# Research Statement

Ziming Zhao

Cybersecurity is of great importance. Given the ever increasing complexity in attack and camouflage techniques (the technology side) and cybercrime workflow (social and human side), it is imperative to have a holistic view of the today's security and privacy landscape from the machine, network, human and society perspectives so as to design effective defense and response mechanisms that span multiple layers in our computing systems and can even utilize human and social factors.

My research not only includes the security and privacy related problems in computer and communications systems but also touches the security implications of human and social behaviors. I consider myself as a *full stack security researcher* who works on multiple layers. In particular, my research foci include system and software security, e.g., utilizing hardware primitives to design and implement secure systems for attack mitigations [16, 15, 9, 7, 8, 6]; network and web security, e.g., designing and implementing novel network defense and response systems for the emerging software-defined network and mobile ad-hoc network [22, 21, 3, 5, 2, 10, 13]; cybercrime and threat intelligence analysis, e.g., developing methodologies and systems to understand the structure of underground communities and the economy and ecosystem of cybercrime [17, 19, 11, 14, 23]; usable and user-centric security, e.g., using characteristics of human perception to find vulnerabilities in existing security solutions and design more secure and usable systems [20, 18, 4, 1]. My interdisciplinary research requires expertise in computer architecture, operating system, computer networks, artificial intelligence, and human factors.

My research outcomes have appeared in the most prestigious computer and communication security conferences and journals, including including IEEE Symposium on Security and Privacy (**Oakland**), USENIX Security Symposium (**SECURITY**), Network and Distributed System Security Symposium (**NDSS**), European Symposium on Research in Computer Security (**ESORICS**), Annual Computer Security Applications Conference (**ACSAC**), ACM Symposium on Access Control Models and Technologies (**SACMAT**), ACM Conference on Data and Applications Security and Privacy (**CODASPY**), IEEE Conference on Communications and Network Security (**CNS**), ACM Transactions on Information and System Security (**TISSEC**), IEEE Transactions on Dependable and Secure Computing (**TDSC**), etc.

My research results are also recognized for their academic and practical impacts, as noted by two best paper awards (ITU Kaleidoscope 2016, ACM CODASPY 2014), the 3rd place in SDN Innovation Challenge (Extreme Networks 2015), and top 10 finalist of best applied security paper award (CSAW 2015). Moreover, my research work has been widely covered in popular media including InformationWeek, Slashdot, NakedSecurity, ACM.org, etc.

## System and Software Security

I believe software and system security is the foundation of our computer and communication security. To defeat a variety of attacks on software and systems, we first need to understand and detect such attacks. My research in this area focuses on using hardware security primitives, program analysis and machine learning to solve security challenges we are facing.

**Binary Code Reuse for Extracting Malware Secret [16].** As promising results have been obtained in defeating code obfuscation techniques, malware authors have adopted protection approaches to hide malware-related data from analysis. Consequently, the discovery of internal ciphertext data in malware is now critical for malware forensics and cyber-crime analysis. We present a novel approach called **ASES** to automatically extract secrets from malware. ASES first identifies and extracts binary code relevant to secret hiding behaviors. Then, ASES relocates and reuses the extracted binary code in a self-contained fashion to reveal hidden information. We implemented ASES and applied it to real-world malwares, which automatically produced plaintext information from embedded ciphertext in malwares.

**Shellcode Detection and Attribution [15].** Binary code injection is still an unsolved problem. Misuse-based detection lacks the flexibility to tackle unseen malicious code samples and anomaly-based detection on byte patterns is highly vulnerable to byte cramming and blending attacks. In addition, it is desperately needed to correlate newly-detected code injection instances with known samples for better understanding the attack events and tactically mitigating future threats. We propose a technique for modeling shellcode detection and attribution through a novel feature extraction method, called instruction sequence abstraction **ISA**, that extracts coarse-grained features from an instruction sequence. ISA facilitates a Markov chain-based model for shellcode detection and support vector machines for encoded shellcode attribution. We implemented ISA and applied it on more than 10,000 shellcode samples from over 10 classes. The results show that ISA can detect and attribute shellcode effectively and efficiently.

