Research Statement

Jun (Jim) Xu  
July 23, 2005

My research focuses on the theoretical and foundational aspects of computer networking. The phenomenal growth of the Internet and its applications has far outpaced the development of fundamental principles of designing network algorithms and performing network measurements. By rigorously establishing such principles, my research has led to much better algorithms and measurement techniques, and has helped to lay the scientific foundation for networking. My work consists of the following three related themes:

The first theme of my research helps establish an emerging research area in computer networking called network data streaming. This area deals with the extraction of important summary information from a stream of packets passing through a high-speed network link (e.g., 40 Gbps), using a small amount of high-speed memory. It has numerous important applications in network measurement, monitoring, and security. I have produced several strong results [II.C.1-15, 17, 19, 20, 21, 23, 24, 25] that are among the earliest in this area, and have received a Best Student Paper Award (for II.C.1-20) from a prestigious international conference (Sigmetrics/Performance 2004). I have also obtained an NSF grant (single-PI) for continued innovations in this area. Section 1 will describe this research theme in detail.

The second theme of my research studies the fundamental tradeoff and lower bound problems in networking and distributed computing [II.B-7, II.C.1-7, 8, 12, 16, 18]. Such results are important since they typically end the fruitless search for a better algorithm that does not exist, or lead to provably optimal algorithms. Obtaining such results is difficult because they require a deep understanding of the domain knowledge and a solid background in theoretical computer science and information theory. My efforts result in the solution of several well-known and long-standing open problems in this area, and inspired many follow-up works by others. I received an NSF CAREER award in 2003 for continuing this effort. Section 2 describes this research theme in detail.

The third theme of my research is the continuation of my Ph.D. thesis research on network security and cryptography. Making the Internet practically secure is a hard, wide-range, and pressing problem. My research effort on this problem (after graduation) is focused on two topics: countering denial of service attacks and anonymizing Internet traffic for research use. I have generated a set of high-impact and pioneering results in this area [II.B-5, 6, 9, II.C.1-10, 11, 13, II.C.2-1] and received an NSF ITR grant (as lead PI) for this effort. Section 3 describes this research theme in detail.

1 Network data streaming

With the rapid growth of the Internet, network link speeds have become faster every year to accommodate more Internet users. Measuring and monitoring the traffic on such high-speed links has become an ever challenging problem. I am among the first to realize that data streaming
algorithms have the potential to become very useful tools for measuring and monitoring high-speed networks. Data streaming is concerned with processing a long stream of data items in one pass using a small working memory in order to estimate certain statistics of the stream. The challenge is to use this small memory to "remember" as much information pertinent to this estimation as possible. However, existing data streaming algorithms are typically designed for processing a single stream of data for a single type of statistic, and most of these algorithms cannot operate at very high link speed. Targeting the deficiency of existing approaches, I have investigated new data streaming paradigms and mechanisms that allow us to perform large-scale distributed data streaming on tens of thousands of high-speed links and nodes, and aggregate, compress, and interpret these streaming results, for better measurement and monitoring of large networks. My data streaming work spans the following three related themes.

The first theme is to design next-generation single-node data streaming algorithms that can operate at high link speeds of 10 to 40 Gigabits and can still provide high measurement accuracy, based on novel methodology called "Lossy data structure + Bayesian statistics = accurate streaming". The main idea of this methodology is to first perform data streaming at very high speed to get the streaming results that are lossy. This loss is inevitable due to the stringent memory and computational complexity constraints. Because of the loss, the streaming results may deviate significantly from the information we would like to estimate. Bayesian statistics will then be used to recover information from the streaming results as much as possible.

