Guest Editorial
Network Infrastructure Configuration

Component configuration is the "glue" for logically integrating network components to set up end-to-end requirements. Requirements can be on security, connectivity, performance and reliability. Each component has a finite number of configuration parameters that are set to definite values to satisfy requirements. Today, the large gap between requirements and configurations is manually bridged. This leads to large numbers of configuration errors whose adverse effects on availability, security, performance and deployment costs are well documented. For example, in his 2004 paper *Computer Security in the Real World*, Turing Award winner Butler Lampson stated that "setting it [security] up is so complicated that it's hardly ever done right. While we await a catastrophe, simpler setup is the most important step toward better security." As if to corroborate this concern, in late 2008, a government report *Securing Cyberspace for the 44th Presidency* states that "incorrect configurations … were responsible for 80% of Air Force's vulnerabilities." A 2008 report from Juniper Networks *What's Behind Network Downtime? Proactive Steps to Reduce Human Error and Improve Availability of Networks* states that "human error [arising out of network complexity] is blamed for 50 to 80 percent of network outages."

Thus, it is critical to develop tools to automatically bridge the gap between requirements and configurations. The fundamental problems that need to be solved to build these tools are:

- What are the primitive requirements and how are they composed into end-to-end requirements?
- What are formal languages for specifying requirements?
- How does one formally verify requirements?
- How does one synthesize correct configurations from requirements?
- How does one diagnose and repair configuration errors?
- How does one visualize logical relationships set up via configuration?
- How does one anonymize configuration so that the new configuration does not reveal sensitive information yet permits researchers to gain insights into how configurations are deployed in practice?
- How does one migrate current configuration into final configuration without disrupting mission-critical services or creating security breaches in the transition?
- How does one test components for whether their configurations have been correctly implemented?

and especially

- How can specification languages be made easy to use by network administrators? Unless requirements are precisely specified, configuration synthesis, migration planning and a substantial amount of diagnosis and repair are impossible.

Answering such questions is inherently hard. Requirements span multiple components at and across multiple protocol layers. A real infrastructure typically contains thousands of components, each with hundreds of configuration parameters. Compounding the challenge is the fact that security interacts with connectivity, performance and reliability. Security is about preventing undesirable behavior while others are about enabling good behavior. Incorrect resolution of this tension can disable mission-critical services and potentially cause as much harm as allowing adversary access to those services. Security and connectivity are often handled by different parts of an organization, and it is not straightforward to share configuration information for end-to-end analysis. Components can work correctly in isolation but not together to support end-to-end services. One cannot diagnose configuration errors by checking component configuration in isolation from that of others. This is because the logical structures into which these have been integrated, via configuration, are broken. Global reasoning is required to check these structures. For the same reason, configuration repair is hard: changing the configuration of one component to restore a requirement may violate another requirement. The change has to be such that all requirements are simultaneously true in the new configuration, not just the falsified one. Avoiding single points of failures requires not only correctly provisioning redundant resources and fault-tolerance protocols at a single layer but also ensuring that these resources are not mapped to the same resource at a lower layer. Even if the final configuration is known, incrementally changing current configuration to final without violating security and functionality invariants is a hard AI planning problem. Convincing administrators that revealing anonymized versions of their configurations will not invite new attacks is an open problem. Network administrators are often not computer scientists so getting them to adopt formalized languages is not straightforward.

This issue brings together nine papers that address some of the above problems in the areas of specification, diagnosis, repair, synthesis and anonymization. We received 37 papers in total.

The first paper by Urushidani, Abe, Ji, Fukuda, Koibuchi, Nakamura, Yamada, Shimizu, Hayashi, Inoue, Shiomoto "Design of Versatile Academic Infrastructure for Multilayer Network Services" provides a concrete description of configuration challenges on a large scale. It describes the design and deployment of SINET3, the new Japanese academic infrastructure providing multi-layer network services to more than 700 universities and research institutions.

The second paper (invited) by Bellovin and Bush "Configuration Management and Security" focuses on the

Digital Object Identifier 10.1109/JSAC.2009.090401

challenges of configuring security in a large, heterogeneous environment that includes firewalls, servers, desktops and PDAs. This challenge is compounded by the interaction between security and business policy and by continuous attempts by "enemies" to subvert configurations where enemies could include benign insiders frustrated by security policies.

The third paper by Pappas, Wessels, Massey, Terzis, Lu and Zhang "Impact of Configuration Errors on DNS Robustness" quantitatively documents the impact of configuration errors on reliability of the critical DNS infrastructure. This infrastructure is highly resilient to server failure but only under the assumption that servers fail independently of each other. The paper shows how configuration errors violate this assumption and thereby significantly compromise this infrastructure's resilience.

