Final Report

NSF Workshop on Future Directions in Smart Networking and Communication

Atlanta, GA
May 05, 2017

Workshop Chairs
Hongyi Wu, Old Dominion University
Kaushik Chowdhury, Northeastern University

Workshop Website: https://www.lions.odu.edu/~h1wu/NSFWorkshop/nsfworkshop2017.html

Summary

Emerging “smart networks” are enabling a paradigm shift from the existing classical state-of-the-art network designs through autonomous and intelligent decisions spanning systems, protocols, and architectures. Such networks will lay down the foundations of the next generation networked applications as well as allow unprecedented levels of localized human to device and device to device interaction. This timely NSF workshop on “Future Directions in Smart Networking and Communication”, organized during IEEE INFOCOM’2017, attempted to define the characteristics of smart networks looking ahead in a 10-20 year timeframe, identify the key research challenges within them and generate concrete action points for the community, all of which are elaborated in the main report. Over 40 attendees participated in the discussion that was organized under the following four thrusts: 1. Smart network architectures and applications; 2. Smart network analysis, protocols, and optimization; 3. Security & privacy; 4. City-scale smart network testbed platforms.

An overview of the main discussion points and summary of recommendations is as follows:

- A collective effort is required to fully understand what truly constitutes “smart” networking and communication, what are the driving applications and their needs for “smartness” in future networks, and what are the enabling technologies in context of smart networks.
- The interconnections of wireless devices with supporting capabilities of the network, such as edge computing and storage, should become a part of future architecture and protocol design. Ambitious projects that support this tight integration, possibly by re-engineering the network stack by introducing additional layers for these functions, may be needed.
- For high impact research, core fundamental work may also consider economic cost benefit tradeoffs. This approach will encourage industry participation. NSF can play a major role in creating and facilitating such joint calls, by also defining different near- and far-term performance metrics in the review process that may excite industry partners. PAWR was singled out a significant effort in the right direction for NSF.
- Machine learning is a promising approach and being increasingly used to create smart networks, which may help in both network configuration and security. Though this is a powerful tool it needs careful mapping of the technique to the problem to be solved. The operational time overheads must match that for training and classification.
- Special care must be paid to data access and aggregation. How to ensure utility while protecting privacy is a key challenge. Moreover, different devices must have different access settings, and managing this wide range of configuration options autonomously, while hard, will ultimately increase adoption for the human user.
- Testbeds at all scale must be supported, including small and mid-sized efforts. NSF may provide support for federating them, increasing user accessibility, and broadly encouraging their adoption through new funding modes, while the community catches up with creating new venues specifically designed to showcase research on these testbeds.

The workshop discussions recognized the importance of encouraging large-scale, repeatable and practicable research that invites participation from multiple stakeholders from government and industry, and directly impacts social benefit. Additionally, there was consensus on the need to calibrate recommendations over the next years as well as holistically treat the four main workshop thrusts as equally important in future funding opportunities.

1. Introduction

In recent years we have witnessed a remarkable proliferation of networked devices, collectively enabling a digitization of the physical world. These devices and the architectures they enable are progressively becoming intelligent, and capable of demonstrating autonomous behavior. These emerging class of "smart networks" are able to collect vast amount of information by embedding physical devices with electronics, sensors, actuators, computation resources, intelligent decision modules, and network connectivity that facilitate data gathering and exchange without human intervention. The resulting autonomous and intelligent decisions are spread across a wide spectrum of possibilities: they can be made remotely across a single or multiple existing network infrastructures through centralized intelligence or at the location itself through collective awareness and group action of the deployed devices. Apart from solving the problem of communication, such smart networks present vast opportunities for organizations to utilize data in real time to improve efficiencies, gain competitive advantage by providing better service to end-users, and build new business models that have not yet been envisaged today.

While the transformation of the classical network architecture into such a smart network paradigm is highly desirable, there are many fundamental challenges that remain. Conventional packet-switched networks are designed to keep the network infrastructure and low-layer protocols simple yet efficient, serving as a resilient and scalable communication infrastructure for packet delivery. By following the end-to-end design principle, sophisticated functions and intelligence are pushed to the edge and applications. Thus efforts to embed intelligence within the network core and to the wireless edge are often met with significant complexity and logistical challenges. These challenges may include network design and ambitious optimization goals that must be met with a myopic local view of the environment. As the community continues to invest in deployable wireless systems and devices that are capable of intelligent action, there is also the inherent danger of new attack vectors that can compromise both the device hardware and the protocol that it operates with. Sophisticated network architectures may have unknown safety loopholes that can be exploited through the network, and immediately, the powerful data gathering devices become a critical source of privacy loss and intrusion.

1.1 Workshop Objectives

The purpose of the workshop is to collect community inputs and provide feedback to the National Science Foundation (NSF) on several major research areas related to **smart networking and communication** looking ahead in a 5-10 year timeframe. In particular it aims to discuss a range of fundamental questions including: What new applications are driving the transformation of smart networks? How to design new network architectures for smart networking and communication? How to effectively optimize smart network algorithms and protocols? How to address security and privacy concerns due to new vulnerabilities in smart networks? To this end, the workshop includes the following discussion thrusts:

1. Smart network architectures and applications

2. Smart network analysis, protocols, and optimization
3. Security & privacy
4. City-scale smart network testbed platforms

At the dawn of the transformation into a smart network paradigm, it is critical for the research community to gain deep understanding of the key challenges and opportunities for conducting fundamental and inter-disciplinary research in this area. This workshop serves as a timely platform for the community and the National Science Foundation to develop and articulate a grand vision for future development of smart networking and communication systems.

