Research Statement

Ioannis Papapanagiotou

My research objective is to contribute to the exciting areas of computer networks and services. To this end, my research work falls into the following categories:

1. Data analytics & network compression systems.
2. Functionality distribution in aggregation networks.
3. Performance analysis of wireless networks.

I work ambitiously to bring interesting ideas to mainstream practice. My research is based on a three step approach: (a) identify a unique engineering problem in the computer networking and services domain; (b) develop the appropriate theoretical framework; (c) evaluate through simulation, emulation or prototyping. Heretofore, my ambition is for my ideas to flourish into useful entities for the community. In what follows, I cover my research background and accomplishments. In the last section, I describe my academic goals and future interests.

Data-driven Analysis & Design of Deduplication Systems

As more and more users access the Web through mobile devices, their traffic patterns need to be better understood. This is important both to wireless network carriers and to enterprise network administrators, as they provision their networks for the expected growth of smartphones and tablets. Yet many aspects of smartphone traffic behavior have not been fully apprehended. This may have implications for network design and provisioning, particularly if smartphone traffic is significantly different from other devices. In my work, I have tried to expose some of the differences by comparing the network access behavior of smartphones with a control group, namely laptops.

Device Identification: First, I have devised and patented an Operating System (OS) fingerprinting methodology that identifies each device type such as iOS, Android, BlackBerry, Windows, Mac OS X and Linux. More specifically, this methodology uses association rule mining to determine a potential correlation and trustworthiness of several protocol headers. I prototyped and deployed this idea in two wireless campuses, namely NC State University and IBM TJ Watson Research Center. The former included mostly student smartphone devices, and the latter corporate laptops. My technique was able to identify 98% of the devices in both networks [5].

Differential DHCP lease time: I investigated the impact that new types of wireless devices, such as smartphones, have on DHCP. I used two one-month long traces, collected at the gateway of NC State University and IBM TJ Watson Research Center, and compared side-by-side their DHCP usage patterns. I showed that DHCP implementations vary among device types, and have an effect on DHCP lease durations. To improve network address utilization, without introducing any protocol changes, I proposed a new leasing strategy which takes into account the device types. This strategy, compared to current approaches, provides a sixfold improvement in the address utilization without considerably increasing the DHCP overhead [5].

Web Traffic: Another important aspect is the Web traffic that the mobile devices are introducing to the network and the corresponding implications on caching designs. I used standard metrics such as Web requests and responses, popularity, content type, and multimedia downloads. I focused mainly on those areas where the traffic of the two device categories differ. I looked into the difference on the video delivery in mobile devices, and examined the effectiveness of caching in detail, both at the browser and at the proxy level. Using a simulated environment fed by the collected data, I showed that smartphone browser caches are not as effective as laptop caches. Surprisingly, this was not due to the size of the device cache, as a 10MB smartphone browser cache was enough to provide most of the achievable savings. I looked into techniques that could take advantage of the unique smartphone traffic properties, such as caching objects based on range requests. In addition, I showed that both proxy and browser caches must handle partially downloaded objects; if they do so, they could significantly improve the bandwidth consumption and the end to end delay compared to the default Squid proxy policies. [6].

Hybrid Byte Level Caching: I also looked into techniques that can improve even further the bandwidth savings, by identifying duplicate content at a smaller granularity. For simplicity and consistency, in this statement, I refer to them as *byte caching* to stress the granularity on which redundancy is determined (sequence of bytes). I used wireless network traces and showed that byte caches can provide more savings compared to Web object caches. Then I analyzed the applications that provide most of the byte level savings, and discovered that some applications are more prone to byte level deduplication than others. Hence, I designed and patented an apparatus that uses an intelligent scheduler to decide the optimal deduplication strategy. The proposed implementation showed more than twice as much savings compared to standalone proxy caches, and higher savings compared to byte caches. It also required a third of the memory to store the hashes, and provided a system speedup equal to two [1]. I have also identified design directions that provide even more savings, and I am currently pursuing them through several patents in IBM.

