

# Professional Statement

by Dr. Sergey Gorinsky

March 2009

---

## Interests

My research interests are mostly in the area of networked systems and encompass its following aspects: behavior of distributed entities, utilization of shared network resources, and alignment of network services with application needs. In particular, I am interested in transmission control, access control, link scheduling, buffer management and other means of effective network and system support for media streaming, bulk data transfer, and interactive delay-sensitive applications. Transgressing the traditional paradigm of networking protocol design, I investigate systems where numerous independent stakeholders are capable of behaving differently than specified by protocols. My objective is to design new network architectures and protocols that exhibit both high performance and robustness when exposed to realistic mixes of compliant behaviors, incidental failures, selfish manipulations, and malicious attacks by networked parties. Hence, my research interests span such diverse domains as network security, fault tolerance, multimedia, real-time systems, and performance evaluation. In choosing specific problems to work on, I favor those of fundamental nature so that my solutions will remain relevant over long time and for various future applications.

## Methods

Reflecting my interdisciplinary interests, my research approach employs techniques that bridge gaps between traditionally separate fields. While different fields have different research languages and cultures, I overcome the linguistic and cultural barriers by collaborating with experts in the counterpart disciplines. My successful collaborations include those on real-time scheduling with Dr. Sanjoy Baruah from the University of North Carolina at Chapel Hill, operations research with Dr. Eric Friedman and Dr. Shane Henderson from Cornell University, and network security with Dr. Yongguang Zhang, who is currently at Microsoft Research Asia. By exposing problems in networked systems from alternative perspectives, the interdisciplinary collaborations enable me to advance not only networking knowledge but also other disciplines. For instance, my solutions developed initially for networking problems have also been found viable for classic real-time systems, web server design, and challenge-response security.

While approaching a problem, I identify a theoretical framework that offers a simple and yet realistic model of the studied problem. Control theory, queueing theory, scheduling theory, and combinatorics are among the analytic bases for my work. However, the elegant modeling is just a starting phase of the research. Behind each of my prominent contributions lies a creative idea that sheds unorthodox light on the problem and leads to a new effective solution. Well aware of the pitfalls lurking in model oversimplification, I validate the analysis with extensive simulations in more detailed models, e.g., in ns-2, and by implementing actual systems. In particular, my Applied Research Laboratory at Washington University builds and maintains a remotely accessible networking test bed ONL (Open Network Lab) that interconnects end hosts through a configurable measurable core of 18 programmable gigabit routers. ONL is proving itself a successful platform for prototyping and evaluating networked systems in controlled realistic settings.

Teaching is an integral part of my research. Mentoring a student constitutes both dissemination and creation of knowledge. In explaining a networking problem from first principles, I reexamine validity of the assumptions that underlie the existing solutions. Whereas even an originally sound assumption might become untenable due to technological or societal changes, the assumption reassessment enables me to detect new opportunities for further research and better designs. For example, my contributions to multicast congestion control and network service differentiation largely derive from questioning the universal trust and cooperation assumptions, which are common in these domains. I intentionally bring research and teaching

closer through heavy reliance of my instruction on the modern research infrastructure, including PlanetLab and ONL. All my undergraduate and graduate courses have strong experimental components that give the students valuable hands-on experience with real networked systems. The effectiveness of ONL as means for teaching and learning advanced system design is investigated in my papers published at ACM SIGCSE in 2007 and 2008.

Success in systems research hinges on the ability to build and sustain the human and technological resources needed for a vibrant research group. By awarding multiple grants with the total amount of \$4.5 million, the U.S. National Science Foundation has been the predominant sponsor of my research activities so far. I am committed to persistent securing of the resources necessary to lead a technically strong and impactful research group.

