Meddle: Middleboxes for Increased Transparency and Control of Mobile Traffic

Ashwin Rao
INRIA

Justine Sherry
UC Berkeley

Arnaud Legaut
INRIA

Arvind Krishnamurthy
University of Washington

Walid Dabbous
INRIA

David Choffnes
University of Washington

Categories and Subject Descriptors

C.2 [**Computer Communication Networks**]: General

Keywords

Mobile, Networking, Measurement, Middlebox, Performance.

1. INTRODUCTION

Mobile networks are the most popular, fastest growing and least understood systems in today's Internet ecosystem. Despite a large collection of privacy, policy and performance issues in mobile networks [3, 6] users and researchers are faced with few options to characterize and address them. In this poster we present *Meddle*, a framework aimed at enhancing transparency in mobile networks and providing a platform that enables users (and researchers) control mobile traffic.

In the mobile environment, users are forced to interact with a single operating system tied to their device, generally run closed-source apps that routinely violate user privacy [6], and subscribe to network providers that can (and do) transparently modify, block or otherwise interfere with network traffic [13].

Researchers face a similar set of challenges for characterizing and experimenting with mobile systems. To characterize mobile traffic and design new protocols and services that are better tailored to the mobile environment, we would like a framework that allows us to intercept and potentially modify traffic generated by mobile devices as they move with users, regardless of the device, OS, wireless technology, or carrier. However, implementing this functionality is difficult on mobile devices because it requires warranty-voiding techniques such as jail breaking to access and manipulate traffic at the network layer [3]. Even when using such an approach, carriers may manipulate traffic once it leaves the mobile device [13], thus rendering some research impractical. Furthermore, researchers generally have no ability to deploy solutions and services such as prefetching and security filters, that should be implemented in the network.

In this poster, we present *Meddle*, a framework that combines virtual private networks (VPNs) with middleboxes to provide an experimental platform that aligns the interests of users and researchers. *Meddle* relies on VPN tunnels to access the mobile traffic regardless of the device, OS, wireless technology, and carrier. *Meddle* can thus provide a continuous and comprehensive view of how mobile devices interact with the Internet. Once packets arrive at a *Meddle* server, we use a variety of middlebox approaches to interpose on mobile-device traffic.

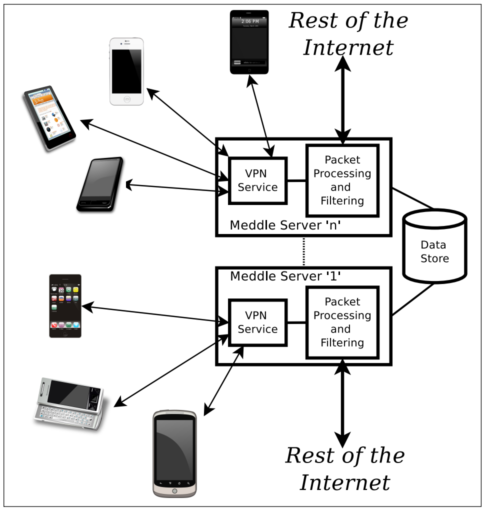

Figure 1: System overview. *Clients use VPNs to connect to a nearby* Meddle *server that interposes traffic.*

Meddle offers new opportunities for measuring and characterizing mobile traffic, and designing new in-network features to improve the mobile experience. For example, by accessing network traffic regardless of the wireless technology we can analyze how different operating systems and apps offload their traffic from cellular networks to Wi-Fi. To improve the user experience, we implement packet filters to block ads; unlike existing packet filters for mobile devices, the packet filters provided by *Meddle* do not require jail-breaking the mobile device. Furthermore, Meddle provides a vantage point for separating mobile-network performance from server-side performance, thus improving bottleneck identification for mobile applications. *Meddle* also enables researchers to investigate what-if scenarios for the impact of new middleboxes as if they were deployed in carrier networks. For example, *Meddle* can be used to deploy anonymization systems such as Privad [5].

2. MEDDLE ARCHITECTURE

Meddle uses VPNs as a portable mechanism to tunnel the data traffic from mobile devices to a machine where users and researchers

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CoNEXT Student'12, December 10, 2012, Nice, France.
Copyright 2012 ACM 978-1-4503-1779-5/12/12 ...$15.00.

Device	**Home (Wi-Fi)**	**Work (Wi-Fi)**	**Mobile (Cell)**	**Volume (MB)**
Android	38.6%	58.7%	2.7%	436 MB
iPhone	47.3%	23.1%	0.7%	105 MB
iPad	92.8%	7.2%	N-A	812 MB

Table 1: Traffic share and volume from three devices when using Wi-fi at home, Wi-Fi at work, or a cellular connection. Meddle *provides a comprehensive view of mobile traffic regardless of the access technology.*

can exert control over network flows. VPNs also reduce the barrier to entry for deploying *Meddle* because Android, BlackBerry, and iOS, which represent more than 86% of the mobile device market [4], have native VPN support. As shown in Figure 1, when a mobile device connects to the Internet, we tunnel its traffic via a nearby *Meddle* server in a similar way to how CDNs use DNS to redirect Web clients to nearby content caches [2]. On each *Meddle* server we implement custom services for users such as packet filtering, caching, and intrusion detection. *Meddle* thus takes two well-known technologies – VPNs and middleboxes – and combines them in unintended ways for the mobile environment.

