Statement of Purpose

Han Zhang

When I first decided to pursue computer science, I did it simply because I enjoy applying the power of logic to build products and to solve challenging problems. Upon starting my undergraduate degree at the University of Michigan, I thought the best I was going to get out of it was a comprehensive view of how computers and networks work. What ended up surprising me was that the deeper my understanding of the technologies became, the better I realized how dramatically they impact our modern society. Millions of people all over the world are using third party services at the moment. In the meantime, tremendous amount of information flows through computer networks. The unprecedented scale at which technology thrives implies that now is an exciting time to dive into computer science research, which will have significant impacts on people's lives in the upcoming decades.

Despite the growing enthusiasm of entering the industry in computer science, I am more inclined to pursue a Ph.D. degree and start a career in academic research to make greater contribution. This is not only because I believe that tackling challenging research problems is more important and impactful than simply building products for consumers with existing technology, but, more importantly, because I enjoy the process of conducting research, which I learned from my previous undergraduate research experiences.

For the past year and a half, I have been working with Professor Harsha V. Madhyastha on a project to design a mobile infrastructure to infer implicit user recommendations while preserving privacy. I joined this project at a relative early stage, and I found many open challenges in our way to alleviate the scarcity of reviews on most recommendation services, like Yelp. To better define our problem and solidify the potential challenges on passive recommendation inferences, I collected measurements on existing recommendation services to demonstrate the current shortage of reviews, developed several location-based inference prototypes, and implemented them on Android to test their feasibility. As for tackling the privacy-accuracy tradeoffs, we examined many prior works in similar and different fields, explored multiple different methods in gathering sensor inputs and data storage, and concluded with a system design that preserves privacy to the greatest extent while enabling recommendation services to make inferences. Our work is published at ACM HotNets 2016, and we are currently building a prototype to resolve the obstacles addressed in our paper.

Prior to joining Prof. Madhyastha, I worked with Dr. Michalis Kallitsis at Merit Network, which is a regional ISP in Michigan, analyzing anomalous network traffic in their R&D department during my sophomore year. Upon starting, I wasn't familiar with the networking concepts necessary for network monitoring and designing anomaly detection algorithms. I read a couple of textbooks as well as went over many online resources like open courses and slides to make up for the knowledge gap. In order to display real-time analysis output in a more interactive way, I

also taught myself full-stack development and built an interactive web dashboard. Nowadays, my dashboard is still in use to monitor network traffic. This early-stage research opportunity at Merit equipped me with useful network knowledge, and led me into the field of academic research. I got hands-on experience with the real world ISP operators and learned more about networking than what I could get only from classrooms.

As for my future Ph.D. research, I would like to spend the following years tackling various issues around distributed and networked system developments. **One of the most interesting areas is computer networking, and Software Defined Networking(SDN) and Network Functions Virtualization(NFV) in particular.** SDN and middleboxes drastically improve the versatility of networks by introducing functions and smartness into the network. How can we design the architecture to make software-based networking services faster and more reliable? How to optimize existing systems to incorporate more functionalities and control logics into middleboxes? Can we apply machine learning to networking software and benefit from this application? For specific scenarios like improving IoT security or revamping certain network designs, how could we do a better job with the help of all these new tools? SDN and NFV are like a networking programming language, and it is equally important to improve the language itself and to applying it to create killer apps.

Another attractive field is system security and privacy, and I would like to tackle these challenges from a system-building point of view. With the recent technology advancements and their increased popularity, more and more user information is available for developers at low cost. How can we design systems or frameworks to minimize privacy concerns and deter information abuse? Can we redesign networks in such a way that we can provide better user protection, while preserving the system's usability and flexibility to attract more users and developers ? These motives are very similar to the reasons why I joined Prof. Madhyastha's project. I am honored and very glad that I have gained much experience over the years.