentication ceremonies currently used in end-to-end encrypted messengers like Signal or WhatsApp. PairSonic is available as open-source software: <https://github.com/seemoo-lab/pairsonic>

CSCW Companion '24, November 9–13, 2024, San Jose, Costa Rica.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Companion of the 2024 Computer-Supported Cooperative Work and Social Computing (CSCW Companion '24)*, November 9–13, 2024, San Jose, Costa Rica, <https://doi.org/10.1145/3678884.3681818>.

CCS Concepts

• **Human-centered computing** → **Computer supported cooperative work**; **Open source software**; **Collaborative interaction**; **Smartphones**; Mobile computing; • **Security and privacy** → **Authentication**; **Usability in security and privacy**.

Keywords

usable security; group pairing; secure device pairing; trust establishment; acoustic communication; data-over-sound; authentication ceremony; public-key cryptography; end-to-end encryption; instant messaging; online collaboration; contact management; key verification; peer-to-peer-authentication; MitM attacks; device association; binding; bonding; coupling; smartphone; prototype; security; privacy; social cybersecurity; interoperability; ad-hoc; decentral; spontaneous; nearby; proximity; open-source; WiFi; NFC; Android

ACM Reference Format:

Florentin Putz, Steffen Haesler, Thomas Völkl, Maximilian Gehring, Nils Rollshausen, and Matthias Hollick. 2024. PairSonic: Helping Groups Securely Exchange Contact Information. In *Companion of the 2024 Computer-Supported Cooperative Work and Social Computing (CSCW Companion '24)*, November 9–13, 2024, San Jose, Costa Rica. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3678884.3681818>

1 Introduction

Effective communication and collaboration require a trusted environment where participants clearly know each other's identities and do not have to fear unauthorized access [4]. This trust is vital for facilitating open discussions among friends, family, colleagues, business partners, and especially vulnerable groups like journalists and activists [2, 3, 7]. While end-to-end encrypted (E2EE) messengers like Signal support group collaboration through real-time messaging and video calls, they cannot, by default, protect against unauthorized access and impersonation by active adversaries. These threats include rogue service operators and machine-in-the-middle (MitM) attacks. Active MitM attacks in particular have become more prominent due to reports of law enforcement agencies compromising app providers, highlighting a practical and concerning risk for privacy infringements by malicious actors in the future [2].

To mitigate these risks, E2EE tools require users to perform a manual contact verification process known as the *authentication ceremony* [8]. Only after this ceremony is completed is the communication fully protected. The current design of authentication ceremonies in E2EE tools has two major shortcomings: First, they have low usability, requiring considerable user interaction, time, and effort [2, 4, 8]. Second, they do not support larger groups, functioning only between two users [6]. This limitation is significant because collaboration often involves multiple participants. Performing bilateral authentication ceremonies between each pair of group members is infeasible, as the number of required manual verifications scales quadratically with the group size.

To address both challenges, we designed PairSonic, a novel pairing protocol for the fast and secure exchange of contact information that scales efficiently to multiple participants. Our prototype implementation allows users to extend trust from physical encounters to online communication by simply holding their devices together for a few seconds, aligning with intuitive association behavior. This demo accompanies our full paper at CSCW 2024 [6], where we evaluated PairSonic's usability in a lab study, finding it to be highly effective for groups. PairSonic automates the tedious verification tasks required by previous approaches, greatly enhancing usability and making secure group communication more accessible.

2 System Design

PairSonic is a group pairing protocol that enables two or more users meeting in person to spontaneously exchange or verify their contact information. Each participant obtains authentic and verified contact details of all other participants, including their cryptographic public keys. Importantly, our protocol does not rely on external key management infrastructure, prior associations, or shared secrets.

2.1 Protocol

The PairSonic protocol is based on the cryptographic foundation of the SafeSlinger protocol [1], but improves its usability and deployability [6]. Our protocol operates in three main phases (Figure 1):

- (1) **Initialization:** A group of users initiates PairSonic on their smartphones to exchange contact data. One member is selected as the temporary *coordinator* while the others become *participants*. The coordinator inputs the total group size and ensures that all participants' smartphones are ready and within proximity. The coordinator's smartphone sets up a temporary ad-hoc WiFi network and transmits the network details over an acoustic out-of-band (OOB) channel, allowing the other smartphones to connect automatically.
- (2) **Verification:** Once connected, the devices use this WiFi network to perform a cryptographic protocol derived from the SafeSlinger protocol [1]. Participants send nested commitments, containing their contact data and cryptographic public keys, to the coordinator's smartphone. The coordinator aggregates these commitments into a hash value and transmits it over the acoustic OOB channel. All devices monitor this channel, ready to abort the process if an incorrect message is detected. Upon verifying the correct hash value, a blue lock symbol appears on each smartphone, prompting users to confirm the matching symbols on all other devices. If confirmed, the smartphones use the WiFi network to distribute success nonces, allowing the display of verified exchanged contacts. If the lock symbol confirmation fails, abort nonces are released, halting the protocol and ensuring the integrity of the contact exchange. We changed the verification symbol based on feedback from our user study, because some participants misunderstood the previous checkmark as indicating protocol completion without verifying the symbol.
- (3) **Finalization:** Each participant now sees a list of all received contacts and can select which to import.

2.2 Acoustic Out-Of-Band Channel

Previous user studies have indicated that performing authentication ceremonies with current E2EE tools like Signal or WhatsApp is often challenging due to the time and effort required [2, 4, 8]. Our protocol, PairSonic, addresses these issues by automating manual verification tasks such as comparing phrases, digits, or pictures. PairSonic utilizes an acoustic location-limited channel, leveraging the smartphones' built-in speakers and microphones to exchange data within a range of approximately one meter. This method simplifies the inclusion of multiple devices at once, a distinct advantage over near-field communication (NFC) or quick-response (QR) codes, which typically handle one device at a time. Moreover, unlike WiFi

Table 1: Selection of devices where we have verified compatibility with PairSonic, indicating support for audible and inaudible modes. In principle, any smartphone with WiFi Direct and running at least Android 6.0 (released in 2015) should be compatible.

Model	OS	SDK	Audible	Inaudible
Google Pixel 4	Android 13	33	✓	✓
Google Pixel 4a	Android 13	33	✓	✓
Google Pixel 5	Android 13	33	✓	✓
Google Pixel 6 Pro	Android 13	33	✓	✓
LG Nexus 5	Android 6.0.1	23	✓	✓
Motorola Edge	Android 10	29	✓	✓
OnePlus 10 Pro	Android 12	31	✓	✓

or Bluetooth, the acoustic physical layer in our system is inherently location-limited and software-defined, allowing for the incorporation of advanced physical layer security techniques [5]. These enhancements can significantly strengthen the protocol’s defense against sophisticated attacks on the acoustic channel, thus providing an additional layer of security for group pairing.

3 PairSonic Prototype

Our prototype of PairSonic is implemented as an Android app using the Flutter framework. For communication, we utilize two channels: a local ad-hoc WiFi connection, provided by the Android WiFi Direct API, and the acoustic OOB channel for data transmission via sound waves, provided by the ggwave library.¹ During our user study [6], we employed audible frequencies ranging from 1875 Hz to 6375 Hz to evaluate user perceptions of this communication method. Participants expressed concerns about the appropriateness of audible noises in quiet settings like museums or libraries. In response, we developed a variant that uses inaudible near-ultrasonic frequencies between 15 000 Hz and 19 500 Hz. Our testing revealed that these inaudible frequencies not only mitigate the issue of noise in quiet environments but also enhance the robustness of the communication. This improvement addresses another concern raised in our lab study, which noted occasional failures in acoustic communication that required restarting the pairing process.

The combination of an acoustic channel with WiFi Direct in our prototype forms a novel communication stack. To ensure functionality and compatibility, we conducted tests using smartphones from various manufacturers (Table 1). The prototype functioned effectively on all devices tested that support WiFi Direct, provided they run at least Android version 6.0 (API level 23, released in 2015). PairSonic is interoperable, meaning that different device models within the exchange can interact seamlessly.

4 Applications

PairSonic is designed to assist users in establishing trustworthy digital communication channels, making it especially useful for forming ad-hoc collaboration groups among multiple users [3]. This makes PairSonic particularly adaptable to various computer-supported collaboration tools, including group work platforms, educational hubs, and video conferencing systems. Furthermore, our protocol could be integrated into E2EE messengers like Signal, WhatsApp, or Matrix, providing a more user-friendly authentication ceremony to support effective and trustworthy online collaboration.

¹Data-over-sound library ggwave: <https://github.com/ggerranov/ggwave>

Model	OS	SDK	Audible	Inaudible
Oppo Reno 6	Android 11	30	✓	✓
Redmi Note 9 Pro	Android 10	29	✓	✓
Samsung Galaxy A32	Android 12	31	✓	✓
Samsung Galaxy A52s 5G	Android 14	34	✓	✓
Samsung Galaxy S10e	Android 12	31	✓	✓
Samsung Galaxy S20 Ultra 5G	Android 10	29	✓	✓
Samsung Galaxy S22	Android 13	33	✓	✓

Our study [6] highlights these potential applications: participants identified multiple scenarios where PairSonic could serve as a secure method for contact exchange. These include forming private chat groups on Signal, managing participants in Zoom video conferences, and enhancing interaction within online collaboration tools.

Availability

Together with this paper, we provide an overview of the PairSonic project online at <https://seemoo.de/s/pairsonic>. PairSonic is available as open-source software on GitHub: <https://github.com/seemoo-lab/pairsonic>

Acknowledgments

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

References

- [1] Michael Farb, Yue-Hsun Lin, Tiffany Hyun-Jin Kim, Jonathan McCune, and Adrian Perrig. 2013. SafeSlinger: Easy-to-Use and Secure Public-Key Exchange. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom '13)*. Association for Computing Machinery, New York, NY, USA, 417–428. <https://doi.org/10.1145/2500423.2500428>
- [2] Matthias Fassel and Katharina Krombholz. 2023. Why I Can’t Authenticate – Understanding the Low Adoption of Authentication Ceremonies with Autoethnography. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Hamburg, Germany. <https://publications.cispa.saarland/3895/>
- [3] Moses Namara and Bart P. Knijnenburg. 2021. The Differential Effect of Privacy-Related Trust on Groupware Application Adoption and Use during the COVID-19 Pandemic. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2, 405:1–405:34. <https://doi.org/10.1145/3479549>
- [4] Sean Oesch, Ruba Abu-Salma, Oumar Diallo, Juliane Krämer, James Simmons, Justin Wu, and Scott Ruoti. 2022. User Perceptions of Security and Privacy for Group Chat. *Digital Threats: Research and Practice* 3, 2 (Feb. 2022), 15:1–15:29. <https://doi.org/10.1145/3491265>
- [5] Florentin Putz, Flor Álvarez, and Jiska Classen. 2020. Acoustic Integrity Codes: Secure Device Pairing Using Short-Range Acoustic Communication. In *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '20)*. 31–41. <https://doi.org/10.1145/3395351.3399420>
- [6] Florentin Putz, Steffen Haesler, and Matthias Hollick. 2024. Sounds Good? Fast and Secure Contact Exchange in Groups. In *Proceedings of the ACM on Human-Computer Interaction* 8 (CSCW2). Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3686964>
- [7] Erica Shusas, Patrick Skeba, Eric P. S. Baumer, and Andrea Forte. 2023. Accounting for Privacy Pluralism: Lessons and Strategies from Community-Based Privacy Groups. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3544548.3581331>
- [8] Elham Vaziripour, Justin Wu, Mark O’Neill, Daniel Metro, Josh Cockrell, Timothy Moffett, Jordan Whitehead, Nick Bonner, Kent Seamons, and Daniel Zappala. 2018. Action Needed! Helping Users Find and Complete the Authentication Ceremony in Signal. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security (SOUPS '18)*. USENIX Association, USA, 47–62.

Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices

Citation	F. Putz, P. Fortmann, J. Frank, C. Haugwitz, M. Kupnik, and M. Hollick. "Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices." In: <i>ACM Transactions on Internet of Things</i> 7.1, Feb. 2026. DOI: 10.1145/3779439
Contribution	This publication is discussed in Chapter 4. My contribution is described in the section " <i>Collaborations and My Contribution</i> " on page xi.
Rights notice	© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is included in this dissertation in accordance with ACM author rights. The definitive Version of Record was published in <i>ACM Transactions on Internet of Things</i> , DOI: https://doi.org/10.1145/3779439 .

Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices

FLORENTIN PUTZ, PHILIPP FORTMANN, JAN FRANK, CHRISTOPH HAUGWITZ, MARIO KUPNIK, and MATTHIAS HOLLICK, Technical University of Darmstadt, Germany

Acoustic data transmission offers a compelling alternative to Bluetooth and NFC by leveraging the ubiquitous speakers and microphones in smartphones and IoT devices. However, most research in this field relies on simulations or limited on-device testing, which makes the real-world reliability of proposed schemes difficult to assess. We systematically reviewed 31 acoustic communication studies for commodity devices and found that none provided accessible source code. After contacting authors and re-implementing three promising schemes, we assembled a testbed of eight representative acoustic communication systems. Using over 11 000 smartphone transmissions in both realistic indoor environments and an anechoic chamber, we provide a systematic and repeatable methodology for evaluating the reliability and generalizability of these schemes under real-world conditions. Our results show that many existing schemes face challenges in practical usage, largely due to severe multipath propagation indoors and varying audio characteristics across device models. To support future research and foster more robust evaluations, we release our re-implementations alongside the first comprehensive dataset of real-world acoustic transmissions. Overall, our findings highlight the importance of rigorous on-device testing and underscore the need for robust design strategies to bridge the gap between simulation results and reliable IoT deployments.

CCS Concepts: • **Networks** → **Mobile ad hoc networks**; Short-range networks; Network reliability; **Network performance evaluation**; • **Human-centered computing** → **Ubiquitous and mobile devices**; **Smartphones**; • **Hardware** → Signal processing systems.

Additional Key Words and Phrases: acoustic communication; ultrasonic communication; near-ultrasonic; audio-based communication; data-over-sound; reproducibility; conceptual replication; generalizability; dataset; open data; open source; audio recordings; iot; ubiquitous computing; experimental testbed; systematic literature study; re-implementation; performance evaluation; device-to-device communication; smartphone-to-smartphone communication; short-range; ad-hoc; interoperability; spontaneous; nearby; proximity

ACM Reference Format:

Florentin Putz, Philipp Fortmann, Jan Frank, Christoph Haugwitz, Mario Kupnik, and Matthias Hollick. 2026. Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices. *ACM Trans. Internet Things* 7, 1, Article 8 (February 2026), 31 pages. <https://doi.org/10.1145/3779439>

1 Introduction

Smartphones and Internet of Things (IoT) devices can communicate using sound waves via their built-in speakers and microphones [39]. Acoustic data transmission promises to be a practical alternative to Bluetooth and near-field communication (NFC), given the universal presence of audio I/O in devices like smartphones, smartwatches, smart speakers, and home automation systems. Over the past decade, there has been increasing interest in using acoustic communication as a wireless, short-range ad-hoc channel to facilitate ubiquitous and mobile computing use cases, such as mobile payments [68], context-aware computing [40, 41], or location-based services [45, 71]. Its location-limited nature also fits well with ubiquitous computing scenarios, where users intuitively

Authors' Contact Information: Florentin Putz, fputz@seemoo.de; Philipp Fortmann, pfortmann@seemoo.de; Jan Frank, jfrank@seemoo.de; Christoph Haugwitz, christoph.haugwitz@tu-darmstadt.de; Mario Kupnik, mario.kupnik@tu-darmstadt.de; Matthias Hollick, mhollick@seemoo.de, Technical University of Darmstadt, Darmstadt, Germany.

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM Transactions on Internet of Things*, <https://doi.org/10.1145/3779439>.

bring devices into proximity so they can “talk” [11], aligning with intuitive user behavior [6, 25, 26]. The widespread availability of low-cost IoT hardware equipped with speakers and microphones has even led to commercial uptake, such as the smart-speaker manufacturer Sonos employing acoustic communication for spatial awareness and device pairing [27].