1.2 Workshop Structure

The NSF Workshop on Future Directions in Smart Networking and Communication was held on May 5th, 2017, co-located with the IEEE INFOCOM conference in Atlanta, GA, to attract broad participation from the wireless communication, networks and systems community. It was co-chaired by Hongyi Wu (Old Dominion University) and Kaushik Chowdhury (Northeastern University), with the support of steering committee consisting of Suman Banerjee, (University of Wisconsin), Marco Gruteser (Rutgers University), Tom Hou (Virginia Tech), Wenjing Lou (NSF & Virginia Tech), and Thyaga Nandagopal (NSF).

A total of 40 experts in related fields were invited to attend this one-day event, aiming to discuss and understand new challenges, identify open problems, and chart a grand vision for fundamental and inter-disciplinary research in smart networking and communication. The participants were informed of the four discussion thrusts and requested to respond to a Doodle poll to indicate their top two choices prior to the workshop. These choices were considered as preferences and not commitments for joining the corresponding discussion groups. The poll was helpful to estimate the overall interest in each topic and ensure balanced participation of the breakout discussion sessions.

The workshop includes two panels and four breakout discussion sessions. Panel A was led by 4 panelists (Henning Schulzrinne, Eytan Modiano, Rajarathnam Chandramouli, and Samir Das) to discuss thrusts 1 & 2, while Panel B (presented by Marco Gruteser, Yingying Chen, Bhaskar Krishnamachari, and Prasun Sinha) covered topics related to thrusts 3 & 4. In each panel, the panelists first presented their positional statements to talk about their visions and propose an ambitious 5-10 year plan for the research community. Following this initial presentation session, there was an open question-answer session, leading to further discussions that were addressed fully in the later breakout sessions. The panel presentations were insightful in a sense that they encouraged different viewpoints. Each panel was followed by two parallel breakout sessions. Each breakout session was organized by a group lead and a scribe. The detailed workshop agenda is included in Appendix.

This report incorporates the summaries from the breakout sessions as well as inputs from individual attendees and various discussions throughout the workshop.

2. Thrust Discussions

This section summarizes the discussions, debates, questions, and recommendations from the two panels and four breakout sessions.

2.1 Smart Network Architectures and Applications
(Lead: Henning Schulzrinne; Scribe: Saswati Sarkar)

Goal: This thrust aimed to stimulate discussions on new research challenges in emerging smart network architectures and applications. Specifically, it invited responses on the following questions:

- What are the emerging applications that drive the development of smart networking and communication? Examples may include smart radios, Internet of Things (IoT), autonomous vehicular networks, drone networks, mobile sensing, radio-based sensing, mobile social networks, user and object localization and tracking, wearables and implants communication, and immersive virtual reality on mobile devices. Why do they require embedded network intelligence?
- What are the new enabling technologies in context of "smart networks"?
- What are the new fundamental challenges in network architecture design that the community must address in priority?
- How can we harness the knowledge from interdisciplinary domains and yet not lose sight of the core networking challenges?

Summary of discussions: The Internet emerged as the "Rebellion of the dumb" in a sense that the telephone network had an intelligent core to start with, but the Internet was a *dumb* network right from inception, and one that relegated complexity at the edges. Over the last two decades, networking researchers have engaged in designing smart network algorithms, protocols, and architecture for the Internet. This "smartness" mostly emerged as mathematically grounded resource optimization algorithms that could operate diverse inputs and provide configuration outputs.

As we take stock of what has changed in the last few years beyond the above approach of optimization-driven problems, we need to identify the priorities for the next two decades. In the context of the workshop topic, we need to understand what truly constitutes smart networking: Are these simply new constraints from emerging applications that can be fed to the mathematical works on optimization that exist already, or must we design new tools, or a combination thereof? Furthermore, are there functions that render the network smart in terms of architecture, fully divorced from the applications?

We use the example of Internet-of-Things (IoT) to address some of the above questions. IoT promises to usher in a paradigm shift, in the sense that intelligence now needs to be spread out in the networks, and the network needs to respond with a very low latency. Smart phones can facilitate the dispersion of network intelligence, as Apps can do computation and storage, can

operate in poor network, and provide a generic edge-computing functionality not seen before. Smart phones and the Internet-of-things have motivated several societally important applications that require computation, storage and data processing and access within the network. Some examples of such applications would be automotive applications, Big Data and Virtual Reality. Thus, today's wireless devices are capable of computing at unprecedented scales at any stage of the network architecture. This computing capability is a new tool, which must be fully leveraged in future optimization problems in network design and response to events.