Functionality Distribution in Aggregation Networks

The second area of my research activity is on network design & dimensioning problems [2, 3]. This work has been motivated by the lack of established principles on the design of next generation aggregation architectures. Specifically, most of the design problems address the issue of deploying new infrastructure. However, the cost to build out new infrastructure may be prohibitively expensive. In addition, vendors have introduced multipurpose edge "systems" (backplanes) that incorporate different functionalities in modular "sub-systems" (line cards). Finally, Internet Service Providers (ISPs) have to deal with the explosion of video traffic in different formats, e.g. multicast, peer-to-peer, mobile etc.

Hence, it is of vital importance for the ISPs to re-engineer their infrastructure to meet these challenges. Moreover, Network Design Problems (NDPs) have to be revisited to include those multipurpose edge systems, instead of single-functionality devices (such as L2 switches or L3 routers). Considering the above challenges, I set out to:

- Define the possible aggregation architectures: 1. *Centralized Single-Edge*, 2. *Centralized Multi-Edge*, and 3. *Distributed Single-Edge*. We further differentiate based on whether the edge system is clustered or maintains its functionalities in a single box.
- Develop a network design model that goes beyond the well investigated *location* and *dimensioning* problems. Our modeling approach is applied to both designing an aggregation architecture, as well as upgrading the current infrastructure.
- Propose models based on edge "systems", rather than network elements. Edge systems may support different types of intelligence, either as part of the backplane or as part of the modular "sub-systems", i.e. high-end line cards.

Finally, I have evaluated the model with two close-to-real-life scenarios and multiple traffic profiles, based on data provided by EU ISPs. My results yield that ISPs will benefit from distributing the functionalities closer to the subscriber. Furthermore, sensitivity analysis suggests that ISPs should invest on systems that support more line cards and interfaces, rather than on devices with higher bandwidth.

Performance Analysis of Wireless Networks

The final area of my work deals with the analytical decomposition of wireless protocols. Specifically, I studied the Quality of Service (QoS) in WiFi networks (IEEE 802.11e). The motivation for this study was twofold; first, some IEEE 802.11e properties were not properly captured by prior analyses, and second, the current standards do not define principles for strict QoS [10, 15].

I used three common analytical approaches for wireless networks, namely Discrete Time Markov Chains, Closed form Queueing networks, and analysis using elementary conditional probabilities. I proposed several enhancements that included (a) QoS class differentiation, (b) Block-Acknowledgments, as defined in the IEEE 802.11e standard, and (c) erroneous channel conditions. I demonstrated that Block-ACKs may improve the QoS under erroneous channel conditions [7, 8]. I also proposed a methodology to incorporate non-markovian traffic to the analytical approaches. In addition, my colleagues and I devised a scheme that combines TDMA for the VoIP class, and CSMA/CA for the rest of the QoS classes [13]. Our analytical results indicated that the saturation throughput was higher and the end-to-end delay was lower. Moreover, we showed that with the proper Admission Control, the aforementioned scheme may provide QoS guarantees [14]. We verified all these results through Discrete Event Simulations. We also extended the simulation analysis based on a Response Surface Methodology (RSM) in order to determine the effect of parameters from multiple layers of the network architecture in the performance of the wireless client [4]. During my collaboration with the VTT Research Institute, I also applied the same analytical approaches on the energy consumption of WiMAX networks (IEEE 802.16e) [12, 11, 9].

Summary & Future Work

As you may notice from the above, my academic background combines principles from computer engineering and operations research, i.e. queueing theory, simulation, trace-driven analytics, data mining, and optimization. I particularly enjoy working on networking systems and designs that may improve the end user's Quality of Experience. I strive so that my work has both a theoretical and a practical contribution, which is why I pursued a dual-degree PhD in Computer Engineering and Operations Research.

My background also includes a 3-year experience with IBM Research and development as part of the in-house incubator team reporting to the IBM WebSphere CTO. In addition, during my studies, I collaborated with Cisco, Samsung Research and the VTT Research Institute of Finland. These experiences have helped me evolve as a researcher and as a member of an organization.

My *future plans* include investigating problems in the areas of mobile computing, data analytics, and distributed networking systems. In particular, I would like to investigate the effect of distributed services in the mobile space. In this domain, I would like to extract knowledge from mobile data, and determine the changes that are needed by the networking operators in order to support the advancements in mobile computing. I am specifically interested in building software systems that have the intelligence to hide the intrinsic complexity from operators and end users.

