Computer Networks 53 (2009) 759–760

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Guest Editorial

Traffic classification and its applications to modern networks

The observation of network traffic is at the core of many fundamental network operation and maintenance activities, such as the enforcement of quality of service guarantees or traffic engineering, and of security issues such as intrusion detection and prevention. However, the rapid introduction of new categories of applications such as network games and peer-to-peer communication, the increasing presence of malicious traffic, and the widespread use of encryption techniques, make the measurement, analysis and classification of Internet traffic a challenging task. The research community has devoted a great effort to the study of new traffic characterization and classification mechanisms, with the specific intent of surpassing the limitations that traditional approaches based on port and payload analysis face.

The purpose of this special issue is to collect high-quality contributions in this fertile area of research, with the specific intent to present novel classification and characterization techniques, and to discuss new applications of such techniques as well as their current limitations.

More than 60 papers were submitted to this special issue, witnessing the importance of this area of research in today's academic and industrial environment. With the help of more than 120 dedicated reviewers, we selected eight top-quality articles. A first subset deals with novel ***techniques for traffic classification and analyzes their accuracy***. More precisely, "*A non-linear, recurrence-based approach to traffic classification*" presents a new, recurrent-plot based mechanisms for packet-level traffic classification; "*Characterizing network traffic by means of the NetMine framework*" describes a novel methodology to leverage advanced data-mining techniques to classify traffic at the packet level; "*Efficient Application Identification and the Temporal and Spatial Stability of Classification Schem*" analyzes the stability of statistical classification results considering different periods of time. A second set of articles deals with ***optimization mechanisms*** that apply to traffic classification, such as those obtained by considering new hardware architectures "*Architecture and Scalability of a High-Speed Traffic Measurement Platform with a Highly Flexible Packet Classification*" or efficient look-up algorithms for large databases of filtering rules "*Scalable Packet Classification Using Controlled Crossproducting*". Finally, three interesting ***applications of statistical techniques for traffic classification*** are being published here: "*Targeting Spam Control on Middleboxes: Spam Detection Based on Layer-3 E-mail Content Classification*" presents a novel way to detect unsolicited e-mails (yes, we mean spam!) exploiting statistical analysis of e-mail traffic performed at the packet level; "*Profiling and Identification of P2 traffic*" explicitly targets the detection of peer-to-peer traffic, while applications of behavioral traffic classification to intrusion detection systems are discussed in "*McPAD: A Multiple Classifier System for Accurate Payload-based Anomaly Detection*".

We believe that the three sets of accepted papers share several common features, besides being based on novel ideas. Firstly, while they show that this area of research has produced mature results, they also demonstrate that there is still significant room for further studies and improvements.

Secondly, they represent excellent examples of practical applications of traffic classification techniques, particularly the ones based on statistical and behavioral traffic analysis. Indeed, besides the traditional application of these mechanisms to QoS support and traffic management, traffic classification is shown to be effective in spam detection, peer-to-peer traffic identification and security enforcement.

Thirdly, and most importantly, all papers show that the most severe hurdle that slows down research on traffic classification is the lack of standard testing procedures and benchmarking metrics. Most of the experiments described in the papers were carried out on different traffic traces, collected from very different networks and scenarios, which makes it difficult to compare the different proposals. In general, the unavailability of public and accurate measurement data taken from real operative networks represents a showstopper to the progress of research in the areas of traffic analysis and classification. During the reviewing process, the most common question reviewers asked was to extend the provided results considering different scenarios, which underlines the lack of benchmarking dataset and of publicly available traffic traces that would offer a common reference set.

1389-1286/$ - see front matter
doi:10.1016/j.comnet.2008.12.007

There are three factors that have contributed to the lack of publicly available, verified reference data accessible for research activities in traffic measurement and analysis. One is certainly related to the need of protecting user privacy, mandated by ethics and laws. Traffic measurement and classification activities, especially when requiring the inspection of packet payload, traditionally require access to content which is very sensitive with respect to privacy – and possibly to the security – of both end users and network operators. The traditional defense to this problem has been simply to avoid disclosing, or even gathering, most forms of measurement data. A second issue relates to the availability of measurement systems and techniques that can work in an on-line fashion, and that can be easily installed and maintained with ideally no impact on running networks. Until this issue is solved, network operators will have tangible and more than reasonable grounds to discourage the installation of monitoring equipment devoted to research activities. Finally, the research community needs to show how its efforts are not only necessary for the continuing evolution of networking technologies, but that network operators and users can benefit from the availability of advanced classification techniques.

A possible (the only?) solution to these problems is to focus future research efforts towards technologies that move the measurement and classification activities as close "to the bits on the wire" as possible. The future of traffic measurement and analysis is necessarily linked to our ability to devise methodologies that directly analyze bits as they are observed on communication channels, as opposed to capturing and storing them for analysis at a later stage. We think that this is the major challenge that the measurement community will need to tackle in the next few years.

Marco Mellia received his Ph.D. in Telecommunications Engineering in 2001 from Politecnico di Torino. From March to October 1999 he was with the CS department at Carnegie Mellon University as visiting scholar. From February to March 2002 he visited the Sprint Advanced Technology Laboratories Burlingame, California, working on the IP Monitoring Project (IPMON). Since April 2001, he is with Electronics Department of Politecnico di Torino as Assistant Professor. He has co-authored over 100 papers published in international journals and presented in leading international conferences, all of them in the area of telecommunication networks. He participated in the program committees of several conferences including IEEE Infocom, IEEE Globecom, IEEE ICC and ACM SIGCOMM. His research interests are in the fields of traffic measurement, classification and modeling, and of Peer-to-Peer systems.

Antonio Pescapè is an Assistant Professor at the Department of Computer Engineering and Systems of the University of Napoli Federico II. He received the M.S. Laurea Degree in Computer Engineering and the Ph.D. in Computer Engineering and Systems both at University of Napoli Federico II. His research interests are in the networking field with focus on models and algorithms for Internet Traffic, Network Measurement and Management of heterogeneous IP networks, and Network Security. Antonio Pescapè has co-authored a large number of journal and conference publications. He is an IEEE member and he has served and serves on several conference technical program committees including IEEE Globecom, IEEE ICC, IEEE WCNC, IEEE HPSR and he serves as Editorial Board Member of IEEE Survey and Tutorials.

Luca Salgarelli is an Associate Professor of Telecommunications at the University of Brescia, Italy. Prior to joining the University of Brescia, from 1998 to 2005 he was with the Networking Lab of Bell Labs Research (Lucent Technologies) in Holmdel, NJ – USA. From 1995 to 1998 he was a Researcher with the Networking Department of CEFRIEL/Politecnico di Milano, Italy. Until 1998 he was active in the areas of broadband IP networks and QoS provisioning. Since 1999 his research activities have covered design, development and evaluation of systems and protocols for data networks, in particular when they involve mobility and security. His writings have appeared in numerous ACM and IEEE publications, as well as other professional conferences and journals. He is currently serving as Associate Editor of IEEE Transactions on Mobile Computing and as Area Editor of Elsevier's Computer Networks, has served in the committees of several IEEE and ACM conferences, and has contributed to several IETF working groups since 1996. He holds six US and international patents and is co-author of several patents pending.

Marco Mellia
Antonio Pescapè
Luca Salgarelli
E-mail addresses: marco.mellia@polito.it (M. Mellia),
antonio.pescape@unina.it (A. Pescapé),
luca.salgarelli@ing.unibs.it (L. Salgarelli)

Available online 24 December 2008