

Research Statement for Arunesh Mishra

Research Engineer, Mobile, Google Inc..

Overview of Research

My area of research expertise is in wireless networks, security and their applications. I started my research at the University of Maryland, College Park, by working on network solutions provide secure and seamless connectivity to network applications as the user moves within the networks area of coverage. Subsequently, I've worked on challenging research problems in spectrum sharing, channel allocation and energy efficiency in wireless networks. Most of my work has been published in top tier conferences and journals. I've been routinely invited to be a part of the Program Committees of reputed Conferences and Workshops and routinely reviewed papers for top tier Journals. Below, I've described my research performed as a Postdoc at the University of Wisconsin, Madison, followed by my research conducted while pursuing my PhD Dissertation, and finally my latest research contributions (and ongoing projects) at Google.

Research as a Postdoctoral Scientist at Univ of Wisconsin, Madison

Client-driven Channel Management for Wireless LANs

Spectrum is scarce and thus proper utilization of 802.11 channels is critical to the performance of a WLAN. Existing methods for assignment of channels to APs are either static or are "AP-centric", that is they do not capture performance metrics at the clients. My recent work (published in *ACM MCCR 2005 [1]* and *Infocom 2006 [2]*) has shown that such approaches can cause severe under-utilization of spectrum. In this work, we also designed and implemented a "client-driven" approach to channel management that uses simple feedback mechanisms from clients to better assign channels to APs. This work highlighted the broad concept of the necessity of client participation in network management for WLANs.

There are a number of unique aspects of this work compared with previous work. First, we demonstrated that for best client performance, the channel assignment problem should be solved *in conjunction with* the client-AP association problem. Second, our solution approach illustrates some problems associated with a graph-based formulation of the problem and demonstrates that a *set-based* formulation better models the different constraints. We expect this idea to apply in many other wireless scenarios. Third, we developed centralized solutions that apply to enterprise WLANs, as well as distributed variants that would co-exist in an environment where multiple WLANs share the same spectrum. To my knowledge, this was the *first paper to define the client-driven approach and demonstrate its potential* for wireless network management in various environments.

In my follow-up work (published at *ACM Mobicom 2006 [3]*) we tackle the problem of channel management for uncoordinated wireless deployments where cooperation between APs cannot be assumed. These entities can be, for example, neighboring coffee shops, apartments and small businesses who mostly maintain a single-AP WLAN. Such APs that interfere with each other are collectively faced with the challenge of selecting a good assignment of channels. We built a fully distributed mechanism to allocate channels which is client-driven and also utilizes partially-overlapped channels for maximum spectrum use. The solution is built on the concept of channel hopping and is algorithmically shown to provide a fair division of the spectrum resources among competing APs while maximizing the overall network and spectrum utilization.

Spectrum Assignment with Partially-overlapped Channels

My next direction of work allows for further improvement in spectrum utilization beyond channel management by building the theory and techniques to take advantage of partially overlapped channels apart from non-overlapped

ones. We observe that the 2.4 GHz band defines 3 non-overlapping and 11 partially overlapped channels. In our initial work, (*IMC 2005 [4]*) we demonstrated that the standard practice of “avoiding” partially overlapped channels due to the interference between such channels is actually causing underutilization of spectrum. We showed that with proper models of the interference effects in partially-overlapped channels it might in fact be possible to use such channels to attain a significantly better system design.

In our followup work (*ACM Sigmetrics 2006 [5]*), we designed an accurate and efficient model of the performance of partially-overlapped channels for different wireless standards, such as 802.11, 802.16, and so forth. We developed algorithms that use the model to efficiently assign partially-overlapped channels to wireless interfaces, leading to a *factor of three improvement in achieved throughputs* for many hypothesized wireless LAN and mesh network workloads. Through analysis, simulation, and implementation, this work, thus debunked the previous practice of restricting channel assignment choices to non-overlapping channels alone. Our proposed model is fairly general; for example, we demonstrated that using the model, we can take *any* existing channel assignment algorithm that currently assigns only non-overlapping channels and derive another channel assignment allowing partially-overlapped channels. This work has the potential for widespread use in the widely adopted standards for spectrum management for enterprises, homes, and other wireless hotspots.

We have a choice of a multitude of wireless technologies today, each optimized for a single and distinct service model – for example, cellular for voice and WiFi for data. The next challenge is thus, convergence, which places the user in control by allowing for a user-driven model of service rather than the technology dictating what services are available. My thesis has examined one part of this puzzle – understanding the challenges faced by voice applications when using WiFi. The focus has been on handoffs and the related latencies faced by users that are mobile within an 802.11 network. A unique aspect of my thesis work has been the design of practical system-level solutions which work transparently with the standards in reducing the handoff latencies from 1.2 seconds to under 50 ms. Based on the concept of what I call Neighbor Graphs, these solutions provide the same level of authentication, confidentiality and access control properties as the original methods but within the 50ms budget. I believe that this concept could have interesting applications beyond handoffs, for example, in providing better quality-of-service, aiding network management and improving network security.