**Android Emulator Detection [9].** Emulator-based dynamic analysis has been widely deployed in Android application stores. While it has been proven effective in vetting applications on a large scale, it can be detected and evaded by recent Android malware strains that carry detection heuristics. Using such heuristics, an application can check the presence or contents of certain artifacts and infer the presence of emulators. However, there exists little work that systematically discovers those heuristics that would be eventually helpful to prevent malicious applications from bypassing emulator-based analysis. To cope with this challenge, we propose a framework called **Morpheus** that automatically generates such heuristics. Morpheus leverages our insight that an effective detection heuristic must exploit discrepancies observable by an application. Morpheus analyzes the application sandbox and retrieves observable artifacts from both Android emulators and real devices. Afterwards, Morpheus further analyzes the retrieved artifacts to extract and rank detection heuristics. Morpheus is a top 10 finalist in the best applied security paper award, CSAW, 2015.

## Network and Web Security

My research on network security focuses on the security challenges in the newly emerging computer networks, such as software-defined networks and mobile ad-hoc networks.

**Enable Stateful SDN Applications [2, 10].** Because OpenFlow attempts to keep the SDN data plane simple and efficient, it focuses solely on L2/L3 network transport and consequently lacks the fundamental ability of stateful forwarding for the data plane. Also, OpenFlow provides a very limited access to connection-level information in the SDN controller. These inherent limitations of OpenFlow pose significant challenges in supporting network services. To address these challenges, we propose an innovative connection tracking framework called **StateMon** that introduces a global state-awareness to provide better access control in SDNs. StateMon is based on a lightweight extension of OpenFlow for programming the stateful SDN data plane, while keeping the underlying network devices as simple as possible. To demonstrate the practicality and feasibility of StateMon, we implement and evaluate a stateful network firewall and port knocking applications for SDNs, using the APIs provided by StateMon.

**Risk-Aware Mitigation of MANET Routing Attacks [22, 21].** Mobile Ad hoc Networks (MANET) have been highly vulnerable to attacks due to the dynamic nature of its network infrastructure. Among these attacks, routing attacks have received considerable attention since it could cause the most devastating damage to MANET. Even though there exist several intrusion response techniques to mitigate such critical attacks, existing solutions typically attempt to isolate malicious nodes based on binary or naive fuzzy response decisions. However, binary responses may result in the unexpected network partition, causing additional damages to the network infrastructure, and naive fuzzy responses could lead to uncertainty in countering routing attacks in MANET. In this paper, we propose a risk-aware response mechanism to systematically cope with the identified routing attacks. Our risk-aware approach is based on an extended Dempster-Shafer mathematical theory of evidence introducing a notion of importance factors. In addition, our experiments demonstrate the effectiveness of our approach with the consideration of several performance metrics.

## Cybercrime and Threat Intelligence Analysis

Understanding the workflow, ecosystem and economy of cybercrime is of significant importance in combating with cyber threats. My research in this area focuses on developing novel social mining techniques to understand the organizational structure of underground groups and analyzing the ecosystem of underground marketplaces and telephone spam campaigns.

**Automated Threat Intelligence Fusion [12].** The volume and frequency of new cyber attacks have exploded in recent years. Such events have very complicated workflows and involve multiple criminal actors and organizations. However, current practices for threat analysis and intelligence discovery are still performed piecemeal in an ad-hoc manner. Consequently, it is imperative to automatically assemble the jigsaw puzzles of cybercrime events by performing threat intelligence fusion on data collected from heterogeneous sources, such as malware, underground social networks, cryptocurrency transaction records, etc. In this project, we propose an Automated Threat Intelligence fuSion framework **ATIS** that is able to take all sorts of threat sources into account and discover new intelligence by connecting the dots of apparently isolated cyber events. To this end, ATIS consists of 5 planes, namely analysis, collection, controller, data and application planes.

**Underground Social Mining [17, 19].** Existing research on net-centric attacks has focused on the detection of attack events on network side and the removal of rogue programs from client side. However, such approaches largely overlook the way on how attack tools and unwanted programs are developed and distributed. Recent studies in underground economy reveal that suspicious attackers heavily utilize online social networks to form special interest groups and distribute malicious code. Consequently, examining social dynamics, as a novel way to complement existing research efforts, is imperative to systematically identify attackers and tactically cope with net-centric threats. We seek a way to understand and analyze social dynamics relevant to net-centric attacks and propose a suite of measures called **SocialImpact** for systematically discovering and mining adversarial evidence. We demonstrated the feasibility and applicability of SocialImpact with a case study on 4GB underground social network data archived from the Internet.

