8
Conclusion

Contents

8.1. Summary and Key Results

In this thesis, we provided a broad analysis of different types of information that is exposed within applications and on the Internet when users interact with digital communication applications. As we have seen, users do not only share data intentionally but also the use of specific applications can reveal information about them to others. For information that is intentionally made available to others, we mainly focused on exposure reduction features, eventually resulting in data lifetime ending. In the context of information that users unintentionally reveal, we demonstrated two types of applications. Tor represents technology that is explicitly used for privacy purposes such as concealing one's identity and messenger apps represent ever-present tools that are widely adopted and used by billions of users for everyday communication purposes.

8.1.1. Managing Self-Published Online Data

In our systematic review of longitudinal online data management in Chapter 2, we categorized a broad range of technical approaches for managing online data longitudinally and studies analyzing how users interact with such features in existing applications and environments. By contrasting technical and user side, we identified incorrect, incomplete, and missing realizations of users' desires in academic proposals for technical solutions. Based on such conflicts, we then derived a set of technical key challenges evolving around the need for flexibilization of data lifetime ending and its conditions, and to better incorporate user perception of security and trust, and their mental models associated with it. The challenges we identified can serve as recommendations for the development of new mechanisms for users to manage their information exposure online.

In a mobile messaging context, we then explored users' perception of and preferences for message deletion options in messengers in Chapter 3. We particularly focused on whether users were able to assess if messages

were only deleted on their own device or also from devices of recipients. We initiated our study following the roll-out of a new feature in WhatsApp, in which users could explicitly select between these two options. Our initial assumption that the effects of message deletion were ambiguous without a clear choice was confirmed by our preparatory analysis of deletion functions in 17 messengers. In our study, we found that users could better determine where a message was deleted when the effects of deletion were explained, as it was implemented in the newly introduced feature in WhatsApp. Our results show that subtleties such as the integration of a simple dialogue have the potential to facilitate user understanding of app functionality and, thus, improve user experience when handling the data they made available to others.

In order to widen the views and also incorporate other non-technical perspectives to allow for more flexibility in the specification of exposure control mechanisms, we proposed a solution combining technical and legal aspects in Chapter 4. Our approach enables users and online platforms to agree on conditions for reducing exposure of online data up to entirely removing it from public access. In order to incentivize the providers of online platforms to comply with the agreement, we used a penalty mechanism that can be triggered by the users and is verified by a neutral authority. Our prototype implementation involving smart contracts based on the Ethereum cryptocurrency system shows that our concept is technically feasible. It also demonstrates how completely new approaches to controlling one's online exposure, designed from scratch, could work.

8.1.2. Usage-Driven Information Revelation

In Chapter 6, we analyzed the feasibility of traffic analysis attacks on Tor, i. e., revealing users' identities and the services they use, under real-world conditions. Compared to theoretically perfect attack performance, real-world adversaries do only have access to a limited set of Tor relays

and can only deanonymize Tor users when they use relays under adversarial control. In this context, we developed three novel stepping-stone attacks that have the potential to reduce the efforts for adversaries as well as improving their chances to uncover Tor users' identities. We have shown how adversaries can use a timing side channel in the circuit establishment handshake to predict the exit relay of the connection to determine in advance whether traffic analysis can be successful. Additionally, adversaries can actively interfere with the circuit establishment to improve their chances for successful user deanonymization. Since all attacks exploit core defensive mechanisms of Tor's circuit establishment, there is no simple way to mitigate them. One possible countermeasure includes obfuscating timings, i. e., adding artificial delays within the circuit establishment procedure, which reduces the attack performance but comes at the cost of usability.

In Chapter 7, we have shown how the whereabouts of individuals can be leaked by exploiting a surprising timing side-channel in widely adopted everyday communication applications. Messenger users can spy on their contacts by simply sending them instant messages and observing the time it takes for message delivery to be confirmed. Timing distributions differ between locations, most likely due to characteristics of the respective Internet connection. Thus, after an initial training phase, an adversary can send a target user a sequence of five messages and determine their whereabouts with up to 95 % accuracy in the scenarios we evaluated. While accuracy varies between scenarios and the three messengers we tested, our results imply that the side-channel exists independent of the underlying messenger infrastructure.

8.2. Directions for Future Research

In the following, we point towards a set of potential topics that we identified while carrying out the work presented in this thesis and that we consider interesting to be addressed by future research.

8.2.1. Filling Gaps Between Technical and User Research

The basis to determine subsequent research tasks is provided by our systematization in Chapter 2. The challenges we identified comprise gaps between technical and user-centered research and each directly points towards open issues that can only be appropriately resolved by taking both sides into account.

For example, incomplete realizations of expiration or exposure reduction mechanisms provide strong indications that users' intentions have not been appropriately addressed. Whereas it is technically sound to develop mechanisms that let data entirely disappear, such mechanisms neglect users more fine-grained preferences, e. g., to make data unavailable for a general audience on the one hand, while at the same time keeping it available for their core peer group.