We believe that applications will play a big role in future smart networks. They will request fundamentally different needs from the network and require optimization of a large number of critical tradeoffs. There are the following questions that merit collective thought: Are there emerging applications that are not well-served by today's network? What fundamental changes would be required in the current networks to serve such applications? For example can the current networks support relaying of data generated by scientific research, given the volume in question? Can the current networks support the specific needs of automotive applications and virtual reality? These applications might need to access storage and computation frequently. Can storage and computation be made closer to the network points that need them? More specifically, there are great advantages if the wireless network is integrated with storage and computation, and its typical operation considers also the overhead and benefits of accessing them. This brings us to systems design question on where must these peripheral support services be located with respect to the radio front-end. From the protocol viewpoint it is also an open question whether these should be introduced as part of the core network functionality or need we define a new network stack layer for performing the storage/computation tasks.

The smart devices that exist today have enabled Internet access over both cellular and WiFi networks. One key problem that the next generation networks need to determine is when to switch between cellular and WiFi and when to use them together.

Given the sheer number of devices involved in next generation smart networks, and given that most of these would be located as end-users, operation and management need to be made easy in smart networks. The earlier model of having a separate and central network manager will no longer apply owing to scalability issues. There are many open research questions in operations and management, which also need to consider real-world implementation issues by service providers. One possible approach is categorizing networks based on their ease of management. For example, level 1 networks have a central manager, while management would be somewhat more dispersed in level 2 and so on.

The next generation networks need to integrate economic viability organically as part of their design, rather than an add-on after-thought. Thus, all the above questions need to be visited keeping in mind economic tradeoffs, economics for different entities, providers and users.

Machine learning has provided a new tool to designers of next generation smart networks. It can help identify the patterns that can be extracted to enhance performance and security, e.g.,

in preventing misuse of IoT devices for attacking other devices. It can help optimize resource utilization and user experience. There is a distinct possibility that while the networks so far have relied on rigorous optimization, future networks will wholly rely on machine learning to configure themselves.

Action Plan: Based on the panel presentation and breakout discussions, the following actions are recommended by the discussion group.

- Both the NSF and research community must make a collective effort to fully understand what truly constitutes "smart" networking and communication.
- Active research is needed on whether machine learning techniques can be exploited to support "smartness" in future networks.
- As applications are expected to drive the development of smart networks, research should be conducted to explore if there are emerging applications that are not well-served by today's networks. This will guide the fundamental changes that are required in the current networks to serve such applications.
- Future research must be encouraged to organically integrate economic viability into smart networks as part of their designs.
- Given that a sheer number of devices would be involved in future smart networks, more that attention should be given to the real-world implementation. Many challenges are not immediately visible unless there is some scale of testing.

2.2 Smart Network Analysis, Protocols, and Optimization
(Lead: Rajarathnam Chandramouli; Scribe: Jia (Kevin) Liu and Bin Li)

Goal: The massive amount of runtime data generated during smart network operations creates new opportunities to better understand the networks, streamline their designs, create upper layer end-to-end networking protocols and achieve optimized network performance. This thrust posed the following questions:

- How can the community effectively obtain and share real-world data sets to guide future design strategies?
- How can future network design and optimization leverage data analytic techniques, in particular, big data and machine learning technologies, in future smart networks.
- How will upper layer protocol design, such as routing and transport layer protocols, evolve from their well-tested classical wireless counterparts by absorbing the vast data and new tools that are emerging today? Are cross-layer protocol designs a viable path forward?
- Are current protocols and analytical techniques enabled to handle complex interactions between emerging wireless devices and networks, and also between the control and data planes?
- How can future network protocol design and optimization formulations better address actual, real-world problems.

Summary of discussions: While the classical network protocol stack has served well so far, a number of recent research advances in networking protocols, e.g., multipath TCP (MPTCP), routing protocols that consider the underlying technology architectures such as heterogeneous networking (HetNet) and, cross-layer approaches that support application-driven quality of experience (QoE) are promising next steps to enable IoT, connected transportation, and other forms of smart networks.

HetNets play a major role in 5G networks. These inherently include multi-radio access terminal (multi-RAT) devices. Therefore, smart access, management and control of the multiple wireless interfaces is critical. While multipath TCP is showing early promise in HetNets and also being slowly adopted by the industry, its future is still unclear. MPTCP related security issues are not yet well understood. Are other layers in the network protocol stack better suited for HetNets, or must MTCP transcend the boundaries of the protocol stack for more inputs that can guide its path (and hence, technology) selection process? More research might answer this critical question.

New physical layer (PHY) and medium access control (MAC) design for wireless devices that allow seamless connections, enable sensing and channel access simultaneously over multiple bands is an open research problem. Given that newer bands are opening up for unlicensed use, we need to develop scalable protocol solutions for fast and wideband channel utilization. Full duplex communications, battery power optimization, and dynamic spectrum access/management are additional and related broad research areas.

Wireless backhaul and fronthaul designs, mobile edge optimization, software defined wireless controllers are also emerging technology research areas. Academic research typically tends to ignore policy related issues including pricing, spectrum policies, mission critical policies of public safety communication networks, etc. Inter-disciplinary collaboration between networking researchers and policy/economics researchers will result in interesting solutions to these problems.