## Contributions

**Incentive-driven network service differentiation:** While networked applications have grown dramatically in their number and variety, the Internet keeps offering a single best-effort service. There exist numerous proposals that diversify network services to align them better with the divergent needs of the applications. However, these innovative architectures failed to gain wide end-to-end deployment, primarily due to economic and legacy issues rather than technical shortcomings. My SIGCOMM 2008 work introduced and explored a new paradigm for network service differentiation where design principles account explicitly for the multiplicity of Internet Service Providers (ISPs) and users as well as their economic interests in environments with partly deployed new services. The key idea is to base the service differentiation on performance itself rather than price. The proposed RD (Rate-Delay) architecture offers two best-effort services: an R (Rate) service with higher throughput and a D (Delay) service with low queuing delay at a congested network link. By marking a bit in the headers of transmitted packets, a user has complete freedom and built-in performance incentive to choose the R or D service as per the needs of the user's application. An RD router supports the R and D services on a link through simple scheduling of transmission from two FIFO (First-In First-Out) drop-tail queues. Due to the superior support for telephony, networked games, and other applications that are sensitive to end-to-end packet delays, the RD architecture is attractive for adoption by ISPs who seek to boost revenue by increasing the number of customers. Extensive simulations confirmed the effectiveness of the RD service differentiation as well as the fitness of the design for incremental virulent deployment in the Internet.

**Robust congestion control for untrustworthy networked environments:** The above SIGCOMM 2008 work contributed through questioning the assumption of cooperation between ISPs in the rollout of a new service. Another successful line of my research examined the issue of trust in network congestion control. In the context of network-layer multicast, I investigated incentives and abilities of a receiver to acquire unfairly high throughput via manipulation of the congestion control protocol. First, I identified and demonstrated various vulnerabilities of prominent multicast protocols to the receiver misbehavior [NGC 2002]. Then, I focused on protecting the multicast protocols that compose a session from several multicast groups and provide the receivers with rules for subscribing to an appropriate subset of the groups [SIGMETRICS 2003, SIGCOMM 2003, IEEE/ACM Transactions on Networking 2006]. Such protocols are vulnerable to inflated subscription attacks where a receiver ignores the subscription rules and inflates its subscription to gain an unfairly large share of the bottleneck link capacity. My solution against the inflated subscription attacks guards access to multicast groups with dynamic keys and consists of two independent components: DELTA (Distribution of ELigibility To Access) – a novel method for in-band distribution of group keys to receivers that are eligible to access the groups according to the congestion control

protocol, and SIGMA (Secure Internet Group Management Architecture) – a generic architecture for key-based group access at edge routers. I applied DELTA and SIGMA to derive robust versions of several high-profile protocols for layered multicast.

The DELTA and SIGMA approach succeeded by bridging the traditional gap between network capacity allocation and security. While prior schemes for key distribution viewed congestion as an obstacle and strove to overcome congestion-induced packet losses, key-based access was a foreign notion for congestion control protocols. In contrast, DELTA and SIGMA utilized congestion as a natural filter to distribute the keys only to those receivers that were eligible for the group access as per the congestion control protocol. The theme of integrating security with performance was further developed in my unicast congestion control work on protecting TFRC (TCP-Friendly Rate Control) against receiver misbehavior [ICAS/ICNS 2005, Journal of Internet Engineering 2009].

**Multimedia streaming and interactive applications:** In the area of effective architectural support for media streaming, telephony, networked games, and other delay-sensitive applications, I have studied various dimensions such as link buffer sizing [ISCC 2005, Simulation journal 2007] and congestion control [INFOCOM 2007, ISCC 2007]. In particular, the proposed MCP (Multimodal Control Protocol) transmits at a stable rate after converging to the efficient fair state and pioneers the explicit communication that enables a sender to urge all flows on its bottleneck links to switch into the same mode of operation. Also, while relative smoothness of the transmission rate is important for media streaming, I explored queuing delay and loss under smooth transmission as well as their scalability with respect to the number of flows [ICNP 2007]. Whereas layered multicast shows great promise for efficient streaming of video and other large-volume content over heterogeneous networks, my work in this domain included both design of scalable privacy-preserving congestion control SIM (Selective participation, Intra-group transmission adjustment, and Menu adaptation) [ICCCN 2008] and comparison of feedback-based and feedback-free schemes in regard to aligning the service with the receiver capabilities [NOSSDAV 2001].

**Bulk data transfer protocols:** Bulk data transfer is a very different type of networked applications where end-to-end delay of individual packets is mostly irrelevant. Such applications are primarily interested in minimal delay achievable for the whole message. PS (Processor Sharing), other forms of instantaneous fair sharing, and their packet-level approximations provide significantly worse average message delay than with the SRPT (Shortest Remaining Processing Time) scheduling, which is optimally efficient but might starve large messages under heavy load. My research revealed and investigated a class of fair efficient algorithms for message communication where no message is delivered later than under PS. In SFS (Shortest Fair Sojourn) and my other algorithms from this class, messages are transmitted one by one with potential preemption. Although I proved that no fair online algorithm assures the minimal average delay attainable under the fairness constraint, my analysis and extensive experiments also showed that the fair SFS and its variants are consistently more efficient than PS and deliver messages with SRPT-like average delay [INFOCOM 2006, ICCCN 2007, Simulation Modelling Practice and Theory journal 2009].