Despite the growing body of research proposing acoustic data transmission systems [4, 13, 14, 33, 38, 70], many have been tested only through simulations or under very limited real-world conditions. Real-device testing, however, is crucial to account for factors like hardware imperfections, user handling, realistic signal propagation, and background noise—factors that are notoriously difficult to model precisely. Moreover, acoustic channels differ substantially from radio channels: sound waves travel more than 87 000 times slower than electromagnetic waves, leading to delay spreads on the order of tens of milliseconds [33, 59]. Acoustic signals also suffer from highly variable ambient noise and, due to low sample rates on most smartphones and IoT devices, bandwidth is limited to at most 22 kHz, making high-throughput transmissions challenging. Existing evaluations typically focus on individual devices or scenarios, providing little insight into how well a scheme might generalize across diverse real-world settings.

At the same time, there is a distinct gap between academic research and commercial practice. Commercial products for acoustic data transmission on smart devices typically achieve only 10 – 100 bps [9, 13, 27] and occasionally up to 200 bps [62], prioritizing robustness and reliability in noisy environments. In contrast, researchers often report much higher throughput up to 500 – 10 000 bps [4, 14, 70] or beyond [69]. This discrepancy raises fundamental questions about whether these systems can maintain reliability in real-world scenarios, or whether additional constraints—such as hardware limitations, user behavior, or environmental interference—ultimately force practical systems to operate at lower data rates.

A major barrier to resolving these questions is that publications in this field often omit source code, limiting independent validation and hindering follow-up work. In this paper, we take a first step toward consolidating existing findings by investigating how well representative acoustic data transmission schemes generalize to various practical scenarios. Specifically, we address the following research questions:

RQ1: “How challenging is it to obtain or re-implement systems proposed in this field?”

RQ2: “How well do these schemes generalize to ad-hoc smartphone communication use cases?”

RQ3: “Which practical challenges affect nearby acoustic data transmission between smart devices?”

1.1 Contributions

This paper presents the first **independent evaluation of the generalizability of acoustic data communication schemes** proposed by other researchers, comparing them in realistic settings. Our contributions are as follows:

- **Systematic literature study.** We conducted a systematic review and identified 31 publications proposing acoustic data transmission systems for commercial off-the-shelf (COTS) devices (Section 4). Regrettably, none of these studies provide implementation source code.
- **Sourcing and re-implementation of representative schemes.** We reached out to the authors for software implementations. Only three authors supplied functioning code, one provided a non-functioning implementation, and the others either did not respond or had lost access. To better understand the generalizability of acoustic communication research, we additionally re-implemented three promising schemes (Section 5).
- **Evaluation of generalizability.** In total, we evaluated the practical reliability of eight acoustic communication schemes [4, 13, 14, 33, 38, 70], including two variants of the popular

open-source project ggwave [12] (Section 6). Our testbed focuses on diverse smartphone models as a practical and accessible subset of IoT devices, although the underlying insights hold for any device with comparable audio input and output capabilities. We expand prior work by evaluating these schemes across various IoT use cases and analyzing how factors such as device distance, model, user handling, and background noise affect communication reliability (measured by the rate of errors in transmissions).

- **Measurement dataset.** We publish our re-implementations and analysis code along with the first dataset in this field—11 900 real recordings of acoustic data transmissions—to support transparent, reproducible research [51].

1.2 Scope

This work supports the IoT research community by analyzing acoustic data transmission schemes from the perspective of researchers or developers who want to integrate this technology into smart devices for nearby data exchange. From this perspective, acoustic communication should be as seamless to adopt as Bluetooth or NFC, requiring no specialized knowledge of the underlying wireless standards. Accordingly, we evaluate representative schemes on multiple smartphone models, which serve as the most common and versatile IoT form factor. Rather than focusing on physical-layer performance or simulations, we treat the complete communication system as a black box, including all components—such as synchronization and error correction—as originally described by the authors, ensuring we analyze its practical suitability on real devices.

2 System model

This study evaluates *acoustic data transmission schemes*: software implementations that broadcast data from one device to nearby devices using sound waves. From a system developer’s perspective, each scheme acts as a black box that transmits data on one device and receives it on another. For our evaluation, we model this capability as an interface with two functions, depicted in Figure 1:

- (1) A **transmitter**, $wav = TX(b_{TX})$, which converts data bits into a WAV file for playback on a loudspeaker.
- (2) A **receiver**, $b_{RX} = RX(wav)$, which extracts data from a WAV file containing the microphone recording of the transmission.

This interface separates the digital signal processing of the modulator and demodulator from audio playback and recording. Although some schemes handle playback and recording internally, we treat them independently for clarity. In practice, a developer might not store intermediate WAV files but instead stream samples directly from a buffer. Nonetheless, our interface fundamentally reflects real-world usage and facilitates a fair comparison of each scheme under uniform playback and recording conditions.

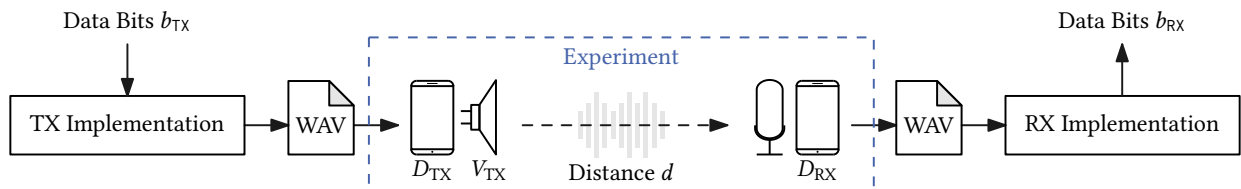


Fig. 1. **Acoustic data transmission model.** In our experiments, we play the WAV files from the TX implementation on device D_{TX} (at volume V_{TX}) and record this on device D_{RX} . The resulting WAV file gets decoded using the RX implementation.

2.1 Scenarios

We focus on the use case of data exchange between nearby smart devices, where software modifications are possible (e.g., by installing apps on smartphones) but changes to hardware or firmware are not. Our emphasis is on broadcast schemes that enable communication without requiring special arrangements with the communication partner. We divide common use cases for acoustic communication (Section 3) into three categories based on the communication distance—a key factor affecting usability. The practical difference in usability is significant, ranging from the need for devices to touch (as in NFC) to the flexibility of holding your device several meters away from your communication partner. Throughout this paper, we focus on these concrete communication scenarios, which represent most current use cases:

- (1) **Near distance** ($d \leq 10$ cm): Transmitting a public key or a small file (≈ 4096 bit) over a short distance, similar to NFC. Examples are mobile payments [8, 63], secure device pairing [31, 64, 70], and access control [8, 45].
- (2) **Medium distance** ($10 \text{ cm} < d \leq 1$ m): Transmitting a public key fingerprint or a hash value (≈ 128 bit) over a medium distance. Examples are IoT device provisioning [27, 58], side channels for air-gapped devices [54], and control channels for personal electronic devices and hearing aids [43].
- (3) **Far distance** ($d > 1$ m): Transmitting a short ID (≈ 16 bit) over larger distances. Examples are providing location-specific information near public signs or in shops [10], indoor IoT communication [18, 23], and enhancing media experiences, such as social television [36] and second screen services via hidden audio data in TV and radio [28].

3 Related work

In this section, we briefly review prior research on acoustic data transmission before presenting our systematic literature study in Section 4.

Although smartphone and IoT audio hardware is primarily intended for media playback and voice communication, it can also be repurposed for sensing [5, 66, 72] or even for aerial acoustic data transmission. In the latter case, speakers encode information in sound waves using frequency or amplitude variations, similar to electromagnetic communication [39], thereby enabling data exchange across a diverse group of devices, including smartphones and voice-controlled appliances, without special radio frequency (RF) hardware [11]. This communication channel is limited by the typical 0 – 22 kHz audio range (due to common sampling rates of 44.1 kHz), pronounced frequency selectivity [21], and non-linearities at higher volumes (see Section 7.3). Data can be transmitted audibly or using near-ultrasonic frequencies, and—unlike Wi-Fi or Bluetooth—no network setup or pairing is required, allowing ad-hoc broadcasts over several meters. Because it is entirely software-defined, acoustic communication can be easily deployed on existing devices and tailored to support a variety of communication needs [38].

Research in ubiquitous and mobile computing has explored acoustic data transmission for context-aware computing [40, 41] and location-limited communication between nearby devices [1, 67]. The limited range has also attracted interest for secure device interactions [23, 52], including device pairing [15, 16, 50, 53, 70] and protecting implantable medical devices [17]. User experience aspects of acoustic transmissions have been evaluated [38, 39, 42, 52], as the audibility of the data exchange is a new aspect distinguishing it from previous wireless technologies. Various commercial applications leverage acoustic data transmission for mobile payments [8, 63], context-aware applications [37], location-based services [9, 60], data transfer [62], and device pairing [31, 64]. Despite its application potential, no previous work has attempted to systematically evaluate these systems on real devices. The absence of published source code hinders replication of prior results.

4 Literature study

We conducted a systematic literature survey to gather all papers that propose new aerial acoustic software-defined communication systems, according to the following scope.

4.1 Scope

To ensure our research aligns with ubiquitous computing and everyday consumer interactions, we focused our literature survey on papers that describe practical implementations of aerial acoustic device-to-device communication specifically for COTS devices, like smartphones. We excluded studies related to other types of acoustic communication, such as underwater or industrial uses, and those that are not applicable to general-purpose data transmission systems.

4.2 Method

To identify relevant literature on acoustic data transmission schemes, we began by collecting a wide array of papers related the subject. We then refined this collection to include only those with research implementations for our target use case. Our method was guided by PRISMA [48], which is a framework for assessing large bodies of medical research in a structured manner. PRISMA provides a flowchart with three steps:

- (1) **Identification:** Define the data sources and search strategy.
- (2) **Screening:** Filter the research according to the desired scope.
- (3) **Inclusion:** Select which papers to include.

4.2.1 Identification. As the information source for this literature study we used the Web of Science (WoS),¹ which is a comprehensive database indexing journals and conference proceedings. We selected WoS due to its extensive, regularly updated catalog and its advanced filtering capabilities with nested logical operators, which are crucial for narrowing the results to our specific scope. Our search strategy balanced breadth and manageability. We designed and executed two mutually exclusive search queries² on May 18, 2023, yielding 631 papers. These queries targeted papers published between 2000 and 2022 in the fields of acoustics, computer science, and engineering, focusing on relevant keywords in the titles. We excluded papers related to nautical or medical fields by blocking specific keywords associated with underwater communication and ultrasonic imaging. The second query broadened the search to include a wider range of keyword combinations. We refer to the query URLs² for details and keywords.

4.2.2 Screening. We processed all identified papers to determine their relevance to our scope. In the first iteration, we reviewed the titles and abstracts of all 631 papers and eliminated 422 papers that were clearly out of scope—such as those focusing on speech-based communication, tracking with audio, or communication through metal barriers. In a second iteration, we examined the full text of the remaining 209 papers, classifying them based on their applicability to ubiquitous computing use cases and whether they describe a testable implementation for COTS devices. We excluded papers dependent on specialized hardware, like microphone or speaker arrays for multiple-input and multiple-output (MIMO) communication, and those embedding data in music, unless the scheme could be utilized as a general-purpose communication system.

4.2.3 Inclusion. In total, our literature study identified 31 papers that directly align with our study's scope and describe a testable communications scheme (as shown in Table 1).

¹<https://clarivate.com/products/web-of-science/>, last access: 2025-03-03

²Search query 1: <https://www.webofscience.com/wos/woscc/summary/4c89e96d-5de5-4543-9167-23d73086d8e6-8a0aeac3/relevance/1>; search query 2: <https://www.webofscience.com/wos/woscc/summary/392d145e-a4ac-4aff-8995-ad2103dd116b-898dca59/relevance/1>.

Table 1. **Surveyed acoustic communication schemes**, sorted by throughput. The right table continues from the left. Data not available in the publications is marked with “?”.

Name	Modulation	Frequencies [kHz]	Inaudible	Throughput [bps]	Name	Modulation	Frequencies [kHz]	Inaudible	Throughput [bps]
Moriyama et al. [44]	ASK	~19	✓	5	Gonçalves et al. [14]	OFDM + PSK	0 – 4	✗	600
Kim et al. [30]	16-FSK	18 – 20	✓	5	PriWhisper [70, 71]	MFSK	9 – 17	✗	1000
Ka et al. [28]	Chirp	18.5 – 19.5	✓	15	DigitalVoices [38]	ASK, (M)FSK, SS	1 – 10	✗	1280
Lee2015 [33]	Chirp	19.5 – 22	✓	16	A-NFC [35]	QPSK	16–21	✓	2200
Lee2020 [34]	Chirp	18.5 – 19.5	✓	16	Dhwani [45]	OFDM + PSK	6 – 7	✗	2400
Hornych et al. [20]	ASK, FSK	18 – 20	✓	33	Hush [47]	OFDM + QAM	17.5 – 21	✓	4900
TagScreen [36]	Chirp	18 – 20	✓	36	He et al. [19]	AM, FSK	0 – 22	✗	6000
Haus et al. [18]	OFDM + ASK	22 – 24	✓	64	Yamamoto et al. [69]	16-QAM	?	?	32000
Miegel et al. [43]	OOK, FSK	16 – 20	✓	84	Arentz et al. [1]	ASK	20–23	✓	?
Rustle [24]	<i>custom</i> ¹	0 – 15	✗	90	Bits-over-Sound [29]	DBPSK	?	✗	?
Nearby [13]	DSSS + MFSK	18.5 – 20	✓	95	Jeon et al. [22]	PSK	20	✓	?
Bang et al. [3]	FSK	18 – 22	✓	170	No Entry [32]	Chirp	?	?	?
Sadeq et al. [54]	FSK	18 – 19	✓	440	Sasano et al. [56]	OFDM + PSK	6.4 – 8	✗	?
HRCSS [4]	Chirp	18 – 22	✓	500	SilentInformer [57]	FSK	17 – 19	✓	?
Dolphin [65, 73]	OFDM + ASK	8 – 20	✗	500					

¹ Rustle’s custom approach is called “furtive modulation”.

5 Obtaining acoustic data transmission implementations

To independently evaluate existing acoustic communication systems and embed them in IoT testbeds, we first needed to obtain functional implementations. We applied three different methods to acquire the implementations (see Section 7.1 for a discussion of the associated challenges).

5.1 Publicly available implementations

None of the 31 papers surveyed referenced a publicly available implementation.

5.2 Author requests

We contacted every author via email as shown in Figure 2. Only six replied and just three of those responses provided a functional implementation:

- **Digital Voices**, proposed by Lopes et al. [38] in 2001, is among the earliest documented aerial acoustic communication schemes. It uses binary amplitude-shift keying (ASK) across eight carrier frequencies. The authors provided their Java source code.
- **HRCSS**, introduced by Cai et al. [4], is a chirp-based communication scheme that uses multiple chirp carriers simultaneously. The MATLAB source code provided to us lacks their bidirectional probing feature, so we only evaluate the data transmission.
- **Gonçalves et al.** [14] proposed a scheme using orthogonal frequency-division multiplexing (OFDM) and phase-shift keying (PSK) on each subcarrier, incorporating a short chirp sequence for synchronization. The authors provided us with the MATLAB source code.

A fourth response (Hush [47]) was unusable on current Android versions.³

³The authors of Hush [47] shared their Android implementation, but we encountered fatal runtime errors that hindered signal transmission and recording. These issues were likely due to changes in the Android ecosystem and incomplete code.

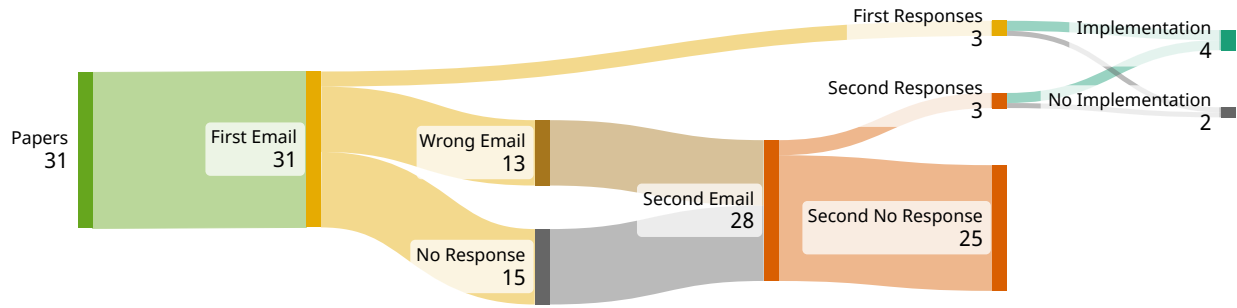


Fig. 2. **Responses from authors** when we requested implementations for 31 publications on acoustic data transmission. We contacted the first authors and all co-authors. For 13 papers, the provided email addresses were outdated, prompting us to search for current contact details manually. We followed up if there was no reply after three weeks. Of the four implementations received, we evaluated three due to runtime errors in the fourth. The other responding authors reported no longer having access to the implementations, citing reasons such as the lead author departing the university or hardware failures.