**Underground Marketplace Analysis [23].** The majority of research examining online underground markets considers small samples of mostly English-language markets; few studies have systematically examined or compared multiple markets over long periods of time. To remedy this, we collected multilingual online underground marketplace data from 12 market forums between December 2005 and July 2011 and systematically examined and compared them to gain a deeper understanding of cybercrime. We presented the underground commerce, including stolen user data, fake identities, and attacking tools and services from the 12 multilingual online marketplaces. The migration trends, items for sale, and seller and buyer characteristics we discovered also reveal commonalities among these fraudulent markets.

**Telephone Spam Ecosystem Analysis [14].** Telephone spam costs United States consumers $8.6 billion annually. In 2014, the Federal Trade Commission has received over 22 million complaints of illegal and wanted calls. Telephone spammers today are leveraging recent technical advances in the telephony ecosystem to distribute massive automated spam calls known as robocalls. Given that anti-spam techniques and approaches are effective in the email domain, the question we address is: what are the effective defenses against spam calls? We analyzed the telephone spam ecosystem, specifically focusing on the differences between email and telephone spam. Then, we also surveyed the existing telephone spam solutions and, by analyzing the failings of the current techniques, derive evaluation criteria that are critical to an acceptable solution.

## Usable and User-centric Security

Human is always the weakest link in any system, and human is also the ultimate asset any computing system tries to protect. My research in this area focuses on understanding the role of human factors in security systems and develop novel techniques that model human perceptions and utilize game theory to enhance the security of our existing systems.

**Graphical and Gesture-based Password [20, 18].** Picture gesture authentication has been recently introduced as an alternative login experience to text-based password on touch-screen devices. In particular, the newly on market Microsoft Windows 8 operating system adopts such an alternative authentication to complement its traditional text-based authentication. We present an empirical analysis of picture gesture authentication on more than 10,000 picture passwords collected from more than 800 subjects through online user studies. Based on the findings of our user studies, we propose a novel attack framework that is capable of cracking passwords on previously unseen pictures in a picture gesture authentication system. Our approach is based on the concept of selection function that models users' thought processes in selecting picture passwords. Our evaluation results show the proposed approach could crack a considerable portion of picture passwords under different settings. Based on the empirical analysis and attack results, we comparatively evaluate picture gesture authentication using a set of criteria for a better understanding of its advantages and limitations.

**Collaborative Access Control in Online Social Networks [4].** Existing online social networks only allow a single user to restrict access to her/his data but cannot provide any mechanism to enforce privacy concerns over data associated with multiple users. This situation leaves privacy conflicts largely unresolved and leads to the potential disclosure of users' sensitive information. To address such an issue, a multi-party access control model was proposed, including a systematic approach to identify and resolve privacy conflicts for collaborative data sharing in OSNs. We further study the problem of analyzing the strategic behavior of rational controllers, where each controller aims to maximize her/his own benefit by adjusting her/his privacy setting in collaborative data sharing in OSNs. We formulate this problem as a multi-party control game and show the existence of unique Nash Equilibrium (NE). In addition, we conduct user studies of the multi-party control game to explore the gap between game theoretic approaches and real human behaviors.

### Future Research: IoT and Automobile Security

I plan to continue my research in the aforementioned four directions. In addition, I have started new research projects in IoT and automobile security. I have formed teams of students and post-docs with strong computer science and electrical engineering background to work on IoT and automobile security. I am applying my research experience in software, system, and network security to these newly emerging areas.

## References

[1] G.-J. Ahn and Z. Zhao. Methods, systems, and media for measuring quality of gesture-based passwords, June 30 2015. US Patent US9069948 B2.

[2] W. Han, H. Hu, Z. Zhao, A. Doupé, G.-J. Ahn, K.-C. Wang, and J. Deng. State-aware network access management for software-defined networks. In *ACM Symposium on Access Control Models and Technologies (SACMAT)*. ACM, 2016.

[3] H. Hu, G.-J. Ahn, W. Han, and Z. Zhao. Towards a reliable sdn firewall. In *2014 Open Networking Summit Research Track (ONS)*. USENIX, 2014.