The fourth paper by Lee, Wong and Kim "NetPiler: Detection of Ineffective Router Configurations" presents a method of identifying BGP routing policies that are "ineffective" in that their removal will not change network behavior. These could either be removed to simplify policies or be repaired to restore network administrator's intent. This method was evaluated with configurations from large ISPs and a university network and identified roughly a hundred misconfigurations.

The fifth paper by Al-Shaer, El-Atawy and Samak "Automated Pseudo-Live Testing of Firewall Configuration Enforcement" addresses the problem of checking whether firewall policies have been correctly implemented. Starting from the firewall policy, the paper describes an algorithm to generate a set of test cases that test for common errors arising in firewall implementation, e.g., in rule ordering. It has been evaluated on firewalls with about 25,000 rules.

The sixth paper by Homer and Ou "SAT-Solving Approaches to Context-Aware Enterprise Network Security Management" describes not only how to diagnose configuration errors but also how to optimally repair these. It exploits properties of minimum cost SAT solvers and of proofs in the logic-based language called Datalog.

The seventh paper by Enck, Moyer, McDaniel, Sen, Sebos, Spoerel, Greenberg, Sung, Rao, Aiello "Configuration Management at Massive Scale: System Design and Experience" addresses the problem of requirement specification and automated configuration synthesis. It defines active templates called configlets whose output can be spliced into a native configuration file thereby reducing the need for administrators to learn a fundamentally new specification language. It has evolved over five years of use in a large enterprise network.

The eighth paper by Wang, Avramopoulos and Rexford "Design for Configurability" addresses the configuration synthesis problem. It presents a system for fine-grained control over interdomain routing policies that is not possible with the current BGP policy language. Specifically, it allows one to express tradeoffs between different BGP policy objectives. The tradeoff is expressed not by a step-by-step ranking of BGP attributes but more naturally by weighting competing objectives within the Analytic Hierarchy Process framework.

The ninth paper by Maltz, Zhan, Hjalmtysson, Greenberg, Rexford, Xie and Zhang "Structure Preserving Anonymization of Router Configuration Data" is based on the observation that enterprises tend to be less secretive about their network's structural information than they are of their identity. Fortunately, the structural information is also of greatest interest to networking researchers. The paper proposes an algorithm to preserve structural information while anonymizing identity information.

We are grateful to Martha Streenstrup, the issue mentor, Laurel Greenidge and Sue Lange of IEEE and the panel of 61 reviewers for their time, expertise and counsel in the preparation of this issue.

Paul Anderson, *Guest Editor*
Researcher
School of Informatics
Edinburgh University

Carl A. Gunter, *Guest Editor*
Professor
Computer Science Department
University of Illinois, Urbana-Champaign

Charles R. Kalmanek, *Guest Editor*
Vice President of Networking & Services Research
AT&T Labs

Sanjai Narain, *Guest Editor*
Senior Research Scientist
Information Assurance and Security Department
Telcordia Applied Technology Solutions

Jonathan M. Smith, *Guest Editor*
Olga and Alberico Pompa Professor of
Engineering and Applied Science
University of Pennsylvania

Rajesh Talpade, *Guest Editor*
Chief Scientist and Director
Information Assurance Group
Telcordia Applied Technology Solutions

Geoffrey G. Xie, *Guest Editor*
Professor
Computer Science Department
Naval Post Graduate School

Martha Steenstrup, *J-SAC Board Representative*

Paul Anderson is a researcher in the School of Informatics at Edinburgh University. He has a background in the practical administration and configuration of large computing installations, and he is particularly interested in bridging the gap between theory and practice.

Carl A. Gunter received his BA from the University of Chicago in 1979 and his PhD from the University of Wisconsin at Madison in 1985. He worked as a postdoctoral researcher at Carnegie-Mellon University and the University of Cambridge in England before joining the faculty of the University of Pennsylvania in 1987 and the University of Illinois 2004. He is a professor, director of Illinois Security Lab, member of the Information Trust Institute (ITI) steering committee, and head of the Systems and Networking Area of the Computer Science Department. He is the chair of the steering committee for the ACM Conference on Computer and Communications Security (CCS) and an editor for IEEE Transactions on Computers. Gunter has made research contributions in the semantics of programming languages, formal analysis of networks and security, and privacy. His contributions to the semantics of programming languages include the interpretation of subtypes using implicit coercions, type inference for continuations and prompts, the use of Grothendieck fibrations as a model of parametric polymorphism, the mixed powerdomain, and the use of Petri nets as a model of linear logic. His 1992 textbook and his chapter in the Handbook of Theoretical Computer Science are standard references on the semantics of programming languages. He has also served extensively as research consultant and expert witness on programming languages. Gunter's contributions to the formal analysis of networks and security include the Packet Language for Active Networks (PLAN), the WRSPM reference model for requirements and specifications, the first formal analyses of Internet and ad hoc routing protocols, the Verisim system for analyzing network simulations, and the use of bandwidth as a DoS countermeasure. His work on privacy includes the first research on certificate retrieval for trust management and the formal analysis of regulatory privacy rules. He founded Probaris Technologies, a company in the Philadelphia area that provides credentials for employees of government agencies such as the Social Security Administration and the Patent and Trade Office. His most recent research directions include the security of control systems, including Building Automation Systems (BASs), power substations, and Advanced Meter Infrastructure (AMI). He is also developing the use of attribute-based systems for messaging and security.