5.3 Re-implementing acoustic data transmission schemes

We re-implemented the schemes from three highly cited papers, chosen to represent our three target use cases, each with distinct throughput and maximum transmission distances:

- **Lee et al.** [33] claim to be the first to create a prototype that uses chirps in practical acoustic communication, targeting the far-distance use case. In their system, each communication frame begins with a long chirp preamble, followed by 16 individual chirps.
- **Nearby** was proposed by Getreuer et al. [13] and has been integrated into several Google products, including Google Nearby. It is tailored for the medium-distance use case and employs a combination of multiple frequency-shift keying (MFSK), direct-sequence spread spectrum (DSSS), and single-sideband modulation (SSB).
- **PriWhisper**, introduced by Zhang et al. [70], is an acoustic communication system for the nearby use case, akin to NFC. It utilizes MFSK and includes friendly jamming for physical-layer security. However, in our study, we focus solely on the communication capabilities of PriWhisper and have not implemented its security features.

We validated our re-implementations using simulations and practical experiments, ensuring that they perform similar to their original evaluation results. The system proposed by Lee et al. [33] is thoroughly detailed, enabling us to implement all aspects as described in their paper. The other two required filling in undocumented details (see Appendix B). We publish our MATLAB re-implementations as part of our replication package [51].

5.4 Summary

We ultimately obtained a subtotal of six academic implementations for state-of-the-art acoustic communication systems. Additionally, we included two non-academic implementations for comparison from the open-source project *ggwave*,⁴ in both its audible (*ggwave_a*) and inaudible (*ggwave_i*) configurations. Thus eight schemes advance to evaluation (Table 2).

The project has not been updated since 2017 and the authors acknowledged in their correspondence that they are aware of the lack of maintenance. Despite this, we still commend Novak et al. [47] for making their source code available.

⁴Most-starred “data-over-sound” project on GitHub: <https://github.com/topics/data-over-sound> (as of 2025-03-03).

Table 2. **Evaluated acoustic communication schemes**, sorted by throughput. The first column specifies the source of the implementation: P – referenced in the paper; A – by contacting the authors; R – by re-implementing them ourselves; N – non-academic solution for comparison. In the last three columns, we classify each scheme’s a priori suitability for the three use case categories defined in Section 2.1, based on their specified throughput and maximum distance.

Via	Name	Modulation	Frequencies	Inaudible	Throughput ¹	Near	Medium	Far
R	Lee et al. [33]	Chirp BOK	19.5–22 kHz	✓	15 bps	✗	✗	✓
A	Digital Voices [38]	MFSK	1–3 kHz	✗	80 bps	✗	✓	– ²
R	Nearby [13]	DSSS + MFSK	18.5–20 kHz	✓	84 bps	✗	✓	✓
N	ggwave _a [12]	MFSK	1.8–6.3 kHz	✗	268 bps	✗	✓	– ²
N	ggwave _i [12]	MFSK	15–19.5 kHz	✓	268 bps	✗	✓	– ²
A	HRCSS [4]	OCSS	18–22 kHz	✓	500 bps	✓	✓	✓
A	Gonçalves et al. [14]	OFDM + QPSK	0.1–4 kHz	✗	600 bps	✓	✗	✗
R	PriWhisper [70]	MFSK	9–17 kHz	✗	729 bps	✓	✗	✗

¹ We list the net data rate, as this is the usable data rate after error correction from the user’s perspective.

² Maximum distance not specified, therefore no a priori assessment.

6 Evaluation

In this section, we systematically evaluate the generalizability and practical suitability of acoustic data transmission across various use cases. For instance, indoor IoT communication [18] requires reliable transmission over distances of several meters, whereas mobile payment systems [8, 63] operate effectively over just a few centimeters. When integrated into smartphone applications (e.g., Google’s Nearby system [13]), acoustic data transmission must accommodate diverse smartphone models with varying audio hardware and frequency responses. Furthermore, transmissions occur under different environmental conditions and must remain robust against various background noises. Our evaluation addresses these considerations systematically.

6.1 Method

We conducted four experiments in which we varied the distance (Section 6.3), device models (Section 6.4), noise interference (Section 6.5), and transmission environment (Section 6.6). We treated each implementation as a black box according to our system model (Figure 1), using it as designed by the authors with their specified parameters.⁵ Our goal was to evaluate each scheme as a complete communication system from a *link layer* (layer two) perspective, including framing, synchronization, and error correction.

We randomly chose the data b_{tx} for transmission. The transmitter converted this data into audio and could utilize the 0–24 kHz baseband freely. Each WAV file was normalized to a specific maximum amplitude before transmitting it from one smartphone to another. The received audio was recorded into a WAV file and processed by the receiver implementation, which demodulated the signal to output data b_{rx} . We then verified whether b_{rx} matches b_{tx} . This procedure allowed us to treat each scheme as a black box, capable of transmitting data from one device to another. Each measurement was repeated at least 20 times for each scheme [22], while the scheme from Lee et al. [33] was repeated 100 times due to its smaller transmission duration. We conducted additional measurements in case of measurement errors. Our experimental setup is shown in Figure 3 and is

⁵Most of the original papers only present one set of PHY parameters. There are two exceptions: (1) PriWhisper [70], which varied the number of frequencies in an M-FSK scheme—we used the highest number to match their claimed throughput; and (2) DigitalVoices [38], which varied PHY parameters solely to assess sound pleasantness, which is outside our scope.

described in detail in [Appendix A](#) to provide further transparency, including full hardware and software configurations.

6.1.1 Metrics. To assess reliability, we introduce the total error rate (TER), a link-layer metric that combines elements of the bit error rate (BER) and the packet error rate (PER), allowing more direct comparisons across different schemes. We also provide the underlying raw BER data in the supplementary material. Under our system model ([Figure 1](#)), the TER measures the final error rate of received data bits *after* error correction by comparing b_{rx} with b_{tx} . If the receiver returns an error (e.g., from synchronization failures) instead of data, we treat the transmission as having a 100% TER.⁶ This classification treats such transmissions equal to transmissions where every bit was decoded incorrectly, as both instances represent completely unusable transmissions from the user's perspective. Previous work on acoustic data transmission often reports only BER [[14](#), [22](#), [57](#)] or PER [[13](#), [70](#)], but each has drawbacks. BER overlooks transmissions that fail to decode, allowing a possible 0% BER even when most transmissions are undecodable, and PER treats a single bit error the same as a thousand bit errors. By combining both metrics, TER better reflects real-world reliability, accounting for individual bit errors as well as complete transmission failures.⁷

6.1.2 Parameters. We conducted four experiments in which we systematically varied the following parameters to determine their impact on reliability. These parameters are the most commonly evaluated ones in prior research using real devices [[13](#), [14](#), [33](#), [70](#)].

- **Distance** ([Section 6.3](#)). The transmission distance d significantly affects suitability for our designated use case categories. For instance, a scheme failing to transmit reliably over multiple meters is unsuitable for the far-distance use case. We set a default distance of $d = 50$ cm, typical for medium-distance use cases.
- **Device models** ([Section 6.4](#)). The transmitter D_{TX} and receiver D_{RX} influence transmission reliability, as the audio hardware in smart devices is not designed for data transmission. Each device has a different frequency response, particularly at near-ultrasonic frequencies [[21](#)]. We selected the Google Pixel 4a smartphone as the default transmitter D_{TX} due to its relatively low distortion at a volume setting⁸ of $V_{TX} = 19/25$, which we experimentally identified to be a good trade-off between high volume and low distortions due to amplifier and speaker nonlinearities. The default receiver D_{RX} was a Samsung Galaxy S20 Ultra smartphone, chosen for its low self-noise.
- **Noise interference** ([Section 6.5](#)). Practical transmissions often occur under suboptimal conditions, such as loud ambient noise or interference from how users handle their smartphones. By default, our testing occurred in a quiet environment without external noise.
- **Environment** ([Section 6.6](#)). The environment influences the multipath profile of the acoustic channel. By default, we evaluated the schemes in a typical indoor setting using quiet office rooms and a hallway for long-distance measurements. Because these environmental

⁶Our measurements show that error rates can exceed 50% when significant portions of the received message are missing. To correctly penalize complete failures—when no message is received—we assign 100% TER.

⁷In summary, the TER is equal to the BER unless the receiver returns an error instead of data, in which case the TER is 100%. Illustrative examples (assuming 100 data bits per frame):

- (1) Receiving 99 correct bits yields a TER of 1% (identical to the BER).
- (2) Failing to decode entirely yields a TER of 100% (where the BER is undefined).
- (3) Receiving all bits incorrectly also yields a TER of 100% (identical to the BER).
- (4) Across 50 transmissions in case (1) and 50 in case (2), the average TER is 50.5%, whereas BER is misleadingly low at 1%, and PER is misleadingly high at 100%.

⁸On smartphones, the volume can typically only be set to one of a few discrete volume indices, where the maximum index varies per device.

conditions cannot be replicated exactly, we also conducted tests in an anechoic chamber (Figure 3) without reverberation or ambient noise, offering a controlled and reproducible testing environment. Finally, we also tested a large lecture room with mechanical ventilation. The background noise levels averaged 53.60 dB SPL in the office, 60.01 dB SPL in the lecture room, and 27.65 dB SPL in the anechoic chamber (all unweighted RMS).

6.2 Preliminary testing

6.2.1 Setup. Before extensive smartphone-based tests, we performed a preliminary evaluation under near-optimal conditions to confirm each implementation's functionality. Using high-quality audio equipment, each scheme was transmitted by a Neumann KH 80 DSP studio speaker and recorded by an Earthworks M23R reference microphone, positioned on-axis at 50 cm in an office room (see Section A.4.1 for details).

6.2.2 Results. The scheme by Lee et al., DigitalVoices, Nearby, PriWhisper, and both ggwave variants achieved a perfect 0% TER (Figure 4). The scheme from Gonçalves et al. was mostly accurate but showed an error rate of 0.8%, likely due to the scheme being designed for closer distances. However, the HRCSS implementation produced extremely high TER, which was unexpected given the high-quality equipment. We unfortunately had to conclude that this implementation is unsuitable for practical tests and therefore excluded it from further testing (see discussion in Section 7.4.3).

6.3 Impact of different distances

6.3.1 Setup. Next, we evaluated each scheme over distances from 5 cm to 40 m. Tests were conducted in both a quiet indoor setting (representative of everyday use) and an anechoic chamber with minimal reverberation and noise. In both environments, we measured distances up to 5 m; for distances beyond 5 m (up to 40 m), we used the hallway adjacent to the office.

6.3.2 Results. In the anechoic chamber (Figure 5a), most schemes achieved low error rates at short distances. The scheme from Lee et al. maintained a zero error rate across all tested distances. Both ggwave schemes performed nearly perfectly, with slight exceptions at 2 m and 5 m. Nearby also excelled, aside from a small outlier at 50 cm. Gonçalves et al.'s scheme kept error rates under 0.5% for distances below 10 cm, but rose to around 13% at longer distances. DigitalVoices and PriWhisper stayed below 1% at close range, then deteriorated notably beyond 20 cm.

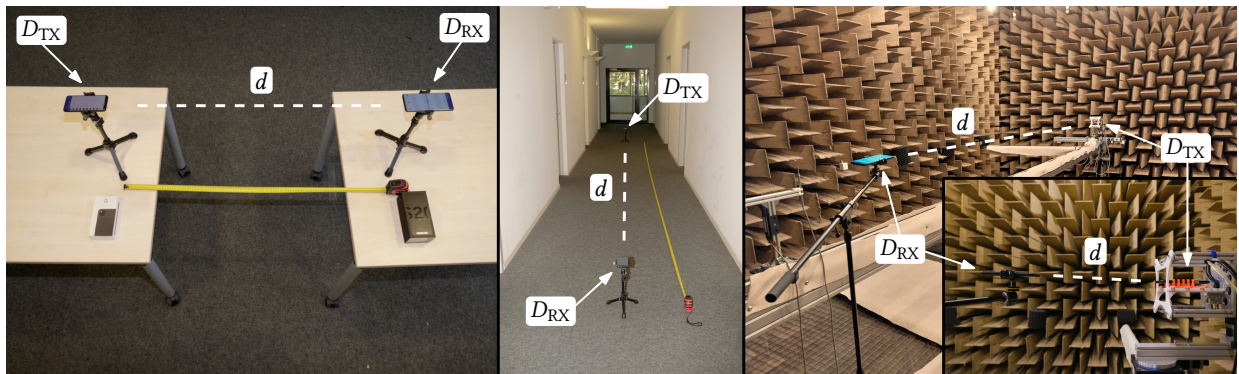


Fig. 3. **Experimental setup.** The left picture shows our office setup, with device D_{TX} transmitting to device D_{RX} at distance d . The middle picture shows our setup in the office hallway for larger distances from 5 m to 40 m. The right pictures show our setup in the anechoic chamber at two different distances.

In the *office environment* (Figure 5b), error rates increased noticeably due to stronger multipath effects and ambient noise. DigitalVoices, PriWhisper, Nearby, and Gonçalves et al.'s scheme all showed a strong distance dependency, degrading over longer ranges. At short distances up to 10 cm, however, all schemes maintained a TER under 0.5% except DigitalVoices. Although DigitalVoices performed worse at close range than in the anechoic chamber, it surprisingly improved between 50 – 500 cm, showing that the DigitalVoices receiver is not yet fully optimized. We investigated the DigitalVoices discrepancy and observed that post-processing and amplifying the recordings significantly reduced its error rates in the anechoic environment. However, we did not include post-processed results in our plots to compare the schemes as described in their publications. Nearby performed well within 1 m, but became unreliable at larger distances, including an outlier at 10 m, likely due to the transition into the hallway. This non-monotonic distance behavior is consistent with Getreuer et al.'s original Nearby evaluation [13, Fig. 13], which attributes such variability to nontrivial multipath behavior. Lee et al.'s scheme performed exceptionally across all distances, with only a minor outlier at 5 m. The inaudible ggwave variant remained effective up to 20 m, whereas the audible variant only worked up to 1 m before failing to decode.

6.4 Impact of different device models

6.4.1 Setup. This experiment assessed how well acoustic data transmission schemes perform across a broad set of smartphones in a typical indoor environment. We selected five smartphones representing a broad spectrum of release dates, prices, and manufacturers (Table 3). We tested each scheme using 21 different transmitter-receiver combinations. Our setup involved one smartphone as the transmitter, positioned 50 cm away on-axis from two to three simultaneous receivers to make the measurement process more efficient. As we had two Pixel 4a smartphones, we also evaluated performance with identical transmitter and receiver models.