[4] H. Hu, G.-J. Ahn, Z. Zhao, and D. Yang. Game theoretic analysis of multiparty access control in online social networks. In *ACM Symposium on Access Control Models and Technologies (SACMAT)*. ACM, 2014.

[5] H. Hu, W. Han, G.-J. Ahn, and Z. Zhao. Flowguard: Building robust firewalls for software-defined networks. In *3rd ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN)*, 2014.

[6] Y. Jing, G.-J. Ahn, H. Hu, H. Cho, and Z. Zhao. Triplemon: A multi-layer security framework for mediating inter-process communication on android. *Journal of Computer Security*, pages 1–22, 2016.

[7] Y. Jing, G.-J. Ahn, Z. Zhao, and H. Hu. Riskmon: Continuous and automated risk assessment for mobile applications. In *ACM Conference on Data and Application Security and Privacy (CODASPY)*. ACM, 2014.

[8] Y. Jing, G.-J. Ahn, Z. Zhao, and H. Hu. Towards automated risk assessment and mitigation of mobile applications. *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2015.

[9] Y. Jing, Z. Zhao, G.-J. Ahn, and H. Hu. Morpheus: Automatically generating heuristics to detect android emulators. In *30th Annual Computer Security Applications Conference (ACSAC)*, 2014.

[10] H. H. K.-C. W. G.-J. A. Z. Z. Juan Deng, Hongda Li and W. Han. On the safety and efficiency of virtual firewall elasticity control. In *24th Network and Distributed System Security Symposium (NDSS)*, 2017.

[11] K. Liao, Z. Zhao, A. Doupé, and G.-J. Ahn. Behind closed doors: Measurement and analysis of cryptolocker ransoms in bitcoin. In *APWG Symposium on Electronic Crime Research (eCrime)*, 2016.

[12] A. Modi, Z. Sun, A. Panwar, T. Khairnar, Z. Zhao, A. Doupé, and G.-J. Ahn. Towards automated threat intelligence fusion. In *IEEE International Conference on Collaboration and Internet Computing (CIC)*, 2016.

[13] N. T. V. D. L. S.-Z. Z. A. D. Sukwha Kyung, Wonkyu Han and G.-J. Ahn. Honeyproxy: Design and implementation of next-generation honeynet via sdn. In *IEEE Conference on Communications and Network Security (CNS)*, 2017.

[14] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn. Sok: Everyone hates robocalls: A survey of techniques against telephone spam. In *37th IEEE Symposium on Security and Privacy (Oakland)*, 2016.

[15] Z. Zhao and G.-J. Ahn. Using instruction sequence abstraction for shellcode detection and attribution. In *1st IEEE Conference on Communications and Network Security (CNS)*, 2013.

[16] Z. Zhao, G.-J. Ahn, and H. Hu. Automatic extraction of secrets from malware. In *Working Conference on Reverse Engineering (WCRE)*, pages 159–168. IEEE, 2011.

[17] Z. Zhao, G.-J. Ahn, and H. Hu. Examining social dynamics for countering botnet attacks. In *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–6. IEEE, 2011.

[18] Z. Zhao, G.-J. Ahn, and H. Hu. Picture gesture authentication: Empirical analysis, automated attacks, and scheme evaluation. *ACM Transactions on Information and System Security (TISSEC)*, 2015.

[19] Z. Zhao, G.-J. Ahn, H. Hu, and D. Mahi. Socialimpact: Systematic analysis of underground social dynamics. In *European Conference on Research in Computer Security (ESORICS)*, pages 877–895, 2012.

[20] Z. Zhao, G.-J. Ahn, J.-J. Seo, and H. Hu. On the security of picture gesture authentication. In *22nd USENIX Security Symposium*. USENIX, 2013.

[21] Z. Zhao, H. Hu, G. Ahn, and R. Wu. Risk-aware mitigation for manet routing attacks. *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 9(2):250–260, 2012.

[22] Z. Zhao, H. Hu, G.-J. Ahn, and R. Wu. Risk-aware response for mitigating manet routing attacks. In *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–6. IEEE, 2010.

[23] Z. Zhao, M. Sankaran, G.-J. Ahn, T. J. Holt, Y. Jing, and H. Hu. Mules, seals, and attacking tools: Analyzing 12 online marketplaces. *IEEE Security & Privacy Magazine*, 2016.