Haystack: *In Situ* Mobile Traffic Analysis in User Space

Abbas Razaghpanah
Stony Brook University

Narseo Vallina-Rodriguez
ICSI

Srikanth Sundaresan
ICSI

Christian Kreibich
ICSI / Lastline

Phillipa Gill
Stony Brook University

Mark Allman
ICSI

Vern Paxson
ICSI / UC Berkeley

Mobile phones now offer users capabilities that rival those of general purpose computers. However, despite the myriad applications they support, mobile systems remain notoriously opaque. Mobile operating systems tightly control access to system resources and network traffic. The opacity of mobile systems to monitoring presents two serious drawbacks. First, it makes it very hard for researchers to understand performance of mobile systems and applications in the wild, requiring them to rely upon instrumenting individual phones to run experiments [1], privileged access to ISPs [2, 4] and even redirecting traffic through instrumented VPNs [3] to obtain large-scale traffic traces, which often lack local phone context (e.g., app context, device status). Second, it opens the door for potentially unchecked exfiltration of personal data by applications, especially since we rely upon these devices for an increasing number of tasks and entrust them with increasingly sensitive data.

We are developing *Haystack*, a platform for Android devices that provides high-