While there is visible trend towards interdisciplinary thinking in wireless networking research, such emphasis must not lost sight of the core disciplinary research problems. Band-aid solutions that attempt to artificially stitch together wireless research with a different discipline may not give high impact. For example, machine learning is a hot research area these days. It may provide novel insights in the design and analysis of cognitive radio networking. But, care must be taken to carefully choose the right approach and justify the choices available that include deep learning, simulated annealing, heuristics, etc. Additionally, cost benefit tradeoffs of machine learning must be carefully thought, especially where such techniques can be implemented. For example, networking controller and the user equipment have access to different scales of data and computational power. Moreover, there are fundamental time-scale differences at the various layers of the protocol stack. If this time-scale of receiving fresh data updates and performing meaningful actuation tasks is not considered carefully, machine learning approaches to solve wireless networking problems are likely going to fail. Therefore, interdisciplinary thinking should be problem driven.

Fundamental problems typically address 10 to 20 year window challenges. A contemporary application-driven problem may not produce fundamental results. Therefore, it is important to build connections between different research communities to arrive at a consensus on the long term, pressing research problems in wireless networking. This will lead to grand challenges that bring together large academic-industry consortia. It is anticipated that the next wave of fundamental research problems will be in digital healthcare, virtual reality, smart manufacturing and connected transportation. All of these require active outreach from the highest levels to build great partnerships.

There is great economic opportunity for major industry players as they progress towards smart networking paradigms. This excitement can potentially change the state of industry support for basic research, which is largely unavailable these days. The community needs to make a stronger case of the shared gains through fundamental wireless networking research, and additional joint proposal calls is a way forward. As the time schedule for industry oriented research is highly compressed, new programs can be designed where the NSF and the industry decide mutually upon a mix of near- and far-term impact research as its merit review criteria. Recent NSF initiatives (e.g., NSF PAWR) that involve a consortium of industry participants (e.g., NSF PAWR) is a welcome addition. Program such as the IUCRC, SBIR, STTR, and SAVI must be strengthened to further encourage industry-academic collaborations both within the US and international cooperation.

Finally, the community needs to identify broader impacts beyond the traditional metrics such as publications in reputed journals or conferences. Does the research result in a tangible product (e.g., a working prototype system)? Who are the end user communities? How does the research improve the day-to-day life of ordinary citizens? Rural broadband connectivity, public safety communications, and telemedicine are a few examples where novel wireless networking solutions can save lives. There needs to be a greater reward and recognition for truly impactful research.

Action Plan: The discussion group proposes the following plan:

- Research on multipath TCP for heterogeneous, dense networks needs more attention, with possibly better integration with security and cross-layer designs.
- Interdisciplinary research that combines expertise of core wireless/wired networking with policy and economics should be encouraged. This will address many future needs of smart networks from a practical viewpoint.
- Future work on machine learning to guide networking decisions, without tight integration, will result in sub-optimal or unreliable outcomes. Instead, the impact on protocol operational time scales, changing environments, training needs must be factored in.
- More support is needed in the domains of digital healthcare, virtual reality, smart manufacturing and connected transportation, as they will majorly benefit from future smart networks.

- NSF may consider revising its evaluation criteria to (i) encourage industry participation and (ii) consider impacts on translation research that will end up directly or indirectly benefitting the community.

2.3 Security and Privacy
(Lead: Marco Gruteser; Scribe: Ming Li)

Goal: The characteristics, performance and security requirements of smart networks vary considerably from one system to another. The endless variety of applications poses an equally wide variety of security and privacy challenges. This thrust posed the following questions:

- What are the fundamentally new wireless security and privacy problems in the emerging smart networks?
- How is network security impacted by the dynamic and diverse network connectivity
- Typical characteristics of emerging networks may include weak device protection, extremely limited computing power (for small IoT), storage space, and energy supply. What techniques may be used to mitigate growing attack vectors under these constraints? Alternately, how do we combine limited local vs. powerful cloud computation to enhance security.
- How do we ensure privacy preservation and data utility at the same time, both in research environments and real world applications?

Summary of discussions: The emergence of IoT gives rise to many exciting smart applications, such as smart health, smart transportation, smart cities, and smart homes, where the key features are the capability of cognition, self-adaptation and self-configuration. For example, in smart health, one can wear smartwatch and various fitness devices that keep monitoring vital signs and health status, so that daily health data can be collected and analyzed for disease prevention and diagnosis; in smart transportation, smartphones and other sensors are deployed in cars to collect state information (such as location, speed, acceleration, or engine status), which enables vehicle fault diagnosis, safety enhancements, as well as global route/traffic optimization. In addition, through understanding the driver's intent, the car sensor/actuators' parameters can be dynamically tuned to better satisfy his/her needs. In smart homes, RFID tags can monitor the food level in the refrigerator and alert the user about empty items beforehand, and the house temperature can be controlled to a comfortable level before the occupant comes back home. Thanks to these smart networked applications, in the future, nearly every aspect of daily life can be planned ahead of time, therefore leading from the Internet-of-Things to an "Internet-of-Plans".