6.4.2 Results. As shown in Figure 6, Lee et al.'s scheme and both ggwave variants consistently exhibited very low error rates across all device combinations. DigitalVoices showed varying performances: the best results occurred with the Pixel 6 Pro transmitting to the Nexus 6P (TER $\approx$ 12%), while the Pixel 4a and the Pixel 6 Pro largely failed as receivers. Nearby generally performed well, except when using the Samsung S20 as the transmitter and for some transmissions of the Pixel 4a. PriWhisper was effective in certain combinations, notably with the Pixel 6 Pro as the transmitter (TER as low as 3%), but failed in others due to synchronization issues. Its poor overall performance

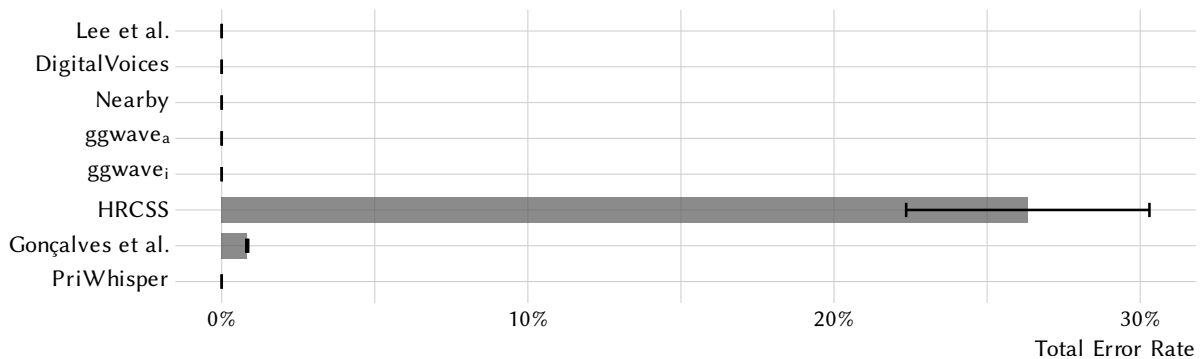
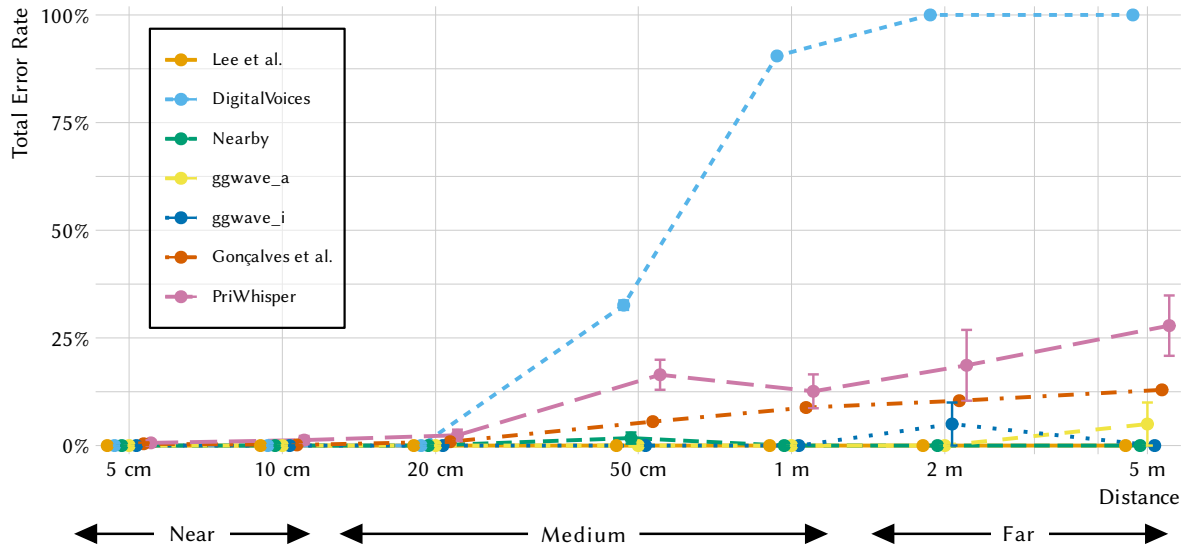
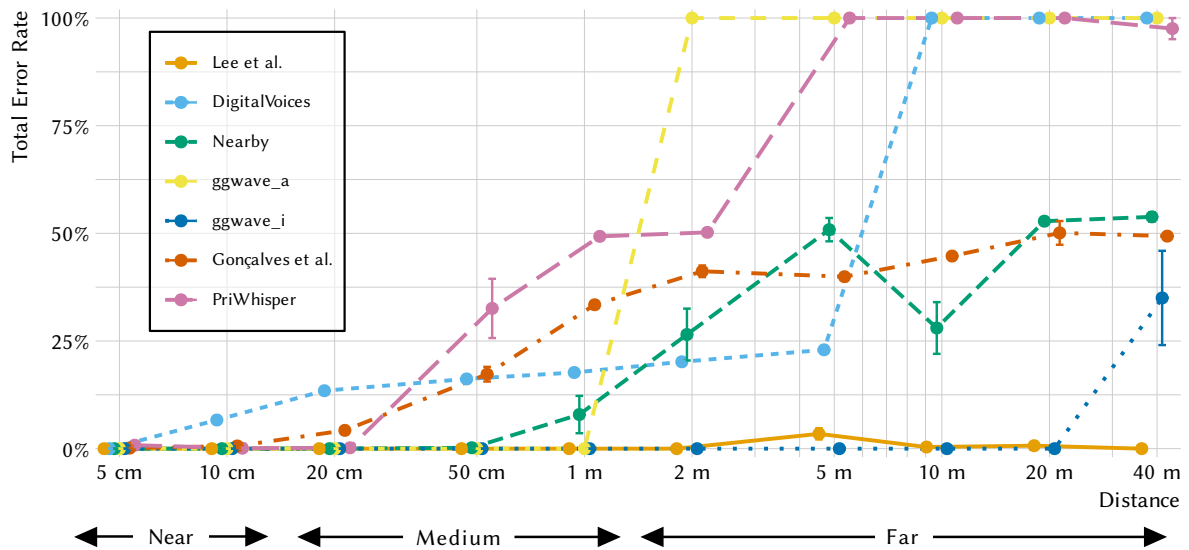


Fig. 4. **Preliminary testing in a best-case hardware scenario.** Shorter bars are better. Schemes are sorted by throughput (see Table 2). The error bars show the standard error. Parameters: $d = 50$ cm; D_{TX} = loudspeaker Neumann KH 80 DSP; D_{RX} = microphone Earthworks M23R; $V_{TX} = -40$ dB interface gain; quiet office environment; $N=240$ measurements.



(a) Anechoic chamber.



(b) Office environment.

Fig. 5. **Reliability of acoustic transmissions with varying distances.** Smaller TER is better. The anechoic chamber allows for measurements up to a maximum distance of 5 m. In the office environment, we were able to measure up to a distance of 40 m, but the measurements at 10 – 40 m were conducted in the hallway. The distance axis is logarithmic, but we only measured at the indicated discrete distances. The error bars show the standard error. The data points of different schemes are dodged to the side to avoid overlap. Parameters: D_{TX} = smartphone Pixel 4a; D_{RX} = smartphone Samsung S20 Ultra; V_{TX} = volume index 19/25; anechoic chamber (Figure 5a) and quiet office environment (Figure 5b); $N=3740$ measurements.

can be attributed to the 50 cm device distance, which highlights its limitations outside its intended usage range of a few cm. In this case, the impact of different device models on the reliability varied significantly, ranging from satisfactory to completely ineffective. The performance of Gonçalves et al.'s scheme was more consistent, with error rates from 5% (best case: Oppo Reno 6 to Pixel 4a) to 32% (worst case: Pixel 4a to Nexus 6P).

Table 3. **Smartphones in the device models experiment** in Section 6.4. The volume column indicates the selected and maximum volume indices for each device. We set each device to a volume of ca. 75% of its maximum volume, which we experimentally identified to be a good trade-off between achieving high volume while maintaining low distortions due to amplifier and speaker non-linearities.

Device Name	Origin	Release Date	Price	Volume
Google Pixel 4a	USA	Oct 2020	349€	19/25
Google Pixel 6 Pro	USA	Oct 2021	900€	19/25
Huawei Nexus 6P	China	Sep 2015	700€	11/15
Oppo Reno 6	China	Sep 2021	499€	12/16
Samsung Galaxy S20 Ultra	South Korea	Feb 2020	1300€	11/15

Overall, the Pixel 6 Pro as the transmitter consistently resulted in the lowest error rates. The Samsung Galaxy S20 Ultra generally performed well as a receiver, except with certain transmitters for DigitalVoices and PriWhisper. This highlights a large diversity in device characteristics, which complicates the generalization of some schemes.

6.5 Impact of different types of noise and device handling

6.5.1 Setup. The preceding tests were conducted in relatively quiet environments, but practical scenarios often contain different types of interference, which can degrade the reliability of acoustic transmissions either due to the ambiance and the surroundings, or due to human handling of the devices. In this experiment, we aimed to assess the influence of different kinds of noises and handling on the receiver. We used the Pixel 4a as the transmitter and the Samsung Galaxy S20 Ultra as the receiver, at a distance of 50 cm aligned on-axis in a quiet indoor environment. We performed experiments for each of the following seven types of noise and handling, designed to mimic real-world disturbances:

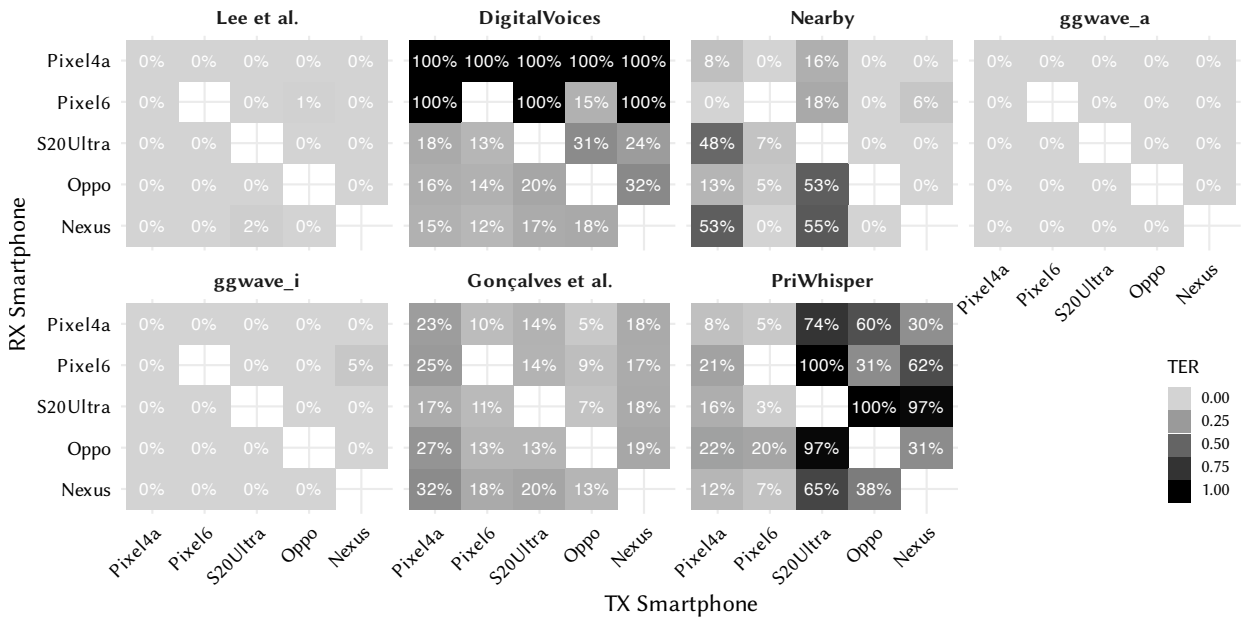


Fig. 6. **Reliability of acoustic transmissions for different device combinations.** Smaller TER is better. Schemes are sorted by throughput. The heatmap shows each scheme's reliability for 21 different TX-RX combinations. Parameters: $d = 50$ cm; $D_{TX}/D_{RX}/V_{TX}$ = see Table 3; office environment; $N=4620$ measurements.

- **Ambient.** Playing ambient field recordings from a café, a train station, and a market place on a Neumann loudspeaker 1 m away from the receiver. The experiment was conducted separately for all three ambient sounds (see [Section A.4.2](#) for detailed setup information).
- **Circle movement.** Shuffling the receiver in a circular motion on the table, which varies the distance by ± 5 cm.
- **Clapping.** Clapping hands near the receiver every 2 s.
- **Pocket.** Inserting the receiver into and removing it from a trouser pocket every 2 s, aligned so that the transmitter's speaker faces the receiver in the pocket.
- **Up and down.** Lifting and firmly placing the receiver on the table approximately every 2 s.
- **Opposite.** The smartphones are positioned with their microphones and speakers facing opposite directions. This arrangement simulates two people holding phones while facing each other.
- **Side-by-side.** The smartphones are placed next to each other, with their long edges adjacent.
- **Handheld.** The transmitter smartphone is held by a person, introducing the natural movement of typical handheld use.

6.5.2 Results. All schemes experienced significant performance degradation when the smartphone was pocketed during transmission ([Figure 7](#)). However, Nearby, the inaudible ggwave variant, and Lee et al.'s scheme were robust against all other noise types tested. The audible ggwave variant was affected by ambient noise, which made it completely unusable, likely because the ambient noise interfered with the transmission bandwidth. Clapping dramatically impaired DigitalVoices, PriWhisper, and Gonçalves et al.'s scheme, rendering them largely ineffective. Noise interactions with the table, whether through circular movement or repetitive lifting and placing, had a relatively minor impact on performance. Notably, DigitalVoices struggled significantly with clapping and pocketing noises ($TER > 50\%$), yet coped better with table-related movements ($TER \approx 6\%$). The side-by-side orientation was more detrimental to reliability than the opposite orientation, although not all schemes were affected.

6.6 Impact of different environments

6.6.1 Setup. To determine the influence of the environment and its multipath profile, we measured transmissions between a Pixel 4a (transmitter) and a Samsung Galaxy S20 Ultra (receiver) aligned on-axis at 50 cm. The evaluation took place in several locations: small and large offices, a meeting room, and a lecture room. In the lecture room, we evaluated both a static setup and a dynamic one in which people walked quietly throughout the room while maintaining an unobstructed line of sight between the devices, dynamically changing the multipath environment (see [Section A.4.3](#)).

6.6.2 Results. As shown in [Figure 8](#), Lee et al.'s scheme, Nearby, and both ggwave variants achieved perfect reliability in all tested environments. The remaining schemes performed worst in the offices, but showed improvement in the meeting and lecture rooms. Interestingly, quiet movements in the vicinity do not seem to negatively impact reliability.

6.7 Summary for each use case

We present a summary of our previous measurement results by scheme and use case in [Figure 9](#). This plot shows how each scheme performs for its intended use case (as per [Table 2](#)) in typical indoor environments, therefore we excluded measurements in the anechoic chamber and with the high-quality studio equipment, as they do not represent typical usage. We discuss these results in more detail in [Section 7.2](#).

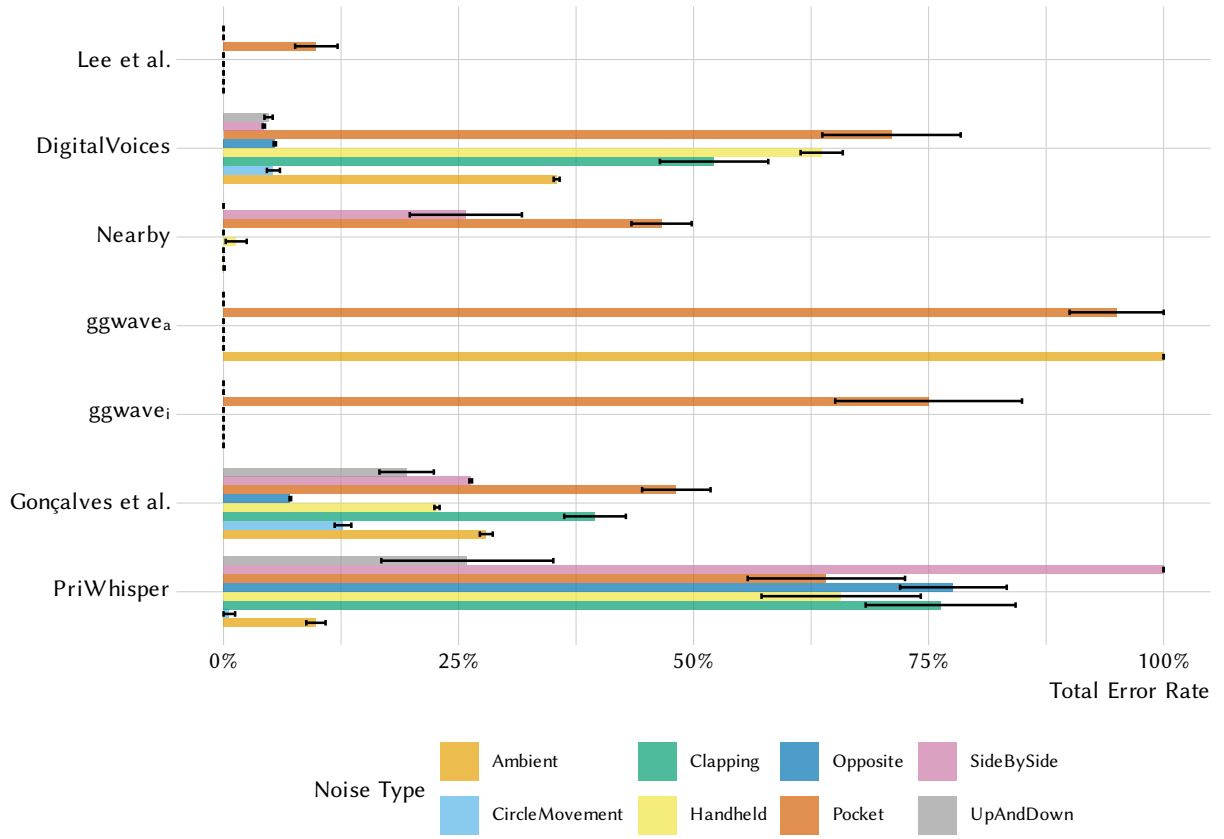


Fig. 7. **Reliability for different types of noise disturbing the transmission.** Smaller bars are better. Schemes are sorted by throughput. For the ambient noise, we performed three separate experiments with different field recordings. The error bars show the standard error. Parameters: $d = 50$ cm; D_{TX} = smartphone Pixel 4a; D_{RX} = smartphone Samsung S20 Ultra; V_{TX} = volume index 19/25; office environment with different types of controlled noise (see Section 6.5); $N=2200$ measurements.