A next broad direction of my work has been in optimizing the performance of WLANs by building practical and intelligent channel assignment algorithms. One unique contribution of this work is the concept of client-driven network management for WLANs and construction of protocol-level methods to realize this. The work shows that client participation in channel assignment and for network management in general, is critical for operating the network efficiently at peak performance. I have also built distributed algorithms to manage 802.11 channels in an uncoordinated wireless environment, for example, 802.11 access points (APs) belonging to individual hotspots and apartment homes which interfere and co-exist with each other. These distributed algorithms perform a fair and near-optimal division of the system’s resources using the concept of channel hopping. Finally, through careful observations and analysis we’ve also debunked the traditional approach of avoiding partially overlapped channels which have been defined by the IEEE standards body. I’ve shown that by careful modeling of the interference between two partially overlapped channels, one can utilize these channels effectively, thus improving the overall system throughput by a factor of three.

Below, I describe each of these research contributions in further detail.

PhD Dissertation Work – Wireless Handoffs and Security

This is research that I performed as a part of my PhD Dissertation work at the University of Maryland, College Park. As a user moves within a wireless network (cellular or WLAN), he changes his point-of-attachment (called an association) to the network through a procedure termed handoff. Handoffs in general, incur certain latencies during which the user is unable to send/receive application traffic. This causes an adverse affect on stateful network protocols such as TCP and also acts as a major deterrent for synchronous multimedia applications such as voice and streaming video. My thesis address this problem in the context of 802.11 WLANs.

• **Analysis of 802.11 Handoff Latencies:** Through careful measurements and experiment designs over a testbed network of 40 APs (built off Soekris NET4521/OpenBSD), I've performed an in-depth analysis of the handoff process for 802.11. This analysis resulted in the first and widely cited paper published at *ACM Sigcomm CCR*, 2003 [6], that explains how handoffs happen in 802.11 networks by presenting the algorithms that various wireless network interface (NIC) vendors implement, provides insights into where the latencies are incurred, what the latency numbers are and the factors that affect them. This paper received considerable attention in both the research community and the industry; it also attracted funding from Samsung Electronics. Through this analysis, we categorized handoff latencies into two: (i) Scanning: The mobile user scans available 802.11 channels in order to search for APs. This latency was measured between 100 - 400 ms depending on the algorithm used for scanning. (ii) Authentication: Based on the IEEE 802.1X standard which provides for robust certificate-based authentication and access control, this phase incurred a high latency of about 800 ms. The combined latency comes to about 1.2 seconds which is far too high for synchronous multimedia applications and also poses a significant hurdle for everyday TCP-style applications.

• **Software Artifacts - Open1x.org Effort:** The measurement of the authentication latency was based on my implementation of the IEEE 802.1X standard. This was made available as opensource on open1x.org which was founded by me in 2002. Today, this code is maintained by a team of enthusiastic graduate students and is popularly used on Linux platforms especially Debian. Through this implementation, we also demonstrated the possibility of mounting Man-in-the-Middle attacks on an 802.1X system. This work gained media publicity (cited on *CNN.com*, Feb 2002) and was also published as a journal paper (*Wiley Journal of Wireless Networking*, 2004 [7] and downloaded about 350 thousand times till date).

• **Concept of Neighbor Graphs:** My thesis is based on the central notion of what I call, *Neighbor Graphs*, defined as a graph structure over the set of APs that comprise a given wireless network. The edges capture unique mobility properties of the users in the network environment. A directed edge is placed between two access points if they act as successive points-of-attachment for users, that is, users handoff between those two access points. Neighbor graphs, which are built over this unique and simple concept, provided the base for building a set of system-level solutions to tackle the scanning and the authentication latencies. The initial work was published at *IEEE Infocom 2004* [8].

• **Solutions for Fast Authentication and Scanning:** By carefully analyzing the authentication mechanism used by 802.1X, I built a key distribution scheme (published in *Proc. of IEEE Wireless Communications*, 2004) based on neighbor graphs, that works with the 802.1X standard, the 802.11i wireless security standard and allows for fast re-authentication with the next access point during handoff. By intelligent use of standard cryptographic techniques and proofs, I showed that this scheme provides the same level of strength as the original 802.1X/802.11i mechanism. By implementing this technique over the 40-AP testbed, we showed that the latency costs were much less at about 5 ms as compared to the 800 ms cost incurred earlier. Called Proactive Key Distribution and Caching, this technique was incorporated as recommended practice into the IEEE 802.11i and IEEE 802.11f standards. Using neighbor graphs, we built a fast scanning algorithm (published at *ACM Mobisys 2004* [9]) that reduced the scanning latencies by avoiding wasteful scan operations without compromising on the quality of the scan results. This was implemented using the Airjack driver for Prism 802.11b wireless cards, and scanning latencies of under 50 ms were achieved.