In general, there are three trends underlying smart networking and communications. First, we see significant increase in the number of networked sensors/devices. The more sensors, the more data we can gather about the physical world surrounding us, which enables us to make more informed decisions about daily life. Imagine a future household where every appliance/food item is going to be equipped with multiple sensors. We can envision that typically the number of sensors/devices is at a scale of 100's to 1000's. Second, there is a significant portion of stationary sensor deployments in the environment, as opposed to only

mobile devices. That makes many of the sensors not easily accessible and replaceable by human users. Also, many sensors are not owned by the users themselves, for instance, security cameras on roads/offices, temperature/activity sensors in public buildings, etc. They can keep collecting data continuously while not being noticed by the users. Third, the heavy use of data-driven and learning-based techniques and applications, which uses data analytics to enhance the cognition and prediction about the environment. This could happen at different scales, including locally such as an autonomous vehicle's perception of its surroundings, or more globally such as city-wide traffic optimization.

Protecting the security and privacy in smart networking and communications is of critical importance and also challenging. On the one hand, since both the actuators and human users' decisions will be highly dependent on the data collected by various sensors/devices, if the data is not genuine or modified by attackers (for example sensors can be hacked), there can be catastrophic consequences. Given the large number of devices, how to manage their security associations/authorizations and correctly configure them is also a challenge. Given the stringent time constraints of some applications (such as V2V communications), security solutions will need to provide a timely response to and recovery from attacks. On the other hand, much of the data collected by smart networking devices is sensitive in nature. For example, personal health data can be used to profile users, justify increases in users' insurance premiums, or deny coverage. It is crucial to protect the privacy of users' data while still being able to obtain useful data analytics.

To structure the above discussion, we identify following key challenges and several priority research areas.

The first and foremost challenge is to understand future threats and risks. Research in privacy and security is most effective if it is guided by reasonable threat models. Such models should take into account economic considerations such as the value of data and services. This is because many attacks in the real-world are motivated by financial gain. For example, the frequently reported data breaches against large companies mostly target data that can be monetized for profit, such as user passwords, social security numbers, credit card numbers, etc. Threat models should also take into account how easily potential attacks scale. Larger scale usually increases an attacker's financial gain and likely the attack's societal impact. For example, in the smart health domain, lots of sensitive personal health information is stored in cloud databases and a single compromise of such a database may yield a larger and more valuable dataset than several compromises of smaller locally stored datasets. Similarly, a remote attack over the Internet can more easily target a larger number of devices than a local attack over a wireless access network. Of course, not all attacks are economically motivated and research identifying such vulnerabilities and developing protections also has value. Some researchers may demonstrate attacks at smaller scales, but when they are related to people's safety and thus can potentially lead to big social impact by raising public awareness of the vulnerability. For example, in smart transportation, the first wireless car hack was demonstrated on tire pressure sensors on vehicles. Although such an attack has not yet happened at large scale, it helped spur developments towards developing more secure cars.

The second challenge is to protect the privacy and confidentiality of networked data. Such data may include locations in the case of a vehicular network, network operation details such as bandwidth usage in the case of an SDN, or smart meter usage data in the case of a smart grid. This form of information is increasingly collected and analyzed, and then utilized to control and enhance the network operation. For example, with vehicle location data one can estimate the traffic density and therefore avoid large-scale traffic congestion. Since such data are often sensitive in nature and data breaches are increasingly frequent nowadays, it is necessary to develop better privacy-enhancing techniques. This can include techniques to enhance the users' awareness and control of their own data, particularly with respect to where the data is transmitted and processed. Questions like who has access to their data and how they will use it arise. Oftentimes, this is not obvious and users are not aware of where the data flows. Thus, it is important to make the use of data more transparent by tracking data flows and to educate users to better understand their privacy risks. It can also include developing techniques for in-network privacy-preserving data analytics or machine learning, which do not require large centralized data stores. One possible approach is to apply tools from secure computation, such as homomorphic encryption. However, computational efficiency is still a challenge that needs to be resolved. Another possible direction is to adopt a secure execution environment, where limited computations can be executed in secure hardware. Aside from that, depending on the application scenarios, data may also be stored in an aggregated form when statistical information is sufficient. However, even if data are anonymized, by combining with other auxiliary data from other datasets, people can often be uniquely identified and their daily behavioral pattern would be disclosed to the public. In that case, to reduce privacy risks, privacy-preserving data publishing mechanisms will be required, and tools such as differential privacy or synthetic data generation may be adopted. A key challenge of using such tools to protect privacy is how to ensure an acceptable level of data utility. In smart networking, the utility could have quite different definitions from database applications, for example, in smart grid this can be defined as the optimality gap between the outputs of the demand response algorithm running over noisy smart meter data versus that over accurate data. On the other hand, studying policy engineering as well as policy and regulation aspects of privacy is equally important, particularly when such data is widely shared across national borders and regulatory domains.

A third challenge is managing many devices at large scale. With trends towards 1000s of devices per person, the following key questions arise: how can we associate their digital identities with their physical identities, and manage credentials and cryptographic keys in a secure and scalable way? How do we enforce different access privileges for different devices? And how do we correctly configure those devices and make sure their software is not tampered with? With such an increase in the number of devices, the amount of human effort in managing each devices needs to significantly decrease. For secure device association, for example, we need more automated ways of pairing devices with their intended targets, which involve little human effort. But automatically capturing the user's intent to pair the correct devices will be a difficult task. Proximity- and context-based authentication techniques can be helpful in this regard. Techniques based on emerging short-range communication techniques such as near-field

communication (NFC) are one example. For device configurations, usability is also a key challenge. In the case of a car, all the devices are carefully designed and integrated by the same manufacturer, so it is relatively easy to configure the devices; but for a building, devices may be installed by many different vendors over a longer span of time. Without a common system integrator and a single point of control, validating the device configurations and program correctness will be more difficult in the latter scenario.