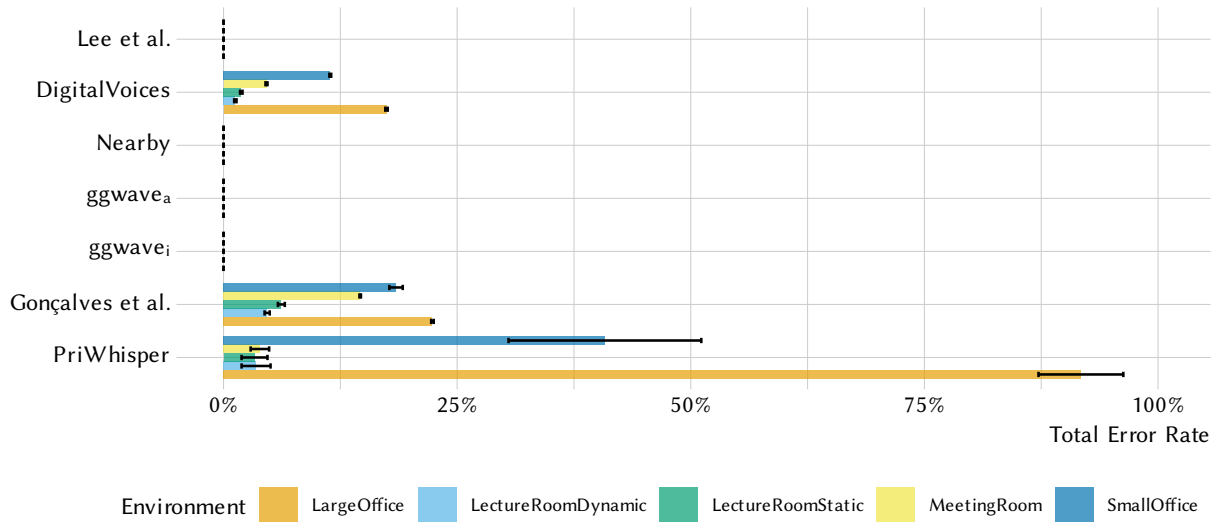


Fig. 8. **Reliability in different transmission environments.** Smaller bars are better. Schemes are sorted by throughput. The error bars show the standard error. Parameters: $d = 50$ cm; D_{TX} = smartphone Pixel 4a; D_{RX} = smartphone Samsung S20 Ultra; V_{TX} = volume index 19/25; $N=1100$ measurements.

6.7.1 Outliers. For the inaudible ggwave variant, one transmission failed during the medium-distance use case ($d = 50$ cm, $D_{TX} = \text{Nexus 6P}$, $D_{RX} = \text{Pixel 6 Pro}$), and some transmissions failed for the far-distance use case ($d = 40$ m, $D_{TX} = \text{Pixel 4a}$, $D_{RX} = \text{Samsung S20 Ultra}$). The outliers for the scheme by Lee et al. were mostly at a distance of 5 m as shown in Figure 5b. Nearby performed well in the medium-distance use case (TER median = 0%), except for a few anomalies in specific device combinations (e.g., $D_{TX} = \text{Samsung S20 Ultra}$, $D_{RX} = \text{Nexus 6P}$).

7 Discussion

7.1 Obtaining software implementations is challenging

To answer our first research question (RQ1) of how researchers can obtain working implementations of acoustic data transmissions systems for evaluation, we conducted a systematic literature study and documented our experience in trying different acquisition methods (Section 5). Out of the 31 publications identified in our survey, none referenced a publicly available implementation, and only four authors provided their implementation upon request. This lack of accessible research artifacts severely hinders replication of research in the field of acoustic data transmission, thereby slowing scientific progress. Without straightforward access to verify and build on existing work, demonstrating improvements over the state of the art becomes significantly more challenging.

We also attempted to re-implement the schemes based on the description in the paper, a process that is both time-consuming and rarely undertaken by other researchers. During this effort, we encountered several inaccuracies and omissions in the original publications, which often required educated guesses to resolve (Appendix B). This challenge is not unique to our field but has been documented in other areas of computer science as well [7]. Such re-implementations carry a

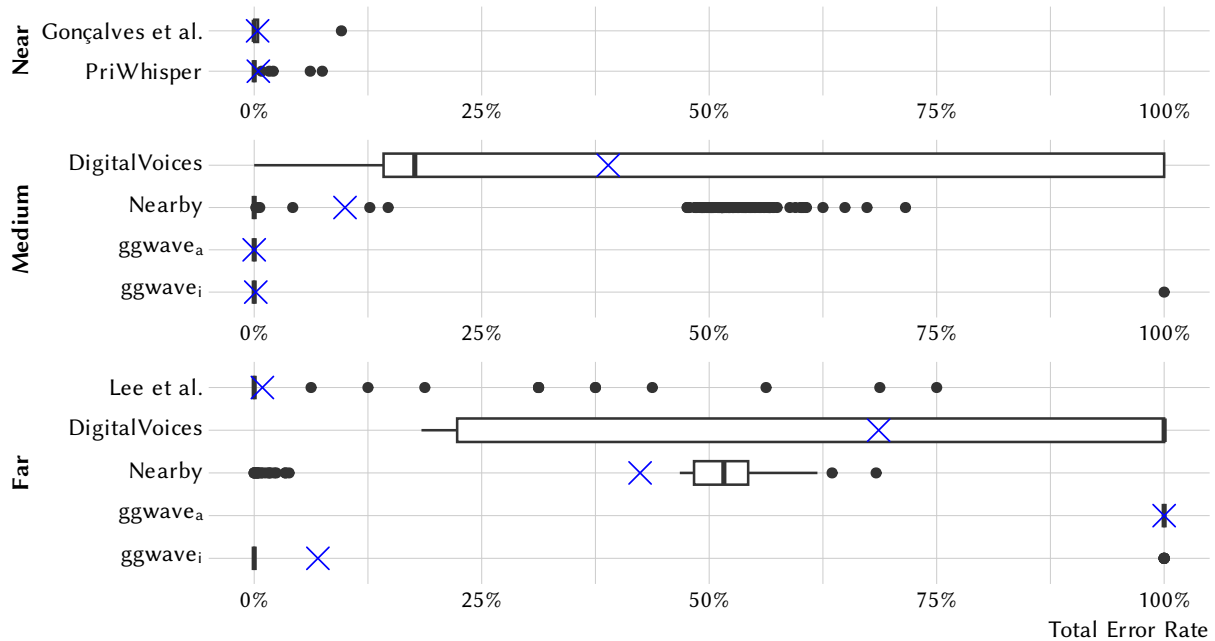


Fig. 9. **Summary of our measurement results from a quiet indoor environment** (Section 6.3, Section 6.4, and Section 6.6), for the three use case categories (defined in Section 2.1 and mapped in Table 2). Smaller TER is better. Schemes are sorted by throughput in each category. The box plots display the median (black bar), mean (blue cross), and outliers (black circles), with hinges marking to the 25th and 75th percentiles. Parameters: d = variable (according to the scenarios in Section 2.1); $D_{TX}/D_{RX}/V_{TX}$ = variable according to Table 3; quiet office environment; $N=3300$ measurements.

Table 4. **Suitability of each scheme for different use case categories (RQ2).** Schemes are sorted by throughput. A scheme is considered suitable (✓) if its mean TER is approximately 0%. Otherwise, it is unsuitable (✗). When only a few outliers appear and the median TER remains at 0%, the scheme may still be suitable but requires additional modifications (e.g., error correction). The highlighted entries indicate novel results that changed compared to prior knowledge (Table 2). The first column indicates how we obtained the implementation: P – referenced in the paper; A – by contacting the authors; R – by re-implementing them ourselves; N – non-academic solution for comparison.

Via	Name	Frequencies	Bandwidth	Inaudible	Throughput ¹	Near	Medium	Far
R	Lee et al. [33]	19.5–22 kHz	2.5 kHz	✓	15 bps	✗	✗	✓
A	Digital Voices [38]	1–3 kHz	2 kHz	✗	80 bps	✗	✗	✗
R	Nearby [13]	18.5–20 kHz	1.5 kHz	✓	84 bps	✗	✓ ²	✗
N	ggwave _a [12]	1.8–6.3 kHz	4.5 kHz	✗	268 bps	✗	✓	✗
N	ggwave _i [12]	15–19.5 kHz	4.5 kHz	✓	268 bps	✗	✓	✓
A	HRCSS [4]	18–22 kHz	4 kHz	✓	500 bps	✗	✗	✗
A	Gonçalves et al. [14]	0.1–4 kHz	3.9 kHz	✗	600 bps	✓ ²	✗	✗
R	PriWhisper [70]	9–17 kHz	8 kHz	✗	729 bps	✓ ²	✗	✗

¹ We list the net data rate, as this is the usable data rate after error correction from the user's perspective.

² Requires further error correction.

high risk of errors or deviations from the original, which is frustrating given the significant time investment. To improve replicability and transparency in research, we encourage other researchers to follow our suggestions in Section 7.5 and publish their software artifacts.

7.2 Generalizability towards ad-hoc smartphone communication use cases

To assess the practicality of the acoustic data transmission schemes proposed in previous studies (RQ2), we evaluated how well these schemes generalize towards the three use case categories defined in Section 2.1. For each use case category, we first reviewed the original publications of each scheme to identify their a priori suitability based on their specified throughput and maximum distance. For example, the scheme by Lee et al. only has a throughput of 15 bps, so by design it is not suitable for transmission of public keys or small files at near distances, as this would take far too long. In many cases, our experiments revealed limited practical suitability of the schemes, which was surprising given the claims in their original publications. Table 4 shows our a posteriori assessment of each scheme based on our evaluation. In particular, all highlighted results in the table represent novel insights which were previously not known based on the limited testing in the current literature. This discrepancy shows the importance of extensive testing across a range of distances and devices to demonstrate practical suitability.

For *far distances* (up to 40 m), the only reliable schemes were the inaudible ggwave variant, effective up to 20 m, and the scheme by Lee et al., which maintained reliability up to 40 m. All other schemes failed at these distances. At *medium distances*, both ggwave variants were effective, with the inaudible one being particularly robust. The Nearby scheme also performed well but struggled to generalize over all tested device combinations. The other schemes are not suited for this use case, either because of high error rates or because their throughput is not high enough. For the *near-distance* use case, only PriWhisper and the scheme by Gonçalves et al. had sufficient throughput, while maintaining low error rates. However, they did experience some erroneous packets, indicating a need for better or additional error correction techniques.

None of our analyzed schemes are suitable for all use cases (see Table 4), as different applications prioritize different trade-offs between reliability, throughput, latency, and audibility. Some schemes offer flexibility; for instance, ggwave can be configured for higher throughput at the cost of reliability or shifted to the near-ultrasonic range to achieve inaudibility.

For applications like secure pairing or mobile payments, where data integrity is paramount, a low TER is the most critical requirement, while high throughput is secondary. The scheme from Lee et al. and PriWhisper are well-suited for this context, although PriWhisper may benefit from additional error correction to guarantee performance in challenging environments. In contrast, use cases such as multimedia applications and content delivery (e.g., second-screen services, location-based information) often prioritize high throughput and inaudibility to ensure a smooth user experience, and can typically tolerate minor, correctable errors. The inaudible ggwave variant is a candidate here, though it cannot reach data rates higher than 268 bps. Notably, schemes designed for higher data rates generally struggled to perform reliably across varied indoor environments.

Given that high throughput generally decreases reliability, a robust design pattern for many practical applications is to treat the acoustic channel as a low-bandwidth trigger: a short, reliable acoustic message (e.g., a unique identifier or hash) can initiate a connection over a high-throughput channel like Wi-Fi or cellular to transfer the bulk of the data.

7.3 Practical challenges for acoustic data transmission

Our measurements show that acoustic data transmission on commodity smartphones and IoT devices faces challenges that rarely appear in simulations or controlled lab settings (RQ3). Because audio hardware and propagation channels differ widely across devices and environments, there is no single transmission specification. Rather than prescribing a new modem architecture, we therefore summarize the *design requirements* that any future scheme should respect for real-world applicability.

7.3.1 Severe multipath propagation. Reverberation produces long delay spreads due to multipath propagation at the relatively slow speed of 343 m/s. Because this delay spread depends on the environment and geometry, a scheme that performs well in one setting may behave differently in another. We observed that even in the same environment, device position and orientation can significantly affect reliability, leading to considerable error rate variability. These findings support the observation by Getreuer et al. that acoustic transmission reliability is sensitive to environmental conditions [13, section 8A]. **Design requirement:** Operate correctly with delay spreads up to tens of milliseconds [33, 59] and the resulting inter-symbol interference (e.g., using guard intervals).

7.3.2 Device heterogeneity. Speaker and microphone responses, amplifier gains, and nonlinearities vary widely: our tests reveal up to 30 dB SPL spread between phones at the same volume setting. Prior research often considered only a specific device model, limiting its practical generalizability. Most acoustic links are unidirectional, so no feedback-based gain control is possible. Developers can only adjust volume relative to each device's maximum output, and our tests show that most devices suffer from nonlinear distortions at volume settings above approximately 75%, though exact thresholds vary. **Design requirement:** Withstand at least 30 dB SNR fluctuation and moderate nonlinear distortion without closed-loop calibration (e.g., using normalization and robust modulation).

7.3.3 Limited bandwidth. IoT audio hardware typically employs ADCs and DACs with sampling rates of 44.1 kHz or 48 kHz, which limits the practical bandwidth to less than 22 kHz. When transmissions must be inaudible, the usable bandwidth reduces to less than 4 kHz, and audio hardware is highly frequency selective in this band [21]. **Design requirement:** When choosing the near-ultrasonic band for inaudibility, prepare for strong frequency-selectivity and limited throughput.

7.3.4 Ambient noise. In contrast to RF communication, the acoustic frequency band is exposed to a wide range of common noise sources—such as traffic, machinery, human speech, and noises from device handling—that occupy a significant share of the available spectrum. Both burst and stationary noise components are common. **Design requirement:** Maintain link integrity in the presence of continuous and impulsive noise (e.g., using strong error correction and time interleaving).

7.3.5 Throughput vs reliability. Some of our surveyed publications in Table 1 pursue very high throughput, but our evaluation shows that increased throughput typically comes at the expense of higher error rates. Commercial systems often target ≤ 200 bps [9, 13, 27, 62], likely because of reliability concerns in practice. **Design requirement:** Favor robustness and aggressive error control unless bi-directional link adaptation is feasible (e.g., HRCSS-style probing), which is often impossible in one-way IoT scenarios.