Overall, through my thesis research, I was able to build a full system for fast handoffs that achieved a 50 ms budget for the handoff latencies (down from 1.2 sec). These solutions did not compromise either on the quality of scan results or the strength of the 802.1X based authentication methods. Moreover, this system was based on a single and powerful concept of neighbor graphs which uniquely captures the topology of the wireless network from the angle of user mobility. This research has given me immense systems expertise and lessons on how to design good practical solutions which I hope to apply to future research challenges.

Research Accomplishments at Google

At Google, I've helped build the core technologies behind the Google's Location System, which is available through a wide range of Google's products, including Search, Toolbar, Desktop and most popularly Google Maps for Mobile. Until this effort from Google, Location-based services had been a promising market that never saw a real adoption. As a part of my work at Google, I am researching and designing innovative location technologies that help find the location of the entity of interest, which could be a user, a device or such. As a core value, the technologies are open, and available freely to web and mobile developers thereby allowing for unrestricted innovation in location-based services and applications.

Cellular Network-based Localization System:

Cellular networks use base-stations to divide the desired region of coverage into smaller 'cells'. A single user is typically within the wireless receiving range of multiple such base-stations possibly from different service providers. At Google, I have contributed towards creating new algorithms for deriving a user's location by observing the base-stations that the user is within range of. By algorithmically analyzing the signal properties of the communication between a user and a base-station, one can obtain an estimate of the distance of separation of the user from the base-station. Combining this with the data from a small number of users who have GPS on their mobile devices, we are able to (i) determine a base-station's location, (ii) determine any remapping of the base-station's location in almost near-realtime and (iii) compute a user's location based on the distance to a base-station and the base-stations location. This is very useful for about 80% of all Google's Mobile users who do not have GPS tracking on their phones. I researched and designed statistical algorithms to determine the base-station's location and any remappings done to the base-station identifier by service providers. The algorithms compute various spatial and temporal statistical metrics to predict a base-station's location and remappings. The system effectively deals with challenges such as noise in a user's contributed data which include errors GPS locations, timestamping methods and other inputs obtained from a user.

Impact: This is a critical piece of the technology that powers Google's Mobile Maps feature called 'MyLocation' launched in early 2008. This technology powers the Google Mobile Maps software for location computations and has grown to become the most successful mobile application. There have been over 100 million installs till date and has received wide publicity in the press. A patent is being filed and I have received an "Innovation Award" from the Executive Management Group at Google. This technology is also being used behind Google's Latitude launch which has seen huge growth and also being used by other Google's mobile products such as search.

Wifi Localization system

One of my core contributions at Google has been the research, design and implementation of a Wifi-based localization system. Based on my expertise in wireless and IEEE 802.11 communication technology (also called Wifi), I was able to design a scalable, fast and accurate positioning technology based on Wifi. Wifi-based positioning is complementary to cellular in that it brings location-awareness to the desktop platforms (such as laptops, smartphones, PDAs). I designed the algorithms to compute a Wifi Access Point's (AP's) location using physical layer attributes such as signal strength, noise, 802.11 frame-type and such. An innovative aspect of this work is the use of data-rate in the localization algorithms as a superior attribute over the traditional signal strength. This is novel in the research literature on localization systems and has been successfully implemented as a part of Google's Wifi Localization System. The technology is patent pending.

Impact: Wifi-based Localization is available as part of Mozilla's Firefox browser, the Gears API which allows any web developer to use this technology on their website. Its also integrated into the Google's Chrome browser and works on the Google's Android platform. Its utilized by Google's Mobile Maps software on phones that have a wifi interface and takes precedence over the cellular localization system since this has a much higher accuracy.

Location Security and Privacy

User location is a type of information that is extremely sensitive. The mere availability of this information can have a wide range of ramifications for a user. Thus, for the success of the web on mobile platforms, users need to be empowered with technologies that let them control and configure who has access to their location and to what granularity. I am actively working on building the next generation protocols for privacy, anonymity and access control to a user's location information. There are two specific projects that I am working on :

- **Probabilistic Spam Detection While Preserving User Anonymity:** One of the core challenges faced by Google's Location Platform is the contribution of bad or malicious GPS data that ties Cellular Base-stations and Wifi Access Points. My work has resulted in the creation of a protocol where its possible to identify a spammer based on the quality of the data points contributed without attaching an identity with every user. This protocol is a novel application of bloom filters alongwith some cryptographic properties.
- **A Scraping-resistant Prefetching Algorithm:** As an optimization to provide faster localization results, Google's clients, such as the Google Maps for Mobile, prefetch location mappings of cellular base-stations and wifi access points and store them for local location computations. Google's Location Server supports this prefetch 'mode' which makes it easy for scrapers. As a part of my work, I've built algorithms where we send a hashed mapping of the location and the identifier (CellID or Wifi) such that unless the user's device knows the identifier (by being at the physical location), the user's software is unable to compute a location. This acts as an important defense and protects the location database from scraping.