Fourth, a key opportunity and challenge is exploiting and validating learning-based techniques. Machine-learning is expected to be widely used in smart networking applications because of its ease of use and the large volumes of training data that are available. Machine learning techniques may also prove useful in securing future networks, not just for detecting intrusions but also for automating the security configurations of systems consisting of vast numbers of devices. It is critical to validate the correctness of learning results and the security properties of such deployed learning algorithms. It has been reported that some classical learning algorithms are easy to fool, for example, slightly modifying a training image dataset will yield completely different image recognition results. Compared to traditional pattern recognition problems, here we need to consider an adversary that intelligently chooses modifications, and tries to hide its trails to evade detection. An additional challenge in deploying learning-based security approaches are frequently high false positive rates, which can cost a lot and degrade the effectiveness of the solution. For example, in financial fraud detection systems, a false positive is directly linked to loss of time of workers or customers and thereby economic losses. Thus, it is worthwhile to study and advance machine learning techniques, in the setting of smart networking and communications under adversarial conditions.

A fifth key challenge is security from a system design perspective. It is important to fully characterize the attack surface of a complex system. Given trends to more software defined networks, the network implementation will be more malleable and the attack surface can be expected to increase. The security of a system always depends on the weakest link. For example, in some modern vehicles, the same CAN bus interconnects communication, control, sensing, entertainment functionalities, which makes it easier for a remote attacker to gain control of critical units like engine and brakes by hacking into another vulnerable sub-system, say entertainment. Thus, it is a good practice to carefully separate important functions, such as control and communication. However, due to cost limitations in the real-world, this often brings a challenge to secure system design. To be compatible with legacy systems, oftentimes deploying a clean-slate design may not be realistic, necessitating more incremental changes. In addition, for systems that have a control plane and data plane (such as SDN), separating these two planes is also desirable. Also, solutions can retain certain key functions in the hardware layer to help secure the software layer, essentially by developing a small trusted computing base rooted in hardware protections.

In addition to system level protection, novel mechanisms are needed to secure the communication among the devices and transmitted data, which should tailor to the unique features of smart networking environments. To ensure the confidentiality, authenticity and integrity of the communication, cryptographic approaches are likely to be used on many

devices. Continuing advances in processing capabilities will allow using such techniques on a wide variety of extremely small sensor devices. However, IoT devices are heterogeneous and we expect classes of devices for new applications that are so energy- and size-constrained that traditional crypto-based protocols are too computationally intensive. It will be necessary to develop lightweight solutions, and to use different levels of protection for different classes of devices. For example, schemes that outsource/offload heavy crypto computations to a cloud while maintaining simpler cryptographic operations on the resource-constrained devices may be worth exploring. Another interesting approach will be to use high-end devices to protect low-end devices. Moreover, since sensors can be hacked/spoofed, ensuring the trustworthiness of gathered data itself is also a priority. This requires more than cryptographic mechanisms, and solutions could exploit the intrinsic redundancy of the data sources. For example, if knowing which data sources are more trusted, allows applying data fusion for secure state estimation. It will also be worthwhile to explore physical properties that are not easily forgeable as trusted sources. For example, emerging communication techniques like visible light communication (VLC) or near-field communications (NFC) have different security properties than the conventional radio frequency wireless channel, which can help secure systems. Also, physical-layer based security approaches can be adopted to complement the security of upper-layer protocols or to secure miniscule devices not capable of executing cryptographic algorithms.

Action Plan: The discussion group makes the following recommendations based on the panel presentation and breakout discussions in the security and privacy thrust.

- Future research on new threat and risk models in future smart networks that take into consideration of both financial gains and social impacts should be encouraged.
- New techniques should be developed to protect the privacy and confidentiality of networked data, e.g., by the means of privacy-preserving data analytics or machine learning, privacy-preserving data publishing, policy engineering, and new mechanisms to enhance the users' awareness and control of their own data.
- The scalability problem should be investigated for securing future smart networks under network densification. Specifically, managing digital identities, cryptographic keys, access privileges, software configurations and human efforts, when the network size grows to 1000s of connected devices per person should be addressed.
- Research should be carried out to understand how to exploit machine learning-based techniques to secure smart networks, how to validate the correctness of learning results, and the robustness of these techniques.
- The workshop also recommends that security should be thoroughly studied from a system design perspective by fully characterizing the complete attack surface of a complex system.

2.4 City-Scale Smart Network Testbed Platforms
(Leads: Bhaskar Krishnamachari, Prasun Sinha; Scribe: Tam Vu)

Goal: Obtaining repeatable experimental results is a critical need of the wireless community today. When the testbed is a "city", running at-scale evaluations is both challenging but

incredibly rewarding, as the outcomes directly benefit the general population. This thrust posed the following questions:

- What kind of experimental systems the community wishes to use at the city-scale, what outcomes/results will be interesting to the community?
- What resources must be made available to encourage development and user participation, and how we ensure the sustainability of the testbeds?
- How best to overcome the learning curve needed to use such platforms?

