My research interest is in networked systems and surrounding areas, such as distributed systems, computer networks, and security/privacy. I have worked on various projects including smartphone security and privacy [2–4, 11, 13], network architecture to protect user privacy [5, 15], reliable distributed systems [9], and mobile systems for voice and vision interactions [7, 8].

In my dissertation, I have focused on security and privacy enhancing solutions. Today, the amount of data users create and share via various devices and platforms is ever increasing with emerging technologies. As it is monetized, this personal data has become more valuable and companies have a greater incentive to obtain it. The exposure of user data to other parties is often out of the user's control, and it is irreversible once it occurs, thus compromising a user's security and privacy. To address these issues, my research takes the following approaches:

Identifying and understanding security and privacy issues in existing systems. Despite the increasing security and privacy risk and much attention from research community, we are often not fully aware of security and privacy issues or their implications. Some issues stem from a lack of security considerations in designing systems, how users or developers use the system, or both. As an example of the first, I studied how users' activities can be correlated in the current Internet architecture [5]. As an example of the second, I designed an information flow tracking system for Android [3], studied the third-party libraries in Android applications using the system [4], and characterized how developers use personal information in smartphones.

Designing and building new systems to address these issues. Understanding security and privacy issues motivates solutions to address these challenges. My approach is to build real prototypes and use them to evaluate design alternatives. Building working systems allows for quantitative analyses to see the system's efficiency in practice, which is an important measure as inefficient solutions are not adopted by users even if they are effective. I then make the systems available to the public as open source software. Often, these systems come with new mechanisms that have theoretical underpinnings. For example, I proposed a client-based Paxos consensus algorithm in [9] and an efficient routing mechanism with cryptographic functions in [5].

In what follows, I describe the projects related to my thesis work, and peripherally related but synergistic work, and then future directions for my research.

Thesis Work

My Ph.D. research addresses security and privacy issues in untrusted environments. While it is impossible to have a single holistic solution given diverse platforms, the thesis tries to expand our understanding of how information is exposed to other parties, and to provide solutions that give users more control over the exposure.

Privacy in smartphones. Smartphones have become a central device in people's daily lives, and they hold a lot of personal information including photos, contacts, location, and so on. This information can be accessed by many applications built by third-party developers to provide users with various functionality. Since applications are not necessarily trusted, it is important to understand how applications use and expose user information.

Information flow tracking is a tool that provides a way to understand what private information is leaked when it goes over the boundary of a system. I contributed to the design and implementation of TaintDroid [3], which enables finer-grained tracking when data is communicated between applications by introducing a new data structure and a mechanism to store taint information during IPC.

While it is useful to know where information goes, TaintDroid itself does not provide any protection mechanisms. My coauthors and I further extended it to create a privacy protection system for Android, called AppFence [11]. AppFence protects private data from being exfiltrated by substituting shadow data in place of the data that users wants to keep private and by blocking network transmissions of sensitive data to untrusted destinations.

As mobile applications become more complicated, many applications are built on third-party libraries, especially for advertising and tracking. Unlike the Web, where third-party libraries can be separated and isolated, all the components
in a mobile application are considered as a single entity. I built a tool to monitor communications containing personal information and conducted month-long user studies with smartphone users [4, 13], which revealed prevalent third-party tracking in mobile applications. Based on the insights gathered from the study, we built a new privacy control mechanism that selectively provides shadow access to sensitive data for third-party components.

Internet architecture for protecting privacy. Personal information accessible by remote parties is not limited to obviously private data, such as what we saw from smartphones, but includes how people use applications or browse web pages. Service providers try to correlate as much activity as possible, for example, tracking user behavior on the web or applications is now a common practice. While there are protection mechanisms against tracking in the application layer, network-level fingerprints (as in the IP address of a client) still disclose a substantial amount of information about the user’s activity. Moreover, users have no control over network-level fingerprints and what information can be inferred from them.

My goal was to design systems that provide users with (i) visibility into which of their behavior can be correlated by remote services and (ii) control over this information exposure. One approach to this issue is to redesign the Internet architecture from scratch; my coauthors and I therefore posed the question, how would we design the Internet if we want to have privacy and security as one of the primary concerns? In "Tor instead of IP" [15, 16], we developed an Internet architecture where network elements use the minimal information to allow communication. Inspired by Tor, we showed that our seemingly radical design—removing IP addresses and using onion circuits, combined with other ideas—can be effective in resolving many of the existing security and privacy issues.

Although such a clean-slate design is appealing, changing the whole Internet is hardly feasible in the short term. In [5], we looked into smaller changes that each network AS can independently adopt. The goal of the project was to address information exposure both in the network-layer and in the application-layer with an efficient design that does not impact network performance. To do so, we proposed a cross-layer architecture that provides users with multiple pseudonyms that are unlinkable to each other. To the user, a pseudonym represents a set of activities that the user is comfortable with being linked to each other, and to the outside world, a pseudonym gives the illusion of a single machine.

