RESEARCH STATEMENT

Junxian Huang (hjx@eecs.umich.edu)

Research Overview

The use of cellular data networks is increasingly popular due to the widespread deployment of 3G/4G technologies and the rapid adoption of smartphones. Besides email and web browsing, a variety of network applications are now available, rendering smartphones potentially useful substitutes for their desktop counterparts. My research focus lies in the networking and energy aspects of smartphone applications and cellular networks.

In this statement, I first discuss my existing works in the cellular network and smartphone space, followed by a few other projects that I have worked on previously. Towards the end, I will talk about the future research agenda.

Existing Research: Network and Energy Efficiency

1. Characterizing Cellular Network Performance:

Since 2008, I have been working on devising systematical methodologies and developing tools for characterizing cellular network performance directly from end users. The tools I have developed (as the lead developer) includes *3GTest* [6], *4GTest* [1] and *MobiPerf* [3], which have cumulatively over 150,000 users from over 190 countries or regions. Notably, *MobiPerf* has received both the *Open Internet App Award* and the *People's Choice App Award* in the *FCC Open Internet Apps Challenge* [2]. It is now an open-source project [5] that I am actively working on and this project is in joint collaboration among University of Michigan, M-Lab, Google and University of Washington. In these measurement tools, I have devised methods to accurately measure round-trip time (RTT), DNS lookup time, uplink/downlink bandwidth, loss rate, and other network performance metrics for 3G, WiMAX and LTE 4G networks and compare those with WiFi networks.

We also study the cellular data network infrastructure and its implications on the performance of CDN services [15]. We observe that placing content servers close to GGSNs can potentially reduce the end-to-end latency by more than 50% excluding the variability from air interface.

2. Anatomizing Smartphone Application Performance:

In order to understand the key factors that affect smartphone application performance, I develop a systematic methodology for comparing this performance along several key dimensions such as carrier networks, device capabilities, and server configurations [10]. I perform detailed analysis to help carriers, phone vendors, content providers, and application developers gain insight. For example, for carriers, I infer various network level problems, *e.g.,* high latency or high loss rate, which they can directly take action on. For phone vendors, I identify performance bottlenecks on the devices or issues associated with the content. These issues can be resolved either independently or by cooperating with content providers. And for application developers, I evaluate factors such as the overhead of HTML rendering and Javascript execution given a particular software configuration. I also perform longitudinal study that compare the determinant factors on smartphone applications in 2010 [10] and 2012 [7]. I identify that the performance bottleneck for web-based applications lies more in the devices processing power than in the network, indicated by the high average CPU usage of 79.3% in LTE network, as well as the underutilized network capacity due to small object size in typical web transactions.

3. Understanding Power Characteristics of 4G LTE Networks:

With the recent advent of 4G LTE networks, there has been increasing interest to better understand its power characteristics, compared with 3G/WiFi networks. I take one of the first steps in this direction [7] and develop the first empirically derived comprehensive power model of a commercial LTE network with less than 6% error rate and state transitions matching the specifications. Using a comprehensive data set consisting of 5-month traces of 20 smartphone users, I carefully investigate the energy usage in 3G, LTE, and WiFi networks and evaluate the impact of configuring LTE-related parameters. Despite several new power saving improvements, I find that LTE is as much as 23 times less power efficient compared with WiFi, and even less power efficient than 3G, based on the user traces and the long high power tail is found to be a key contributor.

4. Optimizing Energy Usage in Cellular Networks:

Todays cellular systems operate under diverse resource constraints: limited frequency spectrum, network processing capability, and handset battery life. We observe that periodic transfers incur tremendous energy and signal overhead in today's smartphone traffic [12]. In addition, we consider a novel and important factor, handset screen status, *i.e.,* whether

the screen is on or off, which was ignored by previous approaches for optimizing cellular resource utilization [8]. We find that off-screen traffic accounts for 58.5% of the total radio energy consumption although their traffic volume contribution is much smaller. Such unexpected results are attributed to the unique cellular resource management policy that is not well understood by developers, leading to cellular unfriendly mobile apps. We then make a further step by proposing screen-aware optimization, by leveraging the key observation that screen-off traffic is much more delay-tolerant than its screen-on counterpart due to a lack of user interaction. Our proposal can better balance the key tradeoffs in cellular networks.

In another work in submission, I develop a novel system, which dynamically and intelligently determines in real time when to deallocate radio resources for a UE based on predicting the network idle time. Using 7-month-long real-world cellular traces, I comprehensively evaluate it using various traffic features and machine learning algorithms. Properly configured, it correctly predicts 85.88% of idle time instances, achieving radio energy savings of 59.07%, at the cost of 91.01% of additional signaling overhead, significantly outperforming existing proposals. it incurs negligible energy overhead and has fast response times, demonstrating the practicality of deploying the system on contemporary smartphones.

5. *Effect of Network Protocol and Application Behavior on Performance for LTE Networks*:

Despite it fast increasing user base, the network performance and the interplay between mobile applications and the network for the LTE networks still remain unexplored. In one work in submission, I thoroughly study the network performance of LTE network with a data set covering around 300,000 LTE users in a large metropolitan area for 10 days. I revisit basic network metrics in the LTE network and compare with previously studied network conditions. I also observe that a high downstream queueing delay, likely due to bufferbloat, has caused TCP congestion window collapse upon one packet loss. With the help of TCP Timestamps option, I have devised a lightweight passive bandwidth estimation algorithm, allowing us to observe that for 71.26% of the large flows, the bandwidth utilization ratio is below 50%. I find that TCP may not fully utilize the fast-varying available bandwidth when RTT is large in the LTE network. Upon further analysis, I identify 52.61% of all downlink TCP flows have been throttled by TCP receive window and data transfer patterns for some popular applications are both energy and network unfriendly.

