Clement Fung - Statement of Purpose for Cornell Computer Science PhD Application

I want to pursue a PhD in computer science, specifically in the field of computer security. I have a particular interest in securing and democratizing interactions with distributed machine learning systems. Given the increasing popularity of decentralized architectures like Google's federated learning and the recent discoveries of attacks on models such as backdooring and attacks on data providers such as property inference, I consider this area to be both incredibly interesting and of paramount importance. Specifically, I am interested in developing robust systems for secure and private multi-party machine learning, in which users have full control and privacy over their own data.

I was first introduced to security research during my masters degree at the University of British Columbia (UBC), where I have been a research assistant in the Networks, Systems and Security (NSS) Lab for the past 2 years. Throughout my masters degree, I was supervised by Professor Ivan Beschastnikh, whose expertise spans from distributed systems to systems security. Together, we completed three different projects in the area of secure and private multi-party machine learning:

- Biscotti[1]: A private and secure distributed ledger for peer to peer machine learning

- FoolsGold[2]: A protocol for detecting and mitigating sybil-based poisoning attacks on federated learning

- TorMentor[3]: A system and protocol for private, secure machine learning over an anonymous network

For each of these projects, the central theme was: how can we modify distributed multi-party machine learning systems in ways that protect the privacy of their users while maintaining the integrity of the learned model?

One solution to providing more privacy and security to distributed multi-party machine learning is to eliminate the centralization present in modern architectures such as federated learning. I worked on developing an alternative peer-to-peer solution that does not rely on a centralized process to store and coordinate the training process, called Biscotti. The peer-to-peer setting requires a novel threat model in ML, and Biscotti adapts elements of distributed ledgers, such as proof of stake, block verification, and cryptographic commitments to ensure a private and secure mechanism for peer-to-peer machine learning through a ledger-based structure. This system was built and designed over a 6 month period in collaboration with another graduate student, in which I focused on developing the distributed machine learning algorithms and implemented machine learning attacks and defenses, while they worked on the consensus and cryptographic elements of the project. This was my first experience in dealing with the teamwork required in successfully sharing the duties of system development, evaluation and paper writing in co-authoring a top tier conference submission. Ultimately, we designed and implemented a system that enables peer-to-peer, private, secure machine learning at scales up to 100 peers, matching model convergence results from federated learning. The system provides state-of-the-art privacy through differential privacy and secure aggregation while using blockchain primitives to prevent sybil attacks.

Sybil attacks are also relevant to federated multi-party learning systems; I am highly interested in defending these systems and thus developed FoolsGold, a mechanism for protecting federated learning systems from sybil-based targeted poisoning attacks. Prior work in this space relies on assumptions of a bounded proportion of attackers, and relies on direct analysis of the training data, which cannot be applied to privacy-preserving machine learning. In FoolsGold, I identified that the similarity of gradients between clients was an effective tool for detecting sybils and designed a penalization function for thwarting targeted poisoning attacks. Unlike prior defenses, this mechanism can actively resist an attack from a system with 99% sybils and unlike prior work, does not rely on observation of client training data, which makes it suitable for federated learning systems. This work is currently in submission at the 2019 IEEE European Symposium on Security and Privacy.

---

[1] ArXiv, November 2018. https://arxiv.org/abs/1808.04866
[2] ArXiv, August 2018. https://arxiv.org/abs/1808.04866
[3] ArXiv, November 2018. https://arxiv.org/abs/1808.04866

In TorMentor, I augmented stronger defenses onto federated learning by using anonymous onion routers as the communication medium in distributed learning. Through anonymous communication, I defined a new learning paradigm called brokered learning, in which data providers and model curators do not need to directly communicate with each other; instead, they coordinate with third-party brokers to perform distributed machine learning. In doing so, model definers are no longer the central authority on the training process: Unlike prior work which requires trusting the model curator to provide privacy, TorMentor gives more control to clients and performs secure and anonymous machine learning in a democratic fashion: providing privacy and control to data providers while concurrently attempting to attain optimal model performance.

My time spent pursuing research at UBC solidified my interest in continuing my research career. I love collaborating with great people on engaging challenges. I was fortunate enough to meet and learn from some superb researchers while at UBC, and working on these unique and varied projects has really reinforced one of my reasons for doing a PhD. I cannot think of a better environment than a PhD, which brings together such students and allows them to challenge the frontier of research. PhDs are great networking tools for meeting interesting, motivated and inquisitive people.

In addition to the systems I have built at UBC, I have experience building large scale systems over 2 internships spanning 8 months at LinkedIn. For both terms, I was a member of the distributed data systems team, working on Apache Kafka and Apache Samza pipeline integration with existing systems on the Voldemort team (distributed database) in 2014, and on the Online Relevance Infrastructure team in 2015. Following the conclusion of my Masters degree in December, I will be furthering my industry experience by spending 6 months working at Oasis Labs[4], an early stage privacy-preserving blockchain research startup founded by Professor Dawn Song from UC Berkeley. I am very excited for the opportunity to continue working with world-class people in a fast-paced environment and look forward to sharpening my skills in systems development, machine learning and security for my eventual PhD research.

While my work so far is a first step towards securing distributed multi-party machine learning, there are still several unanswered questions left. For example, backdoor poisoning attacks that do not target full class labels are still very difficult to detect, and tracing their source is difficult in distributed multi-party settings. I notice that this research problem is part of the larger problem of security and accountability in distributed multi-party machine learning. One particular idea that I would like to pursue builds upon my work in Biscotti to allow retraining of machine learning models after poisoning attacks or information leakages are detected and their data sources are identified. This idea would leverage ideas from data provenance and auditing frameworks on the blockchain to allow model providers to rapidly apply security and privacy patches to models, ideally without requiring the original datasets or the associated clients. There are several methods for tackling this large research problem given the tremendous volume of new attacks and defenses being discovered in this field, coupled with the increase of distributed multi-party architectures, and I am excited to contribute to this important and rapidly growing field of research.

I would like to pursue a PhD in an environment with strong expertise in the field of machine learning security. At Cornell University, I am particularly intrigued by the information leakage work of Professors Vitaly Shmatikov and Thomas Ristenpart, who have developed several methods for attacking and quantifying information leakages in machine learning models. Both professors have done additional work in augmenting privacy and security to federated learning, which is an area I have particular experience with. As federated learning and other distributed machine learning architectures continues to grow in popularity, understanding and defending against privacy leakages will become increasingly urgent.

Thank you for your time in considering me as an applicant of the PhD program at Cornell University.

---

[4]https://www.oasislabs.com/