Charles R. Kalmanek is Vice President of Networking & Services Research in AT&T Labs. Chuck manages research in algorithms and optimization; networking; optical and wireless systems; speech and natural language understanding; and multimedia services. Chuck's lab also supports organizations throughout AT&T as a Center of Excellence for network design and performance analysis. Chuck joined AT&T Bell Labs in 1980. Chuck has extensive experience in network architecture, protocols and distributed systems. Chuck's research background spans IP network management, access network architectures, wireless networks, voice over IP, multimedia streaming, content distribution networks, storage networks, as well as packet switch and host interface design. Chuck received his undergraduate degree from Cornell University, and M.S. degrees in Electrical Engineering and Computer Science from Columbia University and New York University respectively. Chuck is a recipient of AT&T's Strategic Patent and Strategic Standards Awards. Chuck is a former co-chair of the IEEE Internet Technical Committee.

Sanjai Narain is a Senior Research Scientist in Information Assurance and Security Department at Telcordia Technologies, Piscataway, NJ. His current research is on automated planning of secure and reliable infrastructure. This is based on his experience designing, building, testing and analyzing such infrastructure for large enterprises. To support his research, he has obtained funding from major government agencies such as DARPA, DISA, DHS and IARPA. He has organized and led several university-industry teams with partners such as MIT, Princeton, Cornell, Johns Hopkins and Boeing. He has served on editorial boards and program committees of IEEE, USENIX, ACM journals, conferences or workshops. He joined Telcordia in 1990 when it was called Bellcore. His earlier research at Telcordia was on network management tools for SONET, ATM and DSL networks. From 1981 to 1990 he worked at RAND Corporation where he developed technologies to reason about discrete-event simulation models. He has one issued patent on low-cost DSL loop qualification and three filed patents on configuration validation and synthesis. He has over twenty publications in journals, conferences and workshops. His formal training is in mathematical logic, programming languages, and electrical engineering. He studied logic with Professor Alonzo Church at UCLA. He obtained a Ph.D. in Computer Science from UCLA in 1988, an M.S. in Computer Science from Syracuse University in 1988, and a B.Tech. in Electrical Engineering from Indian Institute of Technology, New Delhi, in 1979.

Jonathan M. Smith is the Olga and Alberico Pompa Professor of Engineering and Applied Science at the University of Pennsylvania. He recently returned to Penn after serving as a program manager at DARPA/IPTO, where he initiated research programs in cognitive networking and distributed radio. His current research interests are in terabit-per-second networks and wireless network security. He is a Fellow of the IEEE. Contact him at jms@cis.upenn.edu.

Rajesh Talpade is Chief Scientist and Director of the Information Assurance Group at Telcordia Applied Technology Solutions, with over 15 years experience in Internet, telecom, wireless, and security areas. He currently has responsibility for a new Telcordia software product in IP network management, and has led all stages from concept to market. He has been the Principal Investigator for several Government-funded R&D projects, such as cyber-attack traceback, IP device configuration error detection, and Distributed Denial of Service attack detection. Rajesh has a Ph.D. in Computer Science from the Georgia Institute of Technology, and an MBA from Columbia University. He holds and has several patents pending, and has published numerous refereed papers and IETF RFC 2149.

Geoffrey G. Xie is a Professor and Associate Chair in the Computer Science department at U.S. Naval Postgraduate School. He received his B.S. degree in computer science from Fudan University, China, his M.S. in computer science and his M.A. in mathematics from Bowling Green State University, Ohio, and his Ph.D. in computer sciences from the University of Texas, Austin. His current research interests include clean slate design of IP control plane, static analysis of network configuration, routing glue logic, underwater acoustic networks, and abstraction-driven design and analysis of enterprise networks.