Personal Statement

Wei Xu

My name is Wei Xu. I am currently majoring in communication engineering at Beijing University of Posts and Telecommunications (BUPT). I would like to be enrolled in the Master of Computer and Information Technology program in University of Pennsylvania because I would very much like to become a qualified engineer ready to contribute to the development of big data, and this program can offer me an opportunity to study in computer science.

During my time at BUPT, I have come to understand the basics of communications, and learned to evaluate and design the hardware devices for communications in electric circuit and digital signal processing. But I have to say that programming and networks in the computer science-related courses appealed to me a lot in the first two years because I gained a lot of pleasure and sense of accomplishment in the practice of its fields. For example, I developed a simple Battle City game with C++ as my curriculum project in freshman year, the teaching assistant even doubted whether I wrote it myself!

In order to improve my programming skill, in my sophomore year I started working on coding project in my own time. I developed an Android application to set up a platform for in-school students to share and check their assignments for their courses. Debugging was painstaking but fun and I obtained solid skills in MySQL, Java and server establishment. My exploration in mobile application and network programming helped me understand the basic client/server model and the protocols it uses.

To accomplish something with knowledge about programming language was inspiring to me, but I wanted to try more creative and complex work. Ever since studying cloud computing, I have been fascinated by the power of distributed systems and how it deals with large amounts of data. And I learned that in addition to exchanging data, we could also achieve a higher level of calculating power in a cluster by connecting several computers together. I believe that the way we handle big data will play an important role in analysis of people's behaviors and in prediction for the future in all fields. Cloud computing accelerate calculation of massive statistics and enables us to have easy access to these analyses.

So in summer of my junior year I started my learning and working on distributed systems and cloud computing at the Research Center for Mobile Computing at Tsinghua University. Our research was mainly about reconfigurable processors. We worked on a solution for parallel computing on cluster built with reconfigurable processors. Setting up a distributed system requires knowledge from processor structure to virtualization. I started with the memory allocation of the processor to generate the configuration information of Reconfigurable Processing Unit (RPU), since automatic code compilers for reconfigurable processors were not yet available. Although this meant working with hardware, they were the underlying condition for virtualization and thus for virtual machines, which were strongly needed by cloud computing. After fundamental the hardware system was built, I was able to design a large matrix multiplication on the Hadoop platform for evaluation of the function of the processor and we were then able to run a more complicated video decoding job on it. The results were much better compared with general purpose processors.

Meanwhile, in the second semester of my junior year, I entered the Future Network Theory and Applications Laboratory (FNTAL) at our university and began research on artificial neural networks, the focus of my course project. It was the first time I really learned knowledge of machine learning. I read papers about feed-forward and recurrent networks. I learned how certain artificial neural networks work under supervised or unsupervised learning in order to make predictions and classifications. Then I worked as an assistant of a graduate student, who had proposed two novel Echo State Networks (ESNs), to optimize the ESNs effect on prediction of user behavior and to minimize the error in prediction results.

I had never connected my experience at FNTAL about artificial neural networks with my work on cloud computing, until I found how neural networks were deployed on a Hadoop distributed system and how they were trained with data sets efficiently. It led me to consider the combination of big data and machine learning and showed me a new research and career path I would like to follow.

I am currently completing an internship in the music department at Baidu, Inc. My duties include singer recommendation systems based on data mining and machine learning. In this position I have access to a huge amount of data on users' behaviors in listening to music ad in charge of analyzing these data in order to recommend singers to every user according to their tastes. Since there are tens of thousands of user records input every day and up to millions every week, to find out what kind of singer the users may like is not an easy job, let alone the recommendation for millions of people. We must design algorithms to deal with users' action data from database to calculate the relevance of every singer to every user and rank them in order for recommendation. Research in industry has advantage over that in laboratory because you get sufficient feedbacks promptly. The internship has been very valuable in helping me to develop applicable programming skills in a very short time. I have gained a lot of experience in rewriting MapReduce method in Python scripts and in using Hadoop streaming tools to process such massive data. I also learned how to do data analyses for data mining. In addition, working in a team teaches me how to collaborate properly and efficiently. So far this internship at Baidu has served only to strengthen my passion for computer science.

However, since my undergraduate experience focused more on communication rather than computer science, I have sensed the disadvantages of lacking in fundamental knowledge in computer science in my internship. The MCIT program is just suitable for me because it provides me with more systematic trainings in programming and a wonderful pool of other computer science courses that are necessary to me, who doesn't have a strong background in computer science. Now I'm ready to face challenges I may meet in pursuing my goal and I think this program in Computer and Information Science will no doubt give me an opportunity to prepare me with professional skills for my future career.