7.4 Comparison with original evaluations

Several schemes we tested were also evaluated by their original authors. Although our experimental setup differs intentionally to assess generalizability, we compare our results to determine whether the original findings are *conceptually* replicable [61]. Overall, our results align with the original evaluations—considering that deviations are expected due to differences in device characteristics, environments, and our re-implementations. The only exception is HRCSS [4], whose provided implementation yielded impractically high error rates even under optimal conditions. DigitalVoices and gswave lack prior reliability evaluations.

7.4.1 PriWhisper. Zhang et al. [70] conducted a basic reliability evaluation indoors using a Samsung Galaxy S3 and a Google Nexus S, reporting a packet error rate of 0.5%. In our near-distance tests, we observed a similar bit error rate (around 0.5%), though our packet error rate was somewhat higher. Despite these differences, our results are comparable, given that our re-implementation might not be as optimized as the original and that we used different smartphone models.

7.4.2 Gonçalves et al. Gonçalves et al. [14] evaluated their scheme indoors on two portable computers, reporting bit error rates of approximately 3% at 10 cm and 25% at 100 cm. Our experiments revealed a similar trend. Although their noise tests (using background music from a nearby loudspeaker) lack detailed parameters, our noise experiments at 50 cm confirmed a comparable degradation in reliability.

7.4.3 HRCSS. Cai et al.’s HRCSS [4] initially appeared promising, but the provided MATLAB implementation failed our preliminary tests with high-quality equipment. To rigorously justify its exclusion, we analyzed the provided MATLAB source code in detail. We found that the high error rates stemmed from synchronization failures. In particular, the preamble detection algorithm in the implementation deviates from the method described in the HRCSS paper [4]. Specifically, errors in the code cause the receiver to incorrectly calculate and select signal peaks, preventing it from correctly identifying the start of a transmission. Although we did not attempt to fix the code, we verified synchronization as the root cause by manually supplying the correct preamble start index to the decoder, which substantially improved the TER.

We note that HRCSS occasionally produced excellent results, although its performance was highly inconsistent across distances, volumes, and devices. While their paper mentions a Java implementation used in their evaluations, we only received the MATLAB code, so we were unable to test the Java version. This is particularly unfortunate because their study uniquely benchmarks their scheme against an existing one (by Lee et al. [33]) using real devices.

7.4.4 Lee et al. Lee et al. [33] evaluated their scheme on Android devices in indoor far-distance scenarios, achieving packet success rates above 97% at distances of 10 m and 25 m. Our experiments confirmed these high success rates (over 98% at 10 m, 20 m, and 40 m). Additionally, our compatibility tests across 21 device combinations—using smartphones as transmitters at 50 cm—demonstrated excellent robustness and versatility.

7.4.5 Nearby. Getreuer et al. [13] tested Nearby with a Nexus 6 transmitter and Nexus 5/5X receivers, reporting packet success rates over 94% up to 2 m, though error rates increased at larger distances. Our tests confirmed low error rates below 1 m and higher errors at 1 m and 2 m, likely due to differences in our re-implementation, device models, and packet sizes. Another difference is that Getreuer et al. employed a form of error correction by repeating each transmission three times and using the redundancy to correct bit errors, but we did not evaluate this as it effectively reduces throughput to a third. The packet error rate, however, improves when using this form of error correction.

7.5 Towards more rigorous and reproducible evaluation

Our evaluation reveals that many existing acoustic data transmission systems struggle in practical scenarios. This highlights a critical need for more rigorous and reproducible evaluation methodologies to properly assess how new schemes generalize to real-world conditions. The challenges in replicating prior work—due to environmental variations and incomplete documentation—have historically limited comparative analysis, with most studies evaluating new schemes in isolation.

To address this, we present a set of guidelines for more rigorous and reproducible evaluation. The black-box analysis framework we detail in Section 6 serves as a practical implementation of these principles. A key advantage of our methodology is its generality; being system-agnostic, it can be applied to any unidirectional acoustic data transmission system, ensuring its relevance for evaluating future schemes. We encourage future publications to adopt this framework.

7.5.1 Prioritize real-world evaluation over simulation. While simulations are useful for initial design, final evaluations must be conducted on real devices in scenarios representative of the target use case. Our approach in Section 6 provides a template for how to systematically test the impact of key practical challenges:

- **Distance.** Test at various distances relevant to the intended application.
- **Device heterogeneity.** Use a diverse set of transmitter and receiver devices (e.g., different smartphone models or IoT hardware).
- **Environmental noise.** Assess performance against both ambient noise (using field recordings or white noise) and intermittent burst noise (e.g., from device handling).
- **Multipath effects.** Evaluate in environments of varying sizes and acoustic properties.

7.5.2 Focus on comprehensive and relevant metrics. The TER is arguably the most critical metric, as it directly reflects the system’s reliability from the perspective of a developer wanting to integrate acoustic data transmission schemes into their IoT or smartphone projects. Other factors such as audibility and latency do not have to be experimentally determined as they follow directly from the system’s design. In time-critical applications, the additional processing latency during encoding and decoding can also be of importance.

7.5.3 Ensure replicability through open science practices. To enable verification and build upon prior work, we advocate for the following best practices [2, 7, 46, 55, 61]:

- **Publish software artifacts.** Release complete implementations (transmitter and receiver code), ideally along with clear run instructions and analysis scripts. None of the papers we

reviewed provided their code, forcing error-prone and time-consuming re-implementations. We contribute our re-implementations of Nearby [13], PriWhisper [70], and Lee et al.'s scheme [33] in our replication package [51].

- **Describe the experimental setup.** Clearly specify all system parameters and experimental conditions (e.g., TX/RX devices, volume, gain, payload length, distance, environment, background noise) and include images of the setup. Define the evaluation metrics precisely. Our setup is detailed in [Appendix A](#).
- **Publish data artifacts.** Share raw measurement data (such as audio recordings) to enable both replication and reproduction of the evaluation [49]. This transparency allows independent verification and enables other researchers to improve upon the work (e.g., by designing better decoders). We publish a dataset of 11 900 recorded audio transmissions from eight schemes along with a CSV file containing raw BER results for each recording [51].

7.5.4 Establish standardized testbeds. For maximum reproducibility, evaluations could be performed in a standardized testbed, such as an anechoic chamber. To simulate realistic conditions, controlled noise sources playing back ambient field recordings or white noise can be introduced.

7.6 Limitations

7.6.1 Implementation variability. Due to a lack of research transparency in the field of acoustic data transmission and the absence of publicly available implementations for proposed schemes, we had to re-implement some schemes ourselves. This process was laborious and prone to errors. Our re-implementations of Nearby [13], PriWhisper [70], and the scheme by Lee et al. [33] may differ from the originals because the publications did not provide full implementation details. In part, we had to rely on educated guesses, which we describe in [Appendix B](#). Furthermore, for our re-implementations of Nearby and PriWhisper, we used larger message lengths than those reported in the original publications to meet the demands of their intended use cases in a single transmission. We verified our re-implementations using simulations and practical experiments, which showed that Nearby and PriWhisper can handle these larger message lengths while maintaining high reliability, making them conceptually replicable under these more challenging conditions. This positive result indicates that our re-implementations did not degrade the performance of the original implementations. During our experiments, all of our re-implementations were generally robust and reliable for their intended use case.

7.6.2 Environmental variability. We conducted our experiments in quiet indoor rooms and a hallway, acknowledging that varying environmental conditions could have influenced our results. To ensure accuracy, we repeated any experiments that yielded inconsistent results or were affected by unintended noise interference. However, one measurement stood out: Nearby performed significantly worse in the device models experiment compared to the distance experiment, despite using the same devices and transmission distance. This was not a measurement error; the only difference was the change in office rooms. This supports the observation by the Nearby authors, who also observed that the reliability of acoustic data transmissions is sensitive to environmental conditions [13, section 8A]. Therefore, given that multipath environments such as office rooms are inherently variable and cannot be precisely replicated by others (see [Section 7.3.1](#)), we also conducted tests in an anechoic chamber, which is a controlled and reproducible testing environment.

7.6.3 Burst noise variability. The burst noise experiment presented several challenges not present in the ambient noise experiment, due to the need for manual handling of the devices. This not only made the process time-consuming—requiring manual effort for 560 transmissions—but also introduced variability in handling, which affected the replicability of the experiment. We attempted

to mitigate this by triggering the burst noise at 2 s intervals, but some variability remains. Note that this setup resulted in less burst noise exposure per transmission for the schemes from Gonçalves et al. [14] and Lee et al. [33], which had shorter transmission durations compared to the other schemes (Gonçalves et al. at 0.4 s, Lee et al. at 1.1 s, others approximately 5 s; all with a 2 s pause between transmissions). Consequently, the results from these two schemes are not directly comparable with those from the others.

8 Future work

Based on our evaluation, we have identified a list of design requirements in Section 7.3, which will allow future schemes to be ready for real-world deployment. In particular, our comparison across multiple smartphones demonstrated that most schemes struggle to adapt to different device combinations, due to variations in audio hardware such as frequency responses of microphones and speakers, amplifier non-linearities, and potential on-chip post-processing. Conducting a comprehensive study of acoustic characteristics across a variety of devices, including low-cost IoT hardware and smart devices, could delineate the operational limits within which schemes must reliably function. Such knowledge could streamline future testing efforts and assist in selecting relevant test devices.

Additionally, we observed that most prior studies utilized varying experimental setups when testing with real devices, complicating comparisons between results. Our work highlights the necessity for increased research transparency, specifically in detailing how evaluations are conducted. Future efforts should aim to establish and define a standard for testing acoustic communication schemes. Our approach, which involved designing a unified experimental setup for multiple schemes, could serve as a foundation (Appendix A). However, standardizing testing environments remains a challenge, especially outside of controlled settings like anechoic chambers.

9 Conclusion

We conducted a systematic literature study on acoustic data transmission schemes and analyzed their generalizability by requesting implementations from authors, re-implementing selected schemes ourselves, and evaluating them on real devices in both realistic and controlled environments. Our findings reveal that acoustic data transmission schemes often face reliability issues, particularly at higher data rates and across device models. This problem has not been extensively acknowledged or discussed in previous research, which predominantly relied on theoretical simulations or conducted only limited real-device testing. Moving forward, we recommend stronger emphasis on evaluating real-world performance, and we offer our evaluation methodology as a foundation for future work.

RQ1. *“How challenging is it to obtain or re-implement systems proposed in this field?”* The state of replicability in this field is bleak, as not a single publication in our systematic literature study provided accessible implementations. Contacting authors for their implementations proved largely unsuccessful, and re-implementing these systems is both time-consuming and challenging due to the inconsistent information provided in the publications.

RQ2. *“How well do these schemes generalize to ad-hoc smartphone communication use cases?”* Most of the schemes we tested are not yet reliable for practical ad-hoc data transmission between nearby smartphones. Although many claim high throughput, they fail to sustain it in real conditions. However, the scheme by Lee et al. [33] proved reliable even at longer distances, and the non-academic ggwave [12] is effective for close to medium distances.

RQ3. *“Which practical challenges affect nearby acoustic data transmission between smart devices?”* Acoustic data transmission on smart devices faces additional, often overlooked limitations compared to RF communication, including severe inter-symbol interference from long delay spreads and the added complexity of non-Gaussian noise and device diversity. Because simulations cannot

accurately capture these real-world complexities, we systematically tested acoustic schemes across multiple device models under varying environmental and noise conditions to quantify their impact on reliability and generalizability. Our findings show that thorough practical testing clarifies the trade-off between throughput and reliability for specific use cases, guiding the selection of an optimal system for practical applications. To achieve practical applicability, future schemes should be resistant to delay spreads of tens of milliseconds, withstand large volume fluctuations due to device heterogeneity, and handle both continuous and impulsive noise.

Availability

To provide transparency about our research results, we provide a replication package [51] containing:

- A Methodological Transparency & Reproducibility Appendix (META) containing detailed descriptions of our experimental setup, including the exact settings for transmission and reception (see [Appendix A](#)).
- Our three re-implementations of acoustic communication schemes [13, 33, 70] in MATLAB (see [Section 5](#)).
- A collection of 11 900 WAV files with recorded transmissions from eight acoustic communication schemes, captured both in an office environment and an anechoic chamber (see [Section 6](#)).
- A CSV file with our raw data, including the BER for each scheme's recordings.
- The full MATLAB analysis code that decodes recordings from different schemes in parallel, along with R scripts that generate all result figures.

Acknowledgments

We thank all authors who replied to us when we requested their acoustic data transmission implementations following our literature study ([Section 5](#)). Furthermore, we thank the anonymous reviewers for their helpful suggestions. This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

References

- [1] Will Archer Arentz and Udana Bandara. 2011. Near Ultrasonic Directional Data Transfer for Modern Smartphones. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*. Association for Computing Machinery, New York, NY, USA, 481–482. <https://doi.org/10.1145/2030112.2030181>
- [2] Vaibhav Bajpai, Anna Brunstrom, Anja Feldmann, Wolfgang Kellerer, Aiko Pras, Henning Schulzrinne, Georgios Smaragdakis, Matthias Wählisch, and Klaus Wehrle. 2019. The Dagstuhl Beginners Guide to Reproducibility for Experimental Networking Research. *SIGCOMM Comput. Commun. Rev.* 49, 1 (Feb. 2019), 24–30. <https://doi.org/10.1145/3314212.3314217>
- [3] Green Bang, Myoungbeom Chung, and Ilju Ko. 2016. Data Communication Method Based on Inaudible Sound at near Field. In *2016 18th International Conference on Advanced Communication Technology (ICACT)*. 639–642. <https://doi.org/10.1109/ICACT.2016.7423502>
- [4] Chao Cai, Zhe Chen, Jun Luo, Henglin Pu, Menglan Hu, and Rong Zheng. 2022. Boosting Chirp Signal Based Aerial Acoustic Communication Under Dynamic Channel Conditions. *IEEE Transactions on Mobile Computing* 21, 9 (Sept. 2022), 3110–3121. <https://doi.org/10.1109/TMC.2021.3051665>
- [5] Yongliang Chen, Tao Ni, Weitao Xu, and Tao Gu. 2022. SwipePass: Acoustic-based Second-factor User Authentication for Smartphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3 (Sept. 2022), 106:1–106:25. <https://doi.org/10.1145/3550292>
- [6] Ming Ki Chong and Hans W. Gellersen. 2013. How Groups of Users Associate Wireless Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1559–1568. <https://doi.org/10.1145/2470654.2466207>
- [7] Christian Collberg and Todd A. Proebsting. 2016. Repeatability in Computer Systems Research. *Commun. ACM* 59, 3 (Feb. 2016), 62–69. <https://doi.org/10.1145/2812803>