This thrust was designed to provide inputs to both small to large-scale investment programs by NSF, such as Platforms for Advanced Wireless Research (PAWR).

Summary of discussions: There has been a persistent need for repeatable, verifiable and at-scale experimentation. The important challenges, issues, and potential solutions in building testbeds for smart networking specifically and networking research in general are described in this section. The topics of interest include testbed's sustainability, usability, size and capacity, accessibility, bootstrapping, and the use of testbeds for replicable networking research results.

1. **Sustainability**: A fundamental question on testbed design is "how do we make sure that the testbed will live on even after the initial round of funding ends?" It is common that testbeds, especially ones that include hardware and software updates, require continuation of funding for technical support, software update, and hardware maintenance, apart from other unforeseen causes. The unified line of thinking within the discussion group clearly articulated that government agencies should have a plan to provide funding to keep the testbed alive, which might include budget items for staff, engineers, and student support. This might be reflected in awards that are of a bigger size compared to the current ones. Additionally, there must be active encouragement for researchers to work on the testbeds, and these can through separate credit ratings for outcomes demonstrated on a testbed, requirements for some forms of more applied research to be definitely deployed on the testbed, among others.

2. **Size and capacity**: The current testbed funding model, tends to award larger scale testbed proposals only. The group suggested NSF to fund a more diverse set of proposals with different scale, from small to large, that serves not only the research agenda but also teaching for both local and remote users. There are many projects that need only a relatively small number of nodes (on the order of dozens to a hundred) and it can be much easier for PI's and their students to work with portable devices on their own premises than access them only through remote testbeds. Specifically, there was a sense from the group that researchers would like to see successful NSF grants for more small and medium sized testbeds as well, possibly in conjunction with regular research grants. The testbed section can be weighted to be an important piece of the proposal instead of appearing as an afterthought, as it happens often.

3. **Bootstrapping**: Most testbeds are bootstrapped through NSF funding, with only a few exceptions from industry. The community needs to be more active in seeking support from other funding agencies and broaden the partnership with industry, with NSF help. The question remains as to how to attract industrial partners and what value can we bring into such partnerships to attract a deeper involvement from the industry. Examples include considering

ways to monetize deployed sensor networks, simplifying IP issues, and also at a larger governmental scale, providing financial breaks to the companies who choose to partner for academic research.

4. **Accessibility**: There are several existing testbeds small to medium-scale testbeds that can be federated to create easy access and diverse experimental trials. For example, GENI can provide its users access to Orbit, CloudRack, and many more. That makes it easier for users to create and manage their accounts on different testbeds. The suggestion to NSF is for a call for increase in support for research on innovative ways to connect testbeds from different domains. Furthermore, it would be beneficial for PIs if NSF can lead an effort to put together a list of all available testbeds that it supports, especially the ones from different domains (e.g. health, transportation, smart building, etc.) This can be a "cheat sheet" that is available on the NSF website, with incentives provided to the testbed owner/designer to encourage external use.

5. **Usability**: Many testbeds require a very steep learning curve for its users, which most of the time are graduate and undergraduate students. This sometime reduces the utilization of the testbed. To address this problem, helping testbed users to get the best out of the existing testbed in the least amount of time and learning efforts. One way to address this issue is that NSF should support more testbed workshops in which tutorials are provided to help students and PIs. Each testbed award must have also requirements of creating playbooks that allow a step-by-step use and application examples, early on in the development timeframe. This can also allow the testbed developers to get community feedback. In addition, NSF should provide funding to support summer camps, student exchanges, and visiting professors among laboratories that have testbeds and the ones that need to use the testbed. It is also important for NSF to take stock of the utilization of testbeds across different sub-communities within networking - e.g., there was some sense in the workshop that the GENI testbeds had not been utilized widely in the wireless and mobile networking community compared to networking communities focused on traditional wired networks.

6. **Reproducibility of networking research results**: Most of current networking research results are not replicable. However, it is critical for the community to confirm these research claims and outcomes, especially the fundamental ones, upon which new research will be built. This problem can be addressed, or at least mitigated, by taking advantage of reliable testbeds and benchmarks. That is, if a tutorial of how to replicate the results is provided clearly, and the original experiment was performed on a stable and widely accessible testbed, its results can be replicated. However, the current publishing mechanism does not incentivize researchers to do so. Though flagship conferences have recently started to accept experiment/experience papers, it is not the norm and won't be the mainstream of those conferences in the foreseeable future. Therefore, authors do not have an incentive to invest their time and effort in making their system more repeatable and replicable. A suggestion to NSF is to sponsor organizing competition and challenges to promote reproducible and replicable research results.

Action Plan: The discussion group converged on the following key action items:

- The community felt that NSF should continue to fund both small and large-scale testbeds, with added emphasis on federating them. Continued external usability should be a metric in assessing annual performance of the testbed team, as well some form of "credit" system should be designed to encourage researchers who actually use these testbeds.
- There could be a single, comprehensive directory of NSF funded community infrastructure efforts with playbooks to access them. There could be innovative challenge competitions using existing testbeds to stimulate more use.
- Some changes in the merit review criteria for proposals can help: for core proposals, implementing the research on a testbed and validation plan could be weighted concretely; for proposals specifically on infrastructure, defining and appreciating near-term and long-term goals can help bring in more industry involvement.
- The PAWR model can be scaled down to broaden industry partnerships and share funding commitments for future testbeds. Smaller-sized mixed NSF-industry testbed calls targeting niche areas can be created, with specific list of revolving topics in the solicitations. Solving difficult issues, such as IP sharing, at the central level is key.