This methodology is applied to producing four strong and high-impact results that are also among the earliest in this area. The first result (Infocom'04), based on a novel streaming data structure called Space-Code Bloom Filter (SCBF), solves the open problem of measuring approximate per-flow traffic without keeping per-flow state. This result has been applied to several other network measurement problems, such as detecting superspreaders and our ongoing global iceberg work (described below). The second result (Sigmetrics'04) solves the well-known problem of accurately estimating flow size distribution (FSD), without keeping per-flow state. This result is important also because the flow size distribution can be used to derive any arbitrary statistic of the packet stream such as the mean flow size (the first moment). Our algorithm is orders of magnitude more accurate than the traditional sampling-based approach. This work won the Best Student Paper Award at the ACM SIGMETRICS/IFIP PERFORMANCE 2004 Joint International Conference on Measurement and Modeling of Computer Systems. The third work (Sigmetrics'05) extends the FSD work to solving an important and much more challenging problem: estimating the FSD of any arbitrary subpopulations (e.g., the FSD of all web traffic). The ability to monitor the SubFSD for any arbitrary subpopulation provides a powerful and versatile tool for the characterization of traffic traversing a monitoring point. In particular, several well-known network measurement problems, such as hierarchical heavy hitter detection and the estimation of the number of active flows in a subpopulation, are special cases and/or strictly weaker than SubFSD estimation. They are solved as a byproduct of our work. The fourth result (IMC'05) provides accurate and efficient solutions to the problem of detecting superspreaders, sources/destinations which has communicated a large number of distinct destinations/sources during a short time interval. This problem arises in many network security and measurement tasks such as detecting port scans, worm propagation, DDoS attack, flash crowds, and bottlenecked content distribution servers. Our solution is based on a novel insight that sampling and streaming are often suitable for capturing different and complementary regions of the information spectrum, and a close collaboration between them is a good way to recover the complete information.

The second theme, called distributed coordinated data streaming, is to dramatically improve the efficiency of information dissemination in a large computer network using novel data streaming techniques. We envision the following information dissemination scenario. Each node in the network
will generate a long stream of data over time. Only some (typically a small portion) of the data will be of interest and should be delivered to certain other nodes, and flooding all the data to all the nodes is clearly not scalable and bandwidth-efficient. In this scenario, information routing and dissemination can be performed very efficiently using our coordinated data streaming approach as follows. The stream of data from each node will be filtered and compressed by its neighbors, neighbors' neighbors, and so on. Every node will perform such data streaming for all the other nodes, and maintain a synopsis data structure to store the streaming results. This synopsis data structure can be viewed as a probabilistic forwarding table that provides hints to interested nodes as to where the data are located. We have successfully applied this novel vision to solving two important problems, in the area of peer-to-peer networks and network security, respectively. In our first result (Infocom'05), we have, quite surprisingly, converted the problem of unstructured P2P query routing to a coordinated streaming problem and produced an algorithm orders of magnitude more efficient than existing approaches. Our second result, published in IEEE Symposium on Security and Privacy 2004, is the design of the most resource-efficient and scalable IP traceback scheme to date, based on this coordinated data streaming vision. This work also pioneers the application of information theory, such as mutual information maximization and Fano's inequality, to optimizing streaming data structures.

The third theme, called distributed collaborative streaming, is to solve the important problem of monitoring and analyzing the aggregate traffic passing through a number of high-speed links, for detecting “global” events that are intrinsically distributed through the network. It is hard to detect such events using traditional single-node streaming mechanisms because their signatures are usually too flexible or incomplete to be detected locally. I have proposed a novel methodology to tackle this class of problems. Its main idea is to perform preliminary streaming (barring some similarity to EE's soft-decision coding) at each measurement point, which massages traffic data into digests that are thousands of times smaller, and then to ship these digests for centralized analysis. The challenge is to devise good streaming algorithms that do not lose much relevant information (about the pattern) during this process. We apply this methodology to solving three important problems: (1) estimation of traffic and flow matrices (Sigmetrics'05), (2) detection of unknown common contents in Internet traffic (submitted), and (3) identification of global icebergs over distributed data sets (submitted).

2 Fundamental lower bound and tradeoff problems in networking

I solved several well-known and long-standing open problems concerning lower bounds and tradeoffs in networking. My first result (Sigcomm'02) concerns the minimum computational complexity needed for packet scheduling algorithms to provide a tight delay bound. It has been a long-standing open conjecture in networking research community that there is a minimum computational complexity cost to pay (usually $\Omega(n \log n)$ where $n$ is the number of sessions) for packet scheduling algorithms to achieve tight delay bounds (a type of QoS guarantee). We prove that the lower bound computational complexity of any scheduling algorithm that guarantees $O(1)$ delay bound is indeed $\Omega(n \log n)$ per packet. We also discover that the complexity lower bound remains surprisingly the same even if the delay bound is relaxed to $O(n^a)$ for $0 < a < 1$. This implies that the delay-complexity tradeoff curve is “flat” in the “interval” $[O(1), O(n)]$. We also extend this result to the context of end-to-end delay bounds.