3. CASE STUDIES

We believe that the research enabled by *Meddle* will form a positive feedback loop in which new, proven research artifacts become additional incentives for user adoption, thus enabling further research. We have a *Meddle* prototype which has been running since August 27, 2012. Our prototype currently serves mobile devices that include an Android phone, an iPhone, an iPad, and an iPod Touch. The particpants in our study are using their own devices to interact with *Meddle*, as part of an IRB-approved study to characterize mobile traffic.

***Meddle* example.** We have implemented a DNS-based filter to block ads, analytics, and mediation sites. We believe that such a filter is an important incentive for a user based study because Vallina-Rodriguez *et al.* [12] observe that ads account for 5% of daily traffic from more than 50% of Android users in a large European ISP. Our ad blocking engine relies on the publicly available list of domains for ads and analytics [1]; we augment this list of domains using the recent research on mobile ads [6, 7]. From our initial deployment, we observed a 0.05% to 0.8% reduction in total traffic at each mobile device due to our ad blocking engine. We are investigating whether this difference from previous work is due to incomplete filters, user-specific behavior or other factors.

Measuring Traffic Offload. In Table 1, we summarize the traffic volume observed from an Android phone, an iPhone, and an iPad, for 15 days since October 16, 2012. The three users generated 1.35 GB of traffic that passed through our *Meddle* server. In Table 1, we observe that the iPhone user consumes more cellular bandwidth compared to the Android user. Further, we note that the vast majority of mobile device traffic is traversing Wi-Fi networks instead of cellular networks. We plan to use this comprehensive view of network traffic to investigate the marginal impact of offloading cellular traffic onto Wi-Fi networks.

Overhead. We observe low overheads in terms of power consumption, data quota, and network latency, when tunneling data traffic via a VPN. We observed a 10% increase in power consumption when streaming an HD video from YouTube to an Android device and an iPhone via one of our *Meddle* servers. We measured the overhead of the tunnel in terms of data overhead from IPsec headers and keep-alive messages, finding that it ranges from 8–12% for an Android phone and an iPhone. Our test traffic included Web searches, interaction on social networks, map searches, playing a game, online shopping, downloading popular apps, emailing and reading the news. To mitigate potential additional network delay from routing traffic through *Meddle* we envision a DONAR-style deployment where users are dynamically redirected to a *Meddle* based on network conditions and server load [14]. We observe that PlanetLab nodes have a latency between 3 ms and 13 ms, with a median of 5 ms from the mobile-network egress points. We used the data collected from 10 mobile phones located throughout the US for this measurement study. Thus, when compared to RTTs of 10s or 100s of milliseconds that exist in mobile networks, the additional latencies from traversing a *Meddle* server is expected to be relatively small or even negligible.

4. FUTURE WORK

We are currently recruiting users for an IRB approved study on the network usage profiles of mobile phone users. To enhance the transparency of mobile networks, we plan to allow users to observe how their installed apps use the network and with whom these apps share (or leak) information; a system similar to Mozilla Collusion [8]. We also plan to test algorithms for content coalescing, caching, prefetching, and offloading the work of processing the DOM to speed up page load times [9, 10, 11]. We are currently working towards making the *Meddle* system publicly available along with a framework to enable researchers contribute to the *Meddle* system and analyze the data collected by the *Meddle* servers.

5. REFERENCES

[1] Ad blocking with ad server hostnames and ip addresses. `http://pgl.yoyo.org/adservers/`.

[2] Akamai. Akamai CDN. `www.akamai.com`.

[3] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In *Proc. of USENIX OSDI*, 2010.

[4] Gartner smart phone marketshare 2012 Q1. `www.gartner.com/it/page.jsp?id=2017015`.

[5] S. Guha, B. Cheng, and P. Francis. Privad: practical privacy in online advertising. In *Proc. of USENIX NSDI*, NSDI'11, pages 13–13, Berkeley, CA, USA, 2011. USENIX Association.

[6] P. Hornyack, S. Han, J. Jung, S. Schechter, and D. Wetherall. These aren't the droids you're looking for: retrofitting android to protect data from imperious applications. In *Proc. of CCS*, 2011.

[7] I. Leontiadis, C. Efstratiou, M. Picone, and C. Mascolo. Don't kill my ads! Balancing Privacy in an Ad-Supported Mobile Application Market. In *Proc. of Hotmobile*. ACM, 2012.

[8] Mozilla collusion. `www.mozilla.org/en-US/collusion/`.

[9] Opera mini browser. `www.opera.com/mobile/features/`.

[10] Amazon silk browser. `www.amazon.com/gp/help/customer/display.html?nodeId=200775440`.

[11] SPDY: An experimental protocol for a faster web. `www.chromium.org/spdy/spdy-whitepaper`.

[12] N. Vallina-rodriguez, J. Shah, A. Finamore, Y. Grunenberger, H. Haddadi, K. Papagiannaki, and J. Crowcroft. Breaking for Commercials : Characterizing Mobile Advertising. *Proc. of IMC*, 2012.

[13] Z. Wang, Z. Qian, Q. Xu, Z. Mao, and M. Zhang. An untold story of middleboxes in cellular networks. In *Proc. of ACM SIGCOMM*, 2011.

[14] P. Wendell, J. W. Jiang, M. J. Freedman, and J. Rexford. Donar: decentralized server selection for cloud services. In *Proc. of ACM SIGCOMM*, 2010.