We also study the incorrect caching behavior of smartphone applications have incurred wasted network bandwidth usage [11] and about 30% of the total mobile traffic volume can be saved with correct caching and content compression.

Other Research: Computer Security and Computer Vision

Besides cellular network and energy related projects, I've also worked on a variety of projects, such as computer security and computer vision.

We design and implement Souche [14], a system that recognizes legitimate users early in online services by leveraging social connections established over time. Legitimate users help identify other legitimate users through an implicit vouching process, strategically controlled within vouching trees. In our evaluation on a real dataset of several hundred million users, Souche can efficiently identify 85% of legitimate users early, while reducing the percentage of falsely admitted malicious users from 44% to 2.4%. Our evaluation further indicates that Souche is robust in the presence of compromised accounts. It is generally applicable to enhance usability and security for a wide class of online services.

I also design and implement a novel framework, called SocialWatch [9], to detect online service abuse attacks at a large scale. We explore a set of social graph properties, ranging from those that describe individual user behaviors, to those that capture the interactions among users and their social affinities. We evaluate SocialWatch using a large dataset from a major email provider with more than 682 million users and over 5.75 billion directional relationships. SocialWatch successfully detects 56.85 million attacker-created accounts with a low false detection rate of 0.75% and a low false negative rate of 0.61%. In addition, this work also addresses the challenge of identifying hijacked accounts within the legitimate account set through a Bayesian decision framework.

Back in 2007, when I was an undergraduate in Tsinghua University, China, I worked on a project in computer vision. I design and implement a gesture detection system that use human gesture patterns to detect deception [13].

Future Research Agenda

As future directions, I will continue and extend my previous works in the following aspects:

- *Cellular Network Performance Characterization.* I will continue the development *MobiPerf* and make it into a general mobile measurement platform that can benefit the entire research community. In addition, I will make use of the data set collected by *MobiPerf* to make more analysis on cellular network performance and this data set [4] is also available for other researchers.

- *Energy Optimization for Cellular Networks.* I will extend the inter-burst time prediction framework and implement it as a real functioning system on Android and use fast dormancy to save energy for end users. In addition, I will

improve the prediction algorithm to consider TCP-level and application level context, which may further increase the prediction accuracy.

- *Effect of Network Protocol and Application Behavior on Performance for LTE Networks.* I will improve the design of the bandwidth estimation algorithm that we used to quantify the network utilization ratio of existing applications in the LTE networks. I will also explore the predictability of the available bandwidth in the LTE networks and study its implications and applications.

Summary

In summary, I have received extensive training from research projects in the mobile computing and cellular networking space, as well as a few other areas in computer science before and during my Ph.D. years. I have also worked on projects that deal with large-scale data analysis, involving tens of terabytes of data, billions of users, or billions of packets, which require me to debug/optimize my code following industrial standards. More importantly, I am very interested in the projects that I have worked on and I am enthusiastic to take a job as a researcher in the area of mobile computing and networking.

References

[1] 4GTest. `http://mobiperf.com/4g.html`.

[2] FCC Open Internet Apps Challenge. `http://openinternetapps.challenge.gov/`.

[3] MobiPerf (3GTest). `http://mobiperf.com`.

[4] MobiPerf Open Dataset. `http://openmobiledata.appspot.com/`.

[5] MobiPerf Source Repository. `https://github.com/Mobiperf/MobiPerf`.

[6] Smartphone 3G Test (3GTest). `http://www.eecs.umich.edu/3gtest`.

[7] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. A Close Examination of Performance and Power Characteristics of 4G LTE Networks. In *MobiSys*, 2012.

[8] J. Huang, F. Qian, Z. M. Mao, S. Sen, and O. Spatscheck. Screen-Off Traffic Characterization and Optimization in 3G/4G Networks. In *IMC*, 2012.

[9] J. Huang, Y. Xie, F. Yu, Q. Ke, M. Abadi, E. Gillum, and Z. M. Mao. SocialWatch: Detection of Online Service Abuse via Large-Scale Social Graphs. In *ASIACCS*, 2013.

[10] J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing Application Performance Differences on Smartphones. In *MobiSys*, 2010.

[11] F. Qian, K. S. Quah, J. Huang, J. Erman, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. Web Caching on Smartphones: Ideal vs. Reality. In *MobiSys*, 2012.

[12] F. Qian, Z. Wang, Y. Gao, J. Huang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. Periodic Transfers in Mobile Applications: Network-wide Origin, Impact, and Optimization. In *World Wide Web*, 2012.

[13] F. Xia, H. Wang, and J. Huang. Deception Detection Via Blob Motion Pattern Analysis. In *ACII*, 2007.

[14] Y. Xie, F. Yu, Q. Ke, M. Abadi, E. Gillum, K. Vitaldevaria, J. Walter, J. Huang, and Z. M. Mao. Innocent by Association: Early Recognition of Legitimate Users. In *CCS*, 2012.

[15] Q. Xu, J. Huang, Z. Wang, F. Qian, A. Gerber, and Z. M. Mao. Cellular Data Network Infrastructure Characterization and Implication on Mobile Content Placement. In *SIGMETRICS*, 2011.