- [8] CopSonic Technologies. 2025. *CopSonic - Ultrasonic Authentication*. <https://www.copsonic.com/> [Retrieved 2025-03-03].
- [9] CUE Audio, Inc. 2025. *Cue - Creating Unforgettable Experiences*. <https://www.connectwithcue.com/> [Retrieved 2025-03-03].
- [10] Anind K. Dey. 2001. Understanding and Using Context. *Personal and Ubiquitous Computing* 5, 1 (Feb. 2001), 4–7. <https://doi.org/10.1007/s007790170019>
- [11] Vadim Gerasimov and Walter Bender. 2000. Things That Talk: Using Sound for Device-to-Device and Device-to-Human Communication. *IBM Systems Journal* 39, 3.4 (2000), 530–546–530–546. <https://doi.org/10.1147/sj.393.0530>
- [12] Georgi Gerganov. 2025. *ggwave*. <https://github.com/ggerganov/ggwave> [Retrieved 2025-03-03].
- [13] Pascal Getreuer, Chet Gnegy, Richard F. Lyon, and Rif A. Saurous. 2018/june. Ultrasonic Communication Using Consumer Hardware. *IEEE Transactions on Multimedia* 20, 6 (2018/june), 1277–1290. <https://doi.org/10.1109/TMM.2017.2766049>
- [14] Monira A. Gonçalves, Rafael V. Louredo, Michel P. Tcheou, and Lisandro Lovisolo. 2017. Acoustic Modem Using OFDM and Least-Square Channel Estimation. In *2017 IEEE 9th Latin-American Conference on Communications (LATINCOM)*. 1–5. <https://doi.org/10.1109/LATINCOM.2017.8240178>
- [15] Michael T. Goodrich, Michael Sirivianos, John Solis, Claudio Soriente, Gene Tsudik, and Ersin Uzun. 2009/february. Using Audio in Secure Device Pairing. *Int. J. Secur. Netw.* 4, 1/2 (2009/february), 57–68–57–68. <https://doi.org/10.1504/IJSN.2009.023426>
- [16] Michael T. Goodrich, Michael Sirivianos, John Solis, Gene Tsudik, and Ersin Uzun. 2006. Loud and Clear: Human-Verifiable Authentication Based on Audio. In *26th IEEE International Conference on Distributed Computing Systems (ICDCS. IEEE)*. <https://doi.org/10.1109/icdcs.2006.52>
- [17] Daniel Halperin, Thomas S. Heydt-Benjamin, Benjamin Ransford, Shane S. Clark, Benessa Defend, Will Morgan, Kevin Fu, Tadayoshi Kohno, and William H. Maisel. 2008. Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses. In *2008 IEEE Symposium on Security and Privacy (Sp 2008)*. 129–142. <https://doi.org/10.1109/SP.2008.31>
- [18] Michael Haus, Aaron Yi Ding, Qing Wang, Juhani Toivonen, Leonardo Tonetto, Sasu Tarkoma, and Jorg Ott. 2019. Enhancing Indoor IoT Communication with Visible Light and Ultrasound. In *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. 1–6. <https://doi.org/10.1109/ICC.2019.8762001>
- [19] Yitao He and Junyu Bian. 2018. Novel Ultrasonic Broadcast Communication System. In *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*. 1–7. <https://doi.org/10.1109/WCSP.2018.8555902>
- [20] Yann Hornych, Javier Cañada Toledo, Boyang Wang, Won-Jae Yi, and Jafar Saniie. 2020. Near-Ultrasonic Communications for IoT Applications Using Android Smartphone. In *2020 IEEE International Conference on Electro Information Technology (EIT)*. 407–410. <https://doi.org/10.1109/EIT48999.2020.9208265>
- [21] Veronika Hromadová, Peter Kasák, Roman Jarina, and Peter Brida. 2022. Frequency Response of Smartphones at the Upper Limit of the Audible Range. In *2022 ELEKTRO (ELEKTRO)*. 1–5. <https://doi.org/10.1109/ELEKTRO53996.2022.9803475>
- [22] Kwang Myung Jeon, Hong Kook Kim, and Myung J. Lee. 2016. Noncoherent Low-Frequency Ultrasonic Communication System with Optimum Symbol Length. *International Journal of Distributed Sensor Networks* 12, 5 (May 2016), 9713180. <https://doi.org/10.1155/2016/9713180>
- [23] Xiaoyu Ji, Chaohao Li, Xinyan Zhou, Juchuan Zhang, Yanmiao Zhang, and Wenyan Xu. 2020. Authenticating Smart Home Devices via Home Limited Channels. *ACM Trans. Internet Things* 1, 4, Article 24 (Aug. 2020), 24 pages. <https://doi.org/10.1145/3399432>
- [24] Meng Jin, Yuan He, Yunhao Liu, and Xinbing Wang. 2022. Furtively Connecting IoT Devices with Acoustic Noise. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. 195–207. <https://doi.org/10.1109/IPSN54338.2022.00023>
- [25] Tero Jokela, Ming Ki Chong, Andrés Lucero, and Hans Gellersen. 2015. Connecting Devices for Collaborative Interactions. *Interactions* 22, 4 (June 2015), 39–43. <https://doi.org/10.1145/2776887>
- [26] Tero Jokela and Andrés Lucero. 2014. FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 369–378. <https://doi.org/10.1145/2628363.2628376>
- [27] Daniel Jones. 2022. *How Near-Ultrasonic Audio Adds Spatial Awareness to the Sonos System*. <https://tech-blog.sonos.com/posts/how-near-ultrasonic-audio-adds-spatial-awareness-to-the-sonos-system/> [Retrieved 2025-03-03].
- [28] Soonwon Ka, Tae H. Kim, Jae Y. Ha, Sun H. Lim, Su C. Shin, Jun W. Choi, Chulyoung Kwak, and Sunghyun Choi. 2016. Near-Ultrasound Communication for TV's 2Nd Screen Services. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking*. ACM, 42–54–42–54. <https://doi.org/10.1145/2973750.2973774>
- [29] Vijay A. Kanade. 2018. Audio-Internet [A-I]: A Novel Technique for Streaming [Bits-over-Sound]: SoundData. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 1690–1694.

- <https://doi.org/10.1109/ICACCI.2018.8554953>
- [30] Soohyun Kim, Hyunsu Mun, and Youngseok Lee. 2019. A Data-Over-Sound Application: Attendance Book. In *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. 1–4. <https://doi.org/10.23919/APNOMS.2019.8892996>
 - [31] Greg Kumparak. 2013. *SlickLogin Aims To Kill The Password By Singing A Silent Song To Your Smartphone*. <https://techcrunch.com/2013/09/09/slicklogin-wants-to-kill-the-password-by-singing-a-silent-song-to-your-smartphone/> [Retrieved 2025-03-03].
 - [32] Chulyoung Kwak, Seongwong Kim, Soonwon Ka, Jihwan Lee, and Sunghyun Choi. 2019/june. No Entry: Anti-Noise Energy Detector for Chirp-Based Acoustic Communication. In *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 1–9. <https://doi.org/10.1109/SAHCN.2019.8824986>
 - [33] Hyewon Lee, Tae H. Kim, Jun W. Choi, and Sunghyun Choi. 2015/april. Chirp Signal-Based Aerial Acoustic Communication for Smart Devices. In *INFOCOM*. 2407–2415. <https://doi.org/10.1109/INFOCOM.2015.7218629>
 - [34] Jihwan Lee, Chulyoung Kwak, Seongwon Kim, and Saewoong Bahk. 2020. Reliable and Low-Complexity Chirp Spread Spectrum-Based Aerial Acoustic Communication. *IEEE Access* 8 (2020), 151589–151601. <https://doi.org/10.1109/ACCESS.2020.3017097>
 - [35] Min-Chun Lin, Fu-Yu Huang, and Tzi-Dar Chiueh. 2015. A-NFC: Two-way near-Field Communications (NFC) via Inaudible Acoustics. In *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*. 1–6. <https://doi.org/10.1109/IISA.2015.7388085>
 - [36] Qiongzhen Lin, Lei Yang, and Yunhao Liu. 2017. TagScreen: Synchronizing Social Televisions through Hidden Sound Markers. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*. 1–9. <https://doi.org/10.1109/INFOCOM.2017.8057170>
 - [37] LISNR, Inc. 2025. *Radius 3.0 Ultrasonic SDK*. <https://lisnr.com/radius-ultrasonic-sdk-3/> [Retrieved 2025-03-03].
 - [38] Cristina V. Lopes and Pedro M. Q. Aguiar. 2001. Aerial Acoustic Communications. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*. 219–222. <https://doi.org/10.1109/ASPAA.2001.969582>
 - [39] Cristina V. Lopes and Pedro M. Q. Aguiar. 2003/july. Acoustic Modems for Ubiquitous Computing. *IEEE Pervasive Computing* 2, 3 (2003/july), 62–71. <https://doi.org/10.1109/MPRV.2003.1228528>
 - [40] Anil Madhavapeddy, David Scott, and Richard Sharp. 2003. Context-Aware Computing with Sound. In *UbiComp 2003: Ubiquitous Computing (Lecture Notes in Computer Science)*, Anind K. Dey, Albrecht Schmidt, and Joseph F. McCarthy (Eds.). Springer, Berlin, Heidelberg, 315–332. https://doi.org/10.1007/978-3-540-39653-6_25
 - [41] Anil Madhavapeddy, Richard Sharp, David Scott, and Alastair Tse. 2005/july. Audio Networking: The Forgotten Wireless Technology. *IEEE Pervasive Computing* 4, 3 (2005/july), 55–60. <https://doi.org/10.1109/MPRV.2005.50>
 - [42] Adib Mehrabi, Antonella Mazzoni, Daniel Jones, and Anthony Steed. 2019. Evaluating the User Experience of Acoustic Data Transmission. *Personal and Ubiquitous Computing* (Dec. 2019), 1–14–1–14. <https://doi.org/10.1007/s00779-019-01345-7>
 - [43] Jonathon Miegel, Philip Branch, and Peter Blamey. 2018. Wireless Communication between Personal Electronic Devices and Hearing Aids Using High Frequency Audio and Ultrasound. *The Journal of the Acoustical Society of America* 144, 4 (Oct. 2018), 2598–2604. <https://doi.org/10.1121/1.5063813>
 - [44] Satoshi Moriyama, Junji Takahashi, and Yoshito Tobe. 2017. Creation of Communication Groups Using Sound of Smartphones. In *Intelligent Environments 2017*. IOS Press, 8–14. <https://doi.org/10.3233/978-1-61499-796-2-8>
 - [45] Rajalakshmi Nandakumar, Krishna K. Chintalapudi, Venkat Padmanabhan, and Ramarathnam Venkatesan. 2013. Dhvani: Secure Peer-to-Peer Acoustic NFC. *ACM SIGCOMM Computer Communication Review* 43, 4 (Aug. 2013), 63–74–63–74. <https://doi.org/10.1145/2534169.2486037>
 - [46] B. A. Nosek, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. 2015. Promoting an Open Research Culture. *Science* 348, 6242 (June 2015), 1422–1425. <https://doi.org/10.1126/science.aab2374>
 - [47] Ed Novak, Zhuofan Tang, and Qun Li. 2019. Ultrasound Proximity Networking on Smart Mobile Devices for IoT Applications. *IEEE Internet of Things Journal* 6, 1 (Feb. 2019), 399–409. <https://doi.org/10.1109/JIOT.2018.2848099>
 - [48] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. 2021. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* 372 (March 2021), n71. <https://doi.org/10.1136/bmj.n71>

- [49] Prasad Patil, Roger D. Peng, and Jeffrey T. Leek. 2016. A Statistical Definition for Reproducibility and Replicability. , 066803 pages. <https://doi.org/10.1101/066803>
- [50] Florentin Putz, Flor Álvarez, and Jiska Classen. 2020. Acoustic Integrity Codes: Secure Device Pairing Using Short-Range Acoustic Communication. In *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 31–41. <https://doi.org/10.1145/3395351.3399420> arXiv:2005.08572
- [51] Florentin Putz, Philipp Fortmann, Jan Frank, Christoph Haugwitz, Mario Kupnik, and Matthias Hollick. 2025. *Replication Package: Evaluating Acoustic Data Transmission Schemes for Ad-Hoc Communication Between Nearby Smart Devices*. <https://doi.org/10.5281/zenodo.17661991> [Dataset].
- [52] Florentin Putz, Steffen Haesler, and Matthias Hollick. 2024. Sounds Good? Fast and Secure Contact Exchange in Groups. In *Proceedings of the ACM on Human-Computer Interaction 8* (New York, NY, USA) (CSCW2). Association for Computing Machinery. <https://doi.org/10.1145/3686964>
- [53] Florentin Putz, Steffen Haesler, Thomas Völkl, Maximilian Gehring, Nils Rollshausen, and Matthias Hollick. 2024. PairSonic: Helping Groups Securely Exchange Contact Information. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '24 Companion)*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3678884.3681818>
- [54] Nafis Sadeq and Md. Shohrab Hossain. 2022. Privacy Breach in Android Smartphone Through Inaudible Sound. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. 2441–2446. <https://doi.org/10.1109/WCNC51071.2022.9771721>
- [55] Kavous Salehzadeh Niksirat, Lahari Goswami, Pooja S. B. Rao, James Tyler, Alessandro Silacci, Sadiq Aliyu, Annika Aebli, Chat Wacharamanotham, and Mauro Cherubini. 2023. Changes in Research Ethics, Openness, and Transparency in Empirical Studies between CHI 2017 and CHI 2022. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–23. <https://doi.org/10.1145/3544548.3580848>
- [56] Shigeru Sasano, Fumihito Sasamori, Osamu Takyu, and Shiro Handa. 2014. Performance Evaluation of Acoustic OFDM with Repetition Coding. In *2014 International Conference on Information and Communication Technology Convergence (ICTC)*. 798–799. <https://doi.org/10.1109/ICTC.2014.6983292>
- [57] Utkalika Satapathy, Soumyajit Chatterjee, and Sandip Chakraborty. 2020. SilentInformer: A Framework for Information Dispersion Using Inaudible Acoustic Signals. In *2020 International Conference on COMMunication Systems & NETWORKS (COMSNETS)*. 543–546. <https://doi.org/10.1109/COMSNETS48256.2020.9027493>
- [58] Markus Scheck, Florentin Putz, Frank Hessel, Hermann Leinweber, Jonatan Crystall, and Matthias Hollick. 2023. DEMO: Secure Bootstrapping of Smart Speakers Using Acoustic Communication. (2023). <https://doi.org/10.26083/tuprints-00024180> Demo presented at the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec 2023).
- [59] Liming Shi, Xinheng Wang, Limin Yu, Wenwu Wang, Zhi Wang, Muddesar Iqbal, Charalampos C. Tsimenidis, and Shahid Mumtaz. 2023. A Long-Range Aerial Acoustic Communication Scheme. *Physical Communication* 60 (Oct. 2023), 102135. <https://doi.org/10.1016/j.phycom.2023.102135>
- [60] Shopkick, Inc. 2025. *Shopkick*. <https://www.shopkick.com/> [Retrieved 2025-03-03].
- [61] Patrick E. Shrout and Joseph L. Rodgers. 2018. Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology* 69, Volume 69, 2018 (Jan. 2018), 487–510. <https://doi.org/10.1146/annurev-psych-122216-011845>
- [62] Stimshop. 2024. *Wi-Us by Stimshop. Wireless via Ultrasound*. <https://wi-us.eu/> [Retrieved 2024-11-06].
- [63] The American Business Awards. 2014. *VeriFone's Way2ride*. <https://stevieawards.com/aba/verifones-way2ride> [Retrieved 2025-03-03].
- [64] TrillBit Inc. 2025. *TrillBit - Contactless, Fast and Secure Authentication Protocol*. <https://www.trillbit.com/technology-trillbit> [Retrieved 2025-03-03].
- [65] Qian Wang, Kui Ren, Man Zhou, Tao Lei, Dimitrios Koutsonikolas, and Lu Su. 2016. Messages behind the Sound. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking - MobiCom '16*. ACM Press. <https://doi.org/10.1145/2973750.2973765>
- [66] Tianben Wang, Daqing Zhang, Yuanqing Zheng, Tao Gu, Xingshe Zhou, and Bernadette Dorizzi. 2018. C-FMCW Based Contactless Respiration Detection Using Acoustic Signal. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4 (Jan. 2018), 170:1–170:20. <https://doi.org/10.1145/3161188>
- [67] Weiguo Wang, Jinming Li, Yuan He, Xiuzhen Guo, and Yunhao Liu. 2022/march. MotorBeat: Acoustic Communication for Home Appliances via Variable Pulse Width Modulation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1 (2022/march), 1–24–1–24. <https://doi.org/10.1145/3517255>
- [68] Yu Wang, Xiaojun Zhu, and Hao Han. 2021. ChirpMu: Chirp Based Imperceptible Information Broadcasting with Music. In *2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS)*. 1–9. <https://doi.org/10.1109/IWQOS52092.2021.9521353>

- [69] Katsuyoshi Yamamoto and Hiroshi Kubo. 2022. 32kbps Terrestrial Acoustic Communications of Single Carrier Block Transmission Employing Transmit Precoding and Channel Interpolation. *IEICE Communications Express* 11, 3 (2022), 141–147. <https://doi.org/10.1587/comex.2021XBL0206>
- [70] Bingsheng Zhang, Qin Zhan, Si Chen, Muyuan Li, Kui Ren, Cong Wang, and Di Ma. 2014. PriWhisper : Enabling Keyless Secure Acoustic Communication for Smartphones. *IEEE Internet of Things Journal* 1, 1 (Feb. 2014), 33–45. <https://doi.org/10.1109/JIOT.2014.2297998>
- [71] Xiao Zhang, Jiqiang Liu, Si Chen, Yongjun Kong, and Kui Ren. 2019. PriWhisper+: An Enhanced Acoustic Short-Range Communication System for Smartphones. *IEEE Internet of Things Journal* 6, 1 (Feb. 2019), 614–627–614–627. <https://doi.org/10.1109/jiot.2018.2850524>
- [72] Yanchao Zhao, Yiming Zhao, Si Li, Hao Han, and Lei Xie. 2023. UltraSnoop: Placement-agnostic Keystroke Snooping via Smartphone-based Ultrasonic Sonar. *ACM Trans. Internet Things* 4, 4, Article 22 (Nov. 2023), 24 pages. <https://doi.org/10.1145/3614440>
- [73] Man Zhou, Qian Wang, Kui Ren, Dimitrios Koutsonikolas, Lu Su, and Yanjiao Chen. 2019. Dolphin: Real-Time Hidden Acoustic Signal Capture with Smartphones. *IEEE Transactions on Mobile Computing* 18, 3 (March 2019), 560–573–560–573. <https://doi.org/10.1109/tmc.2018.2842771>

A Methodological transparency & reproducibility appendix

This section offers detailed information about our experimental setup with the goal of improving the transparency and replicability of our work. We document the exact settings we used for transmission and reception. Table 5 lists all the hardware and software we used, including version numbers.