Ongoing/Future Research at Google

I am keen on understanding and solving some of the big challenges we face today as we try to move to a wireless future where the user is in control and not the technology. This vision is attainable through a combination of two interesting directions of research that I might pursue in near future.

- **Fine-grained Spectrum Sharing:** Popular statistics show that over 90% of the licensed spectrum remains unused at any point of time. Such spectrum could alleviate some of the performance issues faced by the WLAN technology which operates in the crowded unlicensed bands. This leads to the vision that tomorrows devices would be able to make effective use of even the licensed bands by operating in a cooperative fashion with the incumbent licensed users. There are many challenges that remain to be addressed both at the wireless and the networking layer before such a system can be realistically deployed. I hope to be able to understand and build protocol-level methods to engineer the next-generation of such networks. Some of my initial research in this direction has just appeared as a paper in *IEEE Hotmobile 2007* [10].

- **Secure Localization :** Traditional methods for localization in wireless networks rely on the correlation of the received signal strength with physical distance. It is also well known, that these mechanisms fail in an adversarial setting due to the lack of robustness of the signal strength property to malicious intent. In this ongoing research, I am studying a property of the wireless medium, which I call 'wireless congruity', that captures the *relative similarities* in wireless media characteristics (such as packet receptions, idle channel time, etc.) as observed by two receivers that are in physical proximity of each other. I've shown that wireless congruity holds promise for secure localization by presenting an results from an empirical study performed through extensive experimentation in a rich indoor wireless environment. This has just appeared as a paper in *IEEE Hotmobile 2007* [11]. I feel that this direction of work holds promise for building secure and robust location authentication methods that could apply to a wide range of wireless technologies.

With this experience, I hope to employ my skills and expertise to new and challenging research problems in the broad domain of systems, wireless/wired networking and security.

References

- [1] Arunesh Mishra, Suman Banerjee, and William Arbaugh, “Weighted coloring based channel assignment for wlans,” *Mobile Computer Communications Review (MC2R)*, vol. 9, no. 3, July 2005.
- [2] Arunesh Mishra, Vladimir Brik, Suman Banerjee, Aravind Srinivasan, and William Arbaugh, “A client-driven approach for channel management in wireless lans,” in *IEEE Infocom*, 2006.
- [3] Arunesh Mishra, Vivek Shrivastava, Dheeraj Agrawal, and Suman Banerjee, “Distributed channel management in uncoordinated wireless environments,” in *ACM Mobicom*, 2006.
- [4] Arunesh Mishra, Eric Rozner, Suman Banerjee, and William Arbaugh, “Exploiting partially overlapping channels in wireless networks: Turning a peril into an advantage,” in *ACM/USENIX Internet Measurement Conference (IMC)*, 2005.
- [5] Arunesh Mishra, Vivek Shrivastava, Suman Banerjee, and William Arbaugh, “Partially overlapped channels not considered harmful,” in *ACM Sigmetrics*, 2006.
- [6] Arunesh Mishra, Minh Shin, and William Arbaugh, “An empirical analysis of the ieee 802.11 mac layer handoff process,” in *Computer Communications Review (ACM SIGCOMM) April*, 2003.
- [7] Arunesh Mishra, Nick L. Petroni Jr., William Arbaugh, and Timothy Fraser, “Security issues in ieee 802.11 wireless-local area networks,” *Wiley Interscience Wireless Communications and Mobile Networking Journal (Wiley Wireless)*, 2004.
- [8] Arunesh Mishra, Min ho Shin, and William A. Arbaugh, “Context caching using neighbor graphs for fast handoffs in a wireless network,” in *Proceedings of IEEE Infocom*, 2004.
- [9] Minh Shin, Arunesh Mishra, and William A. Arbaugh, “Improving the latency of 802.11 hand-offs using neighbor graphs,” in *Proceedings of the MobiSys*, 2004.
- [10] Suman Banerjee, Arunesh Mishra, Vladimir Brik, Vivek Shrivastava, and Victor Bahl, “Towards an architecture for efficient spectrum slicing,” in *IEEE Hotmobile*, 2007.
- [11] Arunesh Mishra, Shravan Rayanchu, Ashutosh Shukla, and Suman Banerjee, “Towards secure localization using wireless congruity,” in *IEEE Hotmobile*, 2007.