The first result has led me to solve another open problem concerning the minimum complexity of tracking GPS clock, an essential function in perfect packet scheduling algorithms such as WFQ and WF$^2$Q. My result (Infocom'04) debunks a long-held misconception that this complexity is
\( \Omega(N) \) per packet, which is cited in many papers and even textbooks. I showed that there is in fact an \( O(\log N) \) algorithm and proved that \( \Omega(\log N) \) is also the lower bound. These two results have generated considerable impact in a short period of time. They have ended the search for scheduling algorithms that would provide better delay-complexity tradeoffs and have been used for proving the optimality of several scheduling algorithms (e.g., UCSD’s Stratified Round Robin). They have also been adapted (by Prof. Anderson and Jeffay at UNC) to the systems area for establishing the complexity lower bounds of CPU scheduling algorithms.

The third open problem I solved (Infocom’03) concerns the fundamental tradeoff between routing table size and network diameter in structured P2P networks. This result has inspired studies on other related tradeoff issues (e.g., by Dr. Dworkadas of Rochester) in P2P networks. It has also led my research team to design Ulysses (ICNP’03), one of the first protocols to achieve the optimal tradeoff, and has inspired some competing protocols designed by others.

3 Research on Network Security

In this wide-range area, I have focused on two pressing and specific research problems: countering distributed denial of service (DDoS) attacks and anonymizing Internet traffic for research use. I have made three major contributions to the first problem (DDoS defense). First, I have designed a practical and effective defense system to protect web services from distributed DoS attacks. It allows a web site to sustain high availability of its services even under severe DoS attacks. This is of paramount importance, because the web is the core technology underlying E-commerce and primary target for recent DoS attacks. A paper describing the design and performance evaluation of this system appears in IEEE Transactions on Computers, Special issue on Reliable Distributed Systems. Second, I have investigated protocol-independent techniques for significantly improving the throughput of legitimate traffic during DDoS attacks (ICNP’02). The baseline methodology is to preferentially drop packets that are more likely to come from attackers, based on the location information about the attackers learned using a generalized form of IP traceback (techniques to trace the origin of an attack). The third result is cross-listed in Sec. 1, which applies the aforementioned distributed coordinated streaming/sampling technique to IP traceback.

On the second problem of anonymizing Internet traffic, we have developed a cryptography-based, prefix-preserving IP address anonymization technique (ICNP’02) that is provably as secure as the existing well-known TCPdpriv scheme, and unlike TCPdpriv, provides consistent prefix-preservation in large scale distributed setting. Through the analysis of Internet backbone traffic traces, we have thoroughly investigated the effect of some types of attacks on the security of any prefix-preserving anonymization algorithm. Our software implementation of this scheme (called CryptoPAN) has been downloaded and used by many network researchers and practitioners for Internet traffic anonymization.

4 Summary

My research contributions have had significant impact on the networking field. The problems I have been working on are considered as exciting, important, and challenging by the research community. This is clear from the fact that I have received three NSF awards (totaling 1M) as the only or lead PI in the past five years. I have published more than a dozen papers in premier conferences and a dozen more in leading journals. These results are well-known and have been used as teaching material at many schools world-wide (KAIST in Korea, Tsinghua Univ. in China, UIUC, UMass, etc.). They have also lead to many follow-up works by other researchers.
My research answers more than theoretical questions: they solve real-world problems. In particular, my data streaming algorithms solve many challenging open problems in high-speed network measurement and monitoring. I have filed for four invention disclosures for these results. In addition, I have just been promised funding (1 RA per year) from Sprint on investigating new data streaming algorithms for detecting traffic anomalies and intrusions in their backbone network. In summary, I have proven to the world many times that my results, although theoretical and mathematical in nature, can have profound impact on the practical design of real networks and systems.