For the network layer, we take advantage of the ample address space in IPv6, where each host can get a large number of addresses that are spread across an enterprise’s or AS’s IP range. These addresses could be randomly allocated inside the IP range, but this would make the routing protocol inefficient as it would have to maintain separate routing entries for each address. On the other hand, if we were to assign addresses with some structure, the resulting allocation would be ineffective as it could be correlated by an external observer. We dealt with this challenge by observing that part of the IP address (following the prefix) can be transformed without affecting routing. In our system, the network (or a proxy) uses encryption to encode the host identifier in a way that is not discernible to outside adversaries. Within the ISP, the identifier is decrypted and used for routing. Based on this observation, I devised a simple encryption-based routing mechanism that is efficient while seemingly random from outside of the network. For the application layer, we focused on web browsing and showed that our approach with pseudonyms is feasible and expressive with a prototype implementation.

Systems for unreliable cloud services. Today, much user data is stored and processed in cloud services. This allows users and applications to access data easily from any device and to process even very large data sets. Unfortunately not all cloud services are reliable or trustworthy. It is not surprising to hear a news story that cloud services leak users’ personal data. With my coauthors, I looked into reliability and privacy issues in file synchronization services, one of the most popular cloud services [9].

Our approach, MetaSync, combines multiple existing providers to provide a secure and reliable file synchronization service. The key challenge was to maintain a globally consistent view of synchronized files across clients using only the service providers’ unmodified APIs without any additional centralized servers. To solve the problem, I devised an efficient client-based Paxos algorithm that maintains globally consistent state among multiple passive storage backends. The approach works with popular storage services, such as Dropbox and OneDrive. Combined with content-based addressing, hash trees, and deterministic file mapping, we were able to show the feasibility of the design. Our prototype even outperformed the underlying services in synchronization speed.
Other Work

In addition to my dissertation work, I have also explored various research directions. Most are still systems research but with connections to other fields, such as ubiquitous computing and applied machine learning.

**Enabling spoken natural language interactions in applications.** Speech interaction is an emerging option for mobile phone users, and would play an even greater role with wearables or IoT (Internet of Things) devices. While commercial products provide natural ways to interact with mobile applications, they work for only a few specific domains. We explored ways to enable any third-party developer to add a spoken natural language interface to their application [8]. The key challenge behind this is that there exists much variability even for a simple command. To address this challenge, we used crowdsourcing and machine paraphrasing to cover as much of the input space as possible, then fully integrated the result into an integrated development environment.

**Continuous mobile vision.** With the advance in wearable devices and computer vision algorithms, applications may use a live video stream as an input in the near future. Our preliminary study [7] showed that these applications would use more resources from the mobile device than previously thought. I then explored ways by which we can provide efficient system support for continuous mobile vision (in collaboration with researchers from Microsoft Research). We tackled this issue from two directions. First, we looked into a camera architecture that uses low-power sensors to selectively enable video sensing [6, 18]. Then, we built a framework for running deep neural networks, MCDNN [10]. With machine learning techniques, MCDNN can generate multiple variants of a model having different resource use and accuracy, then schedule requests across a mobile device and the cloud to get high accuracy while remaining within the energy and cost budgets.

I have also worked on projects in learning graphical models [17], understanding user behavior in team games [14], and security testing [2]. Prior to starting my Ph.D., I participated in research for GPU acceleration of SSL [12], and social network analysis [1].

Future Directions

Looking forward, I intend to continue research in networked systems along with security/privacy and applied machine learning. With advances in hardware enabling new applications, systems face many challenges. The following are examples of future research topics that I am interested in pursuing:

**Scalable anonymous micropayments.** New forms of decentralized cash, such as Bitcoin, have the potential to revolutionize many aspects of electronic payment systems. While Bitcoin has shown that a block chain model can be viable in establishing a banking system without centralized trust, the current design lacks several important properties required to expand its usage (e.g., scalability, short latency, anonymity guarantees). I intend to create a micropayment system that can support anonymous low-cost, high volume transactions based on applying distributed systems ideas, like delegation and sharding, to Bitcoin. Then, we can use the micropayment system in the context of applications that process a large number of transactions; for example, it can be used as a payment method to provide economic incentives in the Tor anonymizing overlay network.

**DNN as a service.** Research in deep neural networks has been rapidly improving the accuracy of various classification tasks. While much research effort has been devoted to faster training, the execution engine has not been studied as much. With the increase in applicability of DNNs, there will be many applications that use DNNs and a huge number of concurrent requests to execute DNN classification tasks. Especially with extensive use of GPUs, I expect that there is need for better scheduling algorithms to exploit hardware parallelism in handling multiple requests, as well as a system for load balancing and scaling. Unlike MCDNN that schedules across a mobile device and the cloud, this work would focus on providing DNN as a cloud service supporting a large number of users and classification tasks.
Security and privacy issues in wearables and augmented reality. Working on both security/privacy and continuous mobile vision naturally leads me to be interested in security and privacy issues with the emerging wearables and augmented reality devices (e.g., hololens, oculus). As envisioned by the continuous mobile vision project, these devices may continuously monitor users’ surroundings and let applications respond to the sensing inputs. This will create new security and privacy challenges as well as research opportunities. For example, much computation on the video inputs may be located in the cloud. As this video stream contains much richer private information, we will need better privacy protection as users may not trust the cloud. Another issue is how to handle such audio and video inputs and determine which applications to process it. Since machine learning cannot always provide a correct answer due to its statistical nature, careful consideration is required when designing the permission and capability system.

These are only few examples of my future research directions. As shown by my previous work, I have very broad interests across many areas in computer science. I am looking forward to working on challenging problems and collaborating with researchers across multiple fields.

References