3. Conclusion

The workshop generated lively discussions on several topics that will undoubtedly influence the growth and adoption of future smart networks. A continued point of debate is what is "smart" compared to the status quo, and if at all the possibilities of wireless technology that lie ahead in a 10-20 year timeframe can be realistically analyzed today. Historical examples from the past decades reveal many instances where prediction accuracy has faltered. Thus, the workshop attendees agree that recommendations need to be continuously revisited and updated over time. In addition, documenting the vision today and returning back to it every few years can also help in answering an important question: is the community as a whole tackling the most relevant problems, or, more ambitious tools, techniques and applications need to be devised to not only keep up with the state of art but anticipate it ahead of time. Smart networking and communication can take many forms, and the workshop as a whole recognized that isolating any single area for investment and improvement, be it systems, architectures, protocols, applications or experimentation, at the expense of one or more of the others may be catastrophic, as they all interconnect at multiple touch-points.

Appendix A. Workshop Program

Friday, May 05, 2017.

7:15 – 8:00am
Breakfast

8:00 am – 8:15 am, Room: Atlanta 1
Opening Remarks, Wenjing Lou, NSF

8:15am – 8:30 am, Room: Atlanta 1
Workshop Overview by Workshop Co-chairs, Hongyi Wu and Kaushik Chowdhury

8:30 am – 9:30 am, Room: Atlanta 1
Panel A: Discussion on Workshop Topics 1 and 2
Panelists: Henning Schulzrinne, Eytan Modiano, Rajarathnam Chandramouli, and Samir Das

9:30 am – 9:45am
Coffee break

9:45 am – 11:15am
Parallel breakout sessions
Session 1 for Topics 1, Lead: Henning Schulzrinne, Scribe: Saswati Sarkar, Room: Atlanta 1
Session 2 for Topics 2, Lead: Rajarathnam Chandramouli, Scribe: Jia (Kevin) Liu & Bin Li, Room: Atlanta 2

11:15 am – 11:45pm, Room: Atlanta 1
Breakout groups reconvene (Group leaders summarize the discussions on Topics 1 and 2)

11:45 am – 12:45 pm
Lunch

12:45 pm – 1:45 pm, Room: Atlanta 1
Panel B: Discussion on Workshop Topics 3 and 4
Panelists: Marco Gruteser, Yingying Chen, Bhaskar Krishnamachari, and Prasun Sinha

1:45 pm – 2:00 pm
Coffee break

2:00 pm – 3:30pm
Parallel breakout sessions
Session 1 for Topics 3, Lead: Marco Gruteser, Scribe: Ming Li, Room: Atlanta 1

Session 2 for Topics 4, Lead: Bhaskar Krishnamachari, Scribe: Tam Vu, Room: Atlanta 2

3:30 pm – 4:00 pm, Room: Atlanta 1
Breakout groups reconvene (Group leaders summarize the discussions on Topics 3 and 4)

4 pm: Adjourn

Appendix B. Workshop Attendees

- Stefano Basagni, Northeastern University
- Guohong Cao, The Pennsylvania State University
- Xiaojun Cao, Georgia State University
- Samir R. Das, Stony Brook University
- J. J. Garcia-Luna-Aceves, University of California, Santa Cruz
- R. (Mouli) Chandramouli, Stevens Institute of Technology
- Yingying Chen, Stevens Institute of Technology
- Kaushik Chowdhury, Northeastern University
- Michelle Effros, California Institute of Technology
- Eylem Ekici, The Ohio State University
- Wei Gao, The University of Tennessee, Knoxville
- Mario Gerla, University of California, Los Angeles
- Marco Gruteser, Rutgers University
- Zygmunt J. Haas, Cornell University
- Tom Hou, Virginia Tech
- Dimitrios Koutsonikolas, University at Buffalo
- Bhaskar Krishnamachari, University of Southern California
- Bin Li, University of Rhode Island
- Ming Li, The University of Arizona
- Jia (Kevin) Liu, The Ohio State University
- Wenjing Lou, NSF & Virginia Tech
- Shiwen Mao, Auburn University
- Janise McNair, University of Florida
- Eytan H Modiano, Massachusetts Institute of Technology
- Tamer Nadeem, Old Dominion University
- Saswati Sarkar, University of Pennsylvania
- Henning Schulzrinne, Columbia University & FCC
- Prasun Sinha, The Ohio State University
- Violet R. Syrotiuk, Arizona State University
- Jian Tang, Syracuse University
- Damla Turgut, University of Central Florida
- Tam Vu, University of Colorado Boulder
- Wenye Wang, North Carolina State University
- Yu Wang, University of North Carolina at Charlotte
- Jie Wu, Temple University
- Hongyi Wu, Old Dominion University
- Linda Xie, University of North Carolina at Charlotte
- Chunsheng Xin, Old Dominion University
- Guoliang Xue, Arizona State University
- Gang Zhou, College of William and Mary