**Scalable congestion control for long high-capacity network paths:** My early but influential work "Additive Increase Appears Inferior" was a precursor of major advances in scalable congestion control for networks with long high-capacity paths. The report revealed the poor scalability of the AIMD (Additive-Increase Multiplicative-Decrease) algorithm, which is a staple component of the TCP (Transmission Control Protocol) congestion control. I argued for alternative control that employs scalable multiplicative increases and thereby improves the utilization of the long high-capacity network paths. When my report became public in 2000, its findings were met with a fair share of skepticism. However, after Dina Katabi et al. proposed XCP (eXplicit Control Protocol) in 2002, scalable congestion control turned into a topic of high interest. Since then, a lot of work is being done in this area and presented at the annual PFLDnet (Protocols

for Fast Long-Distance Networks) and HSN (High-Speed Networks) workshops as well as more general networking fora. My own subsequent related publications include the articles in NEW2AN 2004 and the October 2007 issue of the Journal of Internet Engineering.

**Other research lines:** Beside the aforementioned, my research has advanced various other fields and areas such as real-time scheduling [RTCSA 1999, Real-Time Systems 1999], performance evaluation [ICNS 2008], adaptive generation of TCP acknowledgments [Journal of Internet Engineering in January 2007], fair queuing [IWQoS 2007], and virtual-circuit networking [ICCCN 1997, INFOCOM 1997]. My publications made a truly global impact and were cited by researchers in Switzerland, Austria, Germany, Finland, Sweden, Norway, Netherlands, Belgium, France, Italy, Spain, Portugal, UK, Poland, Romania, Greece, Cyprus, Tunisia, Israel, Palestine, Saudi Arabia, India, China, Korea, Japan, Taiwan, Hong Kong, Singapore, Australia, Brazil, Canada, and USA.

## Plans

I will continue to develop secure networked systems that perform well. A major challenge in designing an efficient robust networked system lies in understanding how the distributed components of the system will behave in reality. Since networked systems are provided for and by humans, the system architecture should be resilient to accidental failures, malicious attacks, or simply selfish behaviors of components. However, the accumulated experience has shown that design for total paranoia takes typically heavy toll on performance and that some degree of trust and cooperation between independent entities is realistic. Hence, my plans involve both modeling of human behaviors and using the analytic insights to build efficient robust systems. Because this research agenda is inherently interdisciplinary, I will collaborate with computer scientists, engineers, psychologists, economists, and service providers.

Another future dimension of my integrative research is to overcome the deficiencies of the traditional layered approach to networked system design. The disregard of inter-layer dependencies has resulted in unrealistic models and suboptimal solutions. The point-to-point link abstraction for wireless communication media and segregation of network congestion control from routing exemplify the negative consequences of the layering. Also, while security is a universal concern, protecting each layer separately overburdens the design without securing the system as a whole. I plan to advance the state of the art in networked systems through development of holistic designs, both theoretically and by building real systems that bridge the gap between software and hardware technologies. My analysis will strive to capture networked system behaviors as simple relationships between macroscopic variables. The research community has already made some progress in this direction, e.g., by expressing the average TCP sending rate as a function of the average loss rate and round-trip time. Whereas such encouraging results are still limited and disconnected, I intend to work toward a unified theory based on a right set of abstractions. My practical aspiration for the integrative research is to design a networked system architecture that supports efficiently, scalably, and robustly the needs of diverse applications including delay-sensitive interaction, bulk data transfer, and media streaming.

The university environment is ideal for my research plans because of its multidisciplinary nature and educational focus. In my teaching, I will continue following the research-oriented approach that gives the students hands-on experience with networked systems, including those based on multicore processors. My research expertise qualifies me to teach undergraduate and graduate courses in communications and networking, distributed systems, operating systems, real-time systems, performance evaluation, system design, and computing theory. I remain committed to disseminating my scientific and engineering results at top research venues such as ACM SIGCOMM, IEEE INFOCOM, and IEEE/ACM Transactions on Networking. My long-term agenda includes publishing a book on the design of efficient robust networked systems.