I also plan to extend my collaboration with the organizations I was affiliated with, and work with graduate students and other faculty members in the department of Computer & Information Technology at Purdue University. I would also like to use my experience in invention disclosure towards building a portfolio of ideas in the above exciting areas of computer engineering.

References

[1] Ioannis Papapanagiotou, Robert D. Callaway, and Michael Devetsikiotis. Chunk and object level deduplication for web. optimization: A hybrid approach. In *International Communications Conference (ICC), 2012 IEEE*, Jun. 2012.

[2] Ioannis Papapanagiotou and Michael Devetsikiotis. Aggregation network design methodologies for triple play services. In *Consumer Communications and Networking Conference (CCNC), 2010 7th IEEE*, pages 1 –5, jan. 2010.

[3] Ioannis Papapanagiotou, Matthias Falkner, and Michael Devetsikiotis. Optimal functionality placement for multiplay service provider architectures. *IEEE Transactions on Network and Service Management*, 9(3), september 2012.

[4] Ioannis Papapanagiotou, Fabrizio Granelli, Dzmitry Kliazovich, and Michael Devetsikiotis. A metamodeling approach for cross-layer optimization: A framework and application to voice over wifi. *Simulation Modelling Practice and Theory*, 19(9):2117–2129, 2011.

[5] Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. Configuring DHCP leases in the smartphone era. In *International Measurement Conference*, IMC '12. ACM, Nov. 2012.

[6] Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. Smartphones vs. laptops: comparing web browsing behavior and the implications for caching. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '12, pages 423–424, New York, NY, USA, 2012. ACM.

[7] Ioannis Papapanagiotou, Georgios S. Paschos, and Michael Devetsikiotis. A comparison performance analysis of QoS WLANs: approaches with enhanced features. *Adv. MultiMedia*, 2007(1):1:1–1:13, April 2007.

[8] Ioannis Papapanagiotou, Georgios S. Paschos, Stavros A. Kotsopoulos, and Michael Devetsikiotis. Extension and comparison of QoS-enabled Wi-Fi models in the presence of errors. In *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, pages 2530 –2535, nov. 2007.

[9] Ioannis Papapanagiotou, Dimitris Toumpakaris, Jungwon Lee, and Michael Devetsikiotis. A survey on next generation mobile wimax networks: objectives, features and technical challenges. *IEEE Communications Surveys Tutorials*, 11(4):3 –18, quarter 2009.

[10] Ioannis Papapanagiotou, John S. Vardakas, Georgios S. Paschos, Michael D. Logothetis, and Stavros A. Kotsopoulos. Performance evaluation of IEEE 802.11e based on ON-OFF traffic model. In *Proceedings of the 3rd international conference on Mobile multimedia communications*, MobiMedia '07, pages 17:1–17:6, ICST, Brussels, Belgium, Belgium, 2007.

[11] Georgios S. Paschos, Ioannis Papapanagiotou, Christos G. Argyropoulos, and Stavros A. Kotsopoulos. A heuristic strategy for ieee 802.16 wimax scheduler for quality of service, 2006.

[12] Georgios S. Paschos, Ioannis Papapanagiotou, and Thomas M. Bonert. *Power Saving Mechanisms in IEEE 802.16e Wireless Networks*. Auerbach Publications, CRC Press, 2009.

[13] Georgios S. Paschos, Ioannis Papapanagiotou, Stavros A. Kotsopoulos, and George K. Karagiannidis. A new mac protocol with pseudo-TDMA behavior for supporting quality of service in 802.11 wireless LANs. *EURASIP J. Wirel. Commun. Netw.*, 2006(3):8:1–8:9, June 2006.

[14] Georgios S. Paschos, Ioannis Papapanagiotou, Efstathios D. Vagenas, and Stavros A. Kotsopoulos. Performance analysis of admission control for 802.11 wlan networks in ad hoc mode. 2006.

[15] John S. Vardakas, Ioannis Papapanagiotou, Michael D. Logothetis, and Stavros A. Kotsopoulos. On the end-to-end delay analysis of the IEEE 802.11 distributed coordination function. In *Internet Monitoring and Protection, 2007. ICIMP 2007. Second International Conference on*, page 16, july 2007.