Therefore, studying users attitudes towards protecting their online data is inevitable in the process of developing new technology. Only when users' intentions are entirely clear, research can equip them with useful tools they need for controlling their information exposure.

8.2.2. Focus on Improving Existing Applications

Developing sound and provably secure concepts and protocols for features such as data deletion represents important foundational work. However, the path to bringing such new mechanisms into effect is not taken by implementing a ground-breaking new tool providing the respective technology and waiting for users to adopt it. Instead, providing better systems most likely entails extending existing applications with additional features that satisfy users needs in interacting with these applications and – ideally – come into effect by default. A related textbook example for this method was showcased by WhatsApp, adapting the Signal protocol for end-to-end encrypted message exchange and turning it on by

default. This procedure effectively enabled end-to-end encrypted communication for more than one billion users world-wide within a moment, most likely without users even noticing.

Following this example, research should explore how newly developed mechanisms for information exposure control are compatible with and can be integrated into popular and widely-adopted applications. This way, many users can immediately profit from latest progress, ideally without being forced to change their behavior, and likewise, new developments can easier find adoption among relevant audiences.

8.2.3. Practicality of Data Revocation Contracts

Our proposal to use agreements based on smart contracts as a means for online data revocation widened the perspectives onto the topic of information exposure control by incorporating legal aspects. While we have presented the fundamental concepts for interaction between users and providers at different stages in the data lifecycle, and demonstrated its technical feasibility with a prototype implementation, the user perspective remains yet unclear.

Since we have sketched how contracts allow for more flexible exposure control mechanisms, we assume that they have the potential to point into the right direction and can better fulfill users' desires in handling their online data. However, for such a mechanism to be eventually deployed in a practical environment, additional research studying users' willingness to adopt it is necessary.

Moreover, the use of currently available cryptocurrency systems blockchains entails additional questions regarding the excessive energy consumption of computationally expensive proof-of-work blockchains. Thus, introducing data revocation smart contracts requires a lot of additional research on the sustainability of the underlying technology and its societal impact, or otherwise comprises a severe burden for currently unresolved environmental challenges.

8.2.4. Trade-offs Between Privacy and Usability

Our technical analyses in the second part of this thesis have showcased two examples for information exposure in digital communication environments unintended by users. For the case of Tor traffic analysis, we have demonstrated how randomized timing delays can help to reduce the attacker success in uncovering users' identities, eventually rendering the attack useless. However, artificially delaying Tor's service comes at the cost of degrading user experience which has also been acknowledged before [67].

Future research could study to what degree users' are willing to tolerate usability cutbacks in exchange for additional privacy or better control of their information exposure. Specifically for the case of privacy-focused applications such as Tor, we discussed ideas evolving around letting users choose between usability and privacy but leave it an open task for future work to explore how such mechanisms could be realized and to study if and how users would be willing to interact with it.

While we have provided evidence that timing delays can indeed mitigate the unintended identity leaks in Tor, such countermeasures might also be effective for the location exposure in messengers which needs to be investigated in practice. Yet again, this entails the need to study potential usability issues, since setting up countermeasures with effects on user experience should not take place without capturing users attitudes towards it before.

Similar to the challenges regarding active data sharing in the first part, we again emphasize the need for joint research from multiple perspectives to address and resolve practical issues as a whole.

8.2.5. Practicality of Traffic Analysis in Messengers

For the location revelation in messenger apps, we have focused on demonstrating its feasibility from a purely technical perspective. Whereas we have shown high accuracy for the location prediction in our experimental

setup, it must be further explored how our results translate into situations in the real world.

Particularly the phase in which the adversary learns the timing patterns of different locations of a contact is presumably trickier to realize than in our experimental setting. Whereas it is not uncommon for contacts to regularly exchange messages, repeatedly sending the same message sequence at constant time intervals will most likely be considered suspicious by a potential target, before any meaningful data could have been collected.

Therefore, studying users might be helpful to determine to what degree sending messages must be throttled for the attack to remain undetected. Additionally, future research could examine if the respective timings could also be collected alongside regular conversations. Actual user behavior is, however, harder to simulate in the lab but reflecting it is a necessary step to fully assess the actual threat under real-world conditions.

8.3. Closing Remarks

In a broader sense, the work presented in this thesis has demonstrated that isolated views on practical topics involving users and applications they are actually using only from a technical security perspective is not enough to build better systems or to improve existing ones. We again emphasize that research studying digital applications that users actually use should always incorporate perspectives from multiple disciplines in order to allow for designing systems and mechanisms that are a benefit for users. This is inevitable to allow for providing technology users actually need and that can help them to better control their own information exposure in digital communication environments.