A.1 General

For every experiment, we repeated the measurements in case of external effects, such as vocal or noise interference, or when there were problems with playback and recording of the transmission.

A.2 Transmission

Following the setup shown in Figure 3, we used the smartphones listed in Table 5. First, we generated transmission WAV files for each scheme, which are part of our dataset. Message lengths were set to the publication’s default or if the implementation supported custom message sizes we adjusted it for an approximate five-second transmission, which we consider a practical upper limit for user wait times. All audio signals were then normalized by dividing the entire signal by the highest absolute sample value and scaling it to –3 dBFS.

We transmitted these files on each smartphone using the open-source Android app *Vinyl*, which allows automating the playback of multiple WAV files using playlists. We compiled a playlist containing the transmission files from all schemes, separated by two-second-long silent WAV files to prevent reverberation from affecting subsequent transmissions. For convenience, we also provide a single WAV file containing repeated transmissions from all schemes and corresponding silent gaps for easy playback. We controlled the output volume using the open-source Android app *Volume Control*, which provides precise control of the Android volume index, as shown in Table 3.

A.3 Reception

On the receiver smartphone, we recorded each transmission using the open-source Android app *Audio Recorder*, which allows detailed control over the microphone source. We selected the MIC source for all smartphones, except for the Google Pixel 4a, which also supported the UNPROCESSED source to bypass Android’s post-processing. Recordings were saved as 24 bit PCM float mono WAV files at a minimum of 44.1 kHz, with software gain at 100% and all extra filtering disabled.

We recorded multiple transmissions consecutively and later segmented them manually in the *REAPER* digital audio workstation. To streamline this process, we used a script⁹ for batch cutting

⁹REAPER script for splitting: <https://stash.reaper.fm/v/25677/SplitX.lua>

and a custom iterator for unique file naming. The segments were rendered as 32 bit PCM mono WAV files at a minimum of 44.1 kHz. Finally, we decoded these files with each scheme's receiver implementation to calculate the TER. If a decoded message was longer than expected, we truncated it; if it was shorter, we counted the missing bits as errors.

A.4 Experimental setup

We conducted tests in several indoor locations and an anechoic chamber. For the experiments in the indoor environment, we mounted all smartphones on small tripods with rubber feet, which isolated them from vibrations caused by their speakers. We positioned the smartphones so their speakers and microphones faced directly toward each other (on-axis), measuring the distance between. The dimensions (length x width x height) of the indoor spaces were: *small office* (5.2 m x 4 m x 3 m), *large office* (13.5 m x 5.3 m x 2.8 m), *meeting room* (7.9 m x 5.3 m x 3 m), *lecture room* (13.5 m x 6.6 m x 3.6 m), and the *hallway* for long-distance tests (40 m x 2 m x 3 m).

The *anechoic chamber* (5.2 m x 7.7 m x 5.8 m) provided a reflection-free environment, with walls lined with mineral wool cones coated in synthetic resin (cone length 100 cm, cone base area 24 cm x 24 cm, reflection coefficient below 0.01 for frequencies above 100 Hz).

A.4.1 Preliminary best-case setup. For the preliminary tests (Section 6.2), we established a best-case baseline using studio equipment instead of smartphones. For transmission, a Neumann KH 80 DSP studio speaker played the signals (local control off, acoustic control “free-standing”, output level 94 dB SPL, input gain at 0 dB). We connected the speaker to an RME ADI-2 Pro FS R BE audio interface (−40 dB gain at a reference level of 24 dBu).

For recording, a calibrated Earthworks M23R reference microphone captured the audio. The signal was fed through a Camden EC1 preamp (35.5 dB gain, no additional filtering) into the audio interface (without additional gain). The preamp's output was split to both interface inputs and summed to mono using mid-side processing, increasing the SNR by 3 dB. We used REAPER for playback and recording. The speaker and microphone were aligned on-axis at a distance of 50 cm.

A.4.2 Ambient noise setup. For the ambient noise experiment (Section 6.5), we used the standard smartphone and introduced noise using the Neumann KH 80 DSP studio speaker. The speaker was placed 1 m from the receiver, orthogonal to the transmission axis. The speaker configuration was the same as in the preliminary tests (Section A.4.1), with the exception of the output gain of our interface, which was set to −35 dB to subjectively match the real-life noise levels of the scenarios at the receiver's location. During each iteration of the experiment, one field recording was played repeatedly while the transmission sounds were played sequentially. This process was repeated for all three recordings (café, train station, marketplace), which are freely available online¹⁰ under a CC-BY license.

A.4.3 Dynamic environment setup. To evaluate performance in a time-varying channel (Section 6.6), we conducted tests in a lecture room with four people walking slowly and quietly (without talking) around the devices. The smartphones were mounted on small tripods placed on tables in the center of the room, 50 cm apart on-axis. For comparison, a static control test was performed with an identical furniture layout but without people.

¹⁰Café field recording (CC BY 3.0): <https://freesound.org/people/AshFox/sounds/172968/>, train station field recording (CC BY 4.0): <https://freesound.org/people/Kyster/sounds/121576/>, marketplace field recording (CC BY 3.0): https://freesound.org/people/le_abbaye_Noirlac/sounds/129677/

Table 5. **Hardware and software used during our experiments.**

Type	Manufacturer	Product	Version
Microphone	Earthworks	M23R	
Preamp	Camden	EC1	
Audio Interface	RME	ADI-2 Pro FS R BE	
Calibration Device	Galaxy Audio	CM-C200	
Loudspeaker	Neumann	KH80 DSP	
Smartphone	Google	Pixel 4a	128GB
Smartphone	Google	Pixel 6 Pro	5G Sub-6 128GB
Smartphone	Huawei	Nexus 6P	32GB
Smartphone	Oppo	Reno 6	Dual-SIM 5G 128GB
Smartphone	Samsung	Galaxy S20 Ultra	5G 128GB
Desk Microphone Stand	K&M	23150-3	
Ball Joint	Roadworx	Universal Ball Joint	
Smartphone Clamp	Roadworx	Smartphone Clamp	
Smartphone App	Adrien Poupa	Vinyl Music Player ¹	v1.5.0
Smartphone App	Stas Shakirov	Volume Control ²	v2.6.0
Smartphone App	axet	Audio Recorder ³	v3.5.15
DAW	Cockos	REAPER ⁴	v6.82

¹ Vinyl: <https://f-droid.org/en/packages/com.poupa.vinylmusicplayer/>² Volume Control: <https://f-droid.org/en/packages/com.punksta.apps.volumecontrol/>³ Audio Recorder: <https://f-droid.org/en/packages/com.github.axet.audiorecorder/>⁴ REAPER: <https://www.reaper.fm/>

A.5 Replicating our experiments

The procedures described above should allow for the replication of our experiments. However, the reliability of acoustic transmissions is sensitive to factors like specific device models, microphone/speaker quality, and environmental acoustics. While we expect the overall trends to be similar, absolute TER values will likely differ.

B Re-implementations

In this section, we detail the challenges we encountered during the re-implementation of Lee et al.'s scheme, Nearby, and PriWhisper.

B.1 Lee et al.

We were able to implement all aspects of this scheme as described in their paper [33]. Their practice of listing most system parameters clearly in a table was particularly helpful. We interpret the 16 bps data rate mentioned by Lee et al. as the gross data rate, as it appears they did not account for the preamble in their calculations. Based on the default system parameters provided, we calculate a net data rate of 14.55 bps.

B.2 Nearby

Implementing the transmitter for the Nearby system [13] was straightforward, following the information provided in the paper. However, re-implementing the receiver proved challenging, as some aspects of the paper were unclear to us.

B.2.1 Synchronization. To locate the transmission's start within the recording, the system processes the signal by iterating over samples, segmenting the signal from each sample offset, and correlating these segments with a period of the code signal. This correlation indicates the likelihood that a symbol begins at each sample offset, as each MFSK symbol was spread by the code signal. After correlating across all time offsets, the results are normalized to compensate for any potential Doppler shifts. These results are termed "normalized acquisition scores" in their paper (Section VI.C in [13]). Next, we identify the spacer symbol's position using the normalized acquisition scores. Ideally, a distinct peak at the spacer symbol's position would appear when plotting these scores. However, since the same code signal spreads all token symbols, peaks appear for each symbol, not just the spacer. We implemented a method to identify the positions of the n highest peaks—where n represents the number of token symbols—and select the earliest peak as the start. The paper does not discuss this issue or how to resolve it.

B.2.2 Block processing. The authors mention that their implementation processes signals in 100 ms blocks to likely improve time efficiency, yet the paper does not detail how this feature is implemented. We chose not to implement this aspect, as it is unlikely to impact our evaluation results.

B.2.3 Ambiguity in Algorithm 2. Algorithm 2 in the paper [13] employs a counter variable n in line 2 that initializes at zero and increments by one in each iteration to compute a "raw acquisition score" for each sample offset n , representing the likelihood of a data symbol starting at that offset. However, the authors do not specify an upper bound for n . We have documented our method for determining this upper bound in the comments of our receiver function, although it remains unclear whether this method aligns with the authors' intended approach.

B.2.4 Ambiguity in Equation 47. In Equation 47 in the paper [13], the variable c , representing the sinc-interpolated code sequence, is indexed at position $(kM_p + m)/F_b$. Here, F_b denotes a sample rate of 12 kHz, which means the division results in a time in seconds at which the code signal should be indexed. Since c is a discrete signal sampled at 48 kHz, it cannot be indexed at arbitrary times. To convert the time $(kM_p + m)/F_b$ to a corresponding sample offset in c , we multiply by 48 000, the sample rate. Again, it is not clear if the authors intended this.

B.2.5 Ambiguity in Algorithm 3. Algorithm 3 in the paper [13] uses the matrix of raw acquisition scores a^{raw} , generated by Algorithm 2. In line 8 of Algorithm 3, it accesses row $n - d$ of a^{raw} , where n is a counter variable starting at zero and increasing by one each iteration. The paper does not specify an upper bound for n , but we assume that n runs up to the number of rows in a^{raw} reduced by one. The variable d is set to 215, leading to the algorithm attempting to access negative row indices during the first 215 iterations, which is invalid. We resolved this issue by omitting line 8 in iterations where $n - d$ results in a negative index. However, as a consequence, the normalized acquisition score matrix ends up with 215 fewer rows than a^{raw} . It is unclear whether this reduction is in line with the authors' intentions.

B.2.6 Ambiguity during noise suppression. The bandpass filter for noise suppression uses a variable η , which is supposed to be "a small positive parameter" (Section VI.A in [13]). Since the paper does not specify an exact value for η , we set it to 1.

B.2.7 Data rate. We interpret the stated data rate of 94.5 bps as the gross data rate. The system can operate without the token structure that includes additional spacers and a parity symbol, in which case 94.5 bps is also the net data rate. However, when using the full token structure as described in the paper and used in our evaluation, the net data rate decreases to 84 bps.

B.3 PriWhisper

When re-implementing PriWhisper [70], the major difficulty was designing a mechanism for synchronization and calibration, as the paper does not clearly explain these aspects. Additionally, the descriptions of the interleaving and checksum processes in the paper were ambiguous. The specified data rate was also unclear, further complicating our implementation efforts.

B.3.1 Synchronization. In the original implementation, a sinusoid preamble is used within the jamming signal to aid synchronization. However, for our study, we did not implement the security features of the system, focusing solely on data transmission. To minimize the impact on the data rate, we combine the synchronization and calibration sequence. Accordingly, we introduced a special symbol at the beginning of each transmission, which contains the sum of all available data waveforms $s(t)$ (Section III.A in [70]). For synchronization, we use this known synchronization symbol and cross-correlate it with the recording. The position of the highest peak in the output is considered the start of the transmission.

B.3.2 Calibration. To determine the frequency of an MFSK symbol, we correlate each symbol duration in the recording with all available frequencies using the following equation from Section III.A in [70]:

$$R_m = \left| \int_0^T r(t) e^{i2\pi(f_c + m\Delta f)t} dt \right|, m \in [0, M-1] \quad (1)$$

After iterating over all frequencies, we select the frequency with the highest correlation value and map it to the binary number it represents. We repeat these steps for all symbols to decode the entire data transmission in the recording. Before demodulation, calibration using the synchronization symbol is crucial due to potential variations in microphone sensitivity across frequencies. Without calibration, Equation 1 might return a lower correlation value for the actual frequency present in a symbol due to the frequency-selectivity of the microphone.

To address this, we apply Equation 1 to the synchronization symbol, containing all frequencies. This establishes a baseline correlation for each frequency. We then calculate a calibration factor for each frequency by dividing the highest of these baseline values by the others. This factor adjusts the correlation values during demodulation, ensuring that each frequency is appropriately represented despite the microphone's bias. However, this calibration step assumes that the channel characteristics remain constant throughout the transmission duration.

B.3.3 Interleaving and checksum. The paper mentions interleaving the data and adding a checksum, but it lacks clear guidance on how to implement the interleaving process and the sequence of these steps. In our implementation, we first perform regular random interleaving and then add the checksum.

B.3.4 Data rate. We calculate a gross data rate of 771 bps for $M = 8$, which differs from the 1027 bps reported by the authors. Our data rate calculation is as follows: Using M different frequencies allows encoding $\log_2(M)$ bits per symbol. For $M = 8$, this results in 3 bits per symbol. With a symbol duration of 2 ms, we transmit 500 symbols per second, equivalent to 1500 bits per second. As one coded data block contains 255 BCH encoded bits, we can transmit $\frac{1500}{255} \approx 5.88$ blocks per second. One block contains 131 bits of pure data, resulting in a gross data rate of $5.88 \times 131 \approx 771$ bps, excluding signal components for synchronization or calibration. The resulting net data rate is 729 bps.

Erklärung zur Dissertationsschrift

*gemäß der
Allgemeinen Bestimmungen der Promotionsordnung der Technischen Universität Darmstadt
sowie der zugehörigen
Besonderen Bestimmungen des Fachbereichs Informatik
in den aktuell gültigen Versionen.*

§ 8 Abs. 1 lit. b/c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§ 8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde.

§ 9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfasst wurde. Die „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Technischen Universität Darmstadt“ und die „Leitlinien zum Umgang mit digitalen Forschungsdaten an der TU Darmstadt“ in den jeweils aktuellen Versionen wurden beachtet.

§ 9 Abs. 2 PromO

Ich versichere hiermit, dass die Arbeit bisher noch nicht zu Prüfungszwecken gedient hat.

§ 9 Abs. 5 PromO sowie Bes. Bestimmungen des Fachbereichs Informatik

Eigenzitate aus vorausgehenden wissenschaftlichen Veröffentlichungen sowie die Urheberschaften und Eigenleistungen der einzelnen Beiträge sind im Kapitel „*Collaborations and My Contribution*“ auf den Seiten x bis xi explizit erläutert.

Darmstadt, 28. Mai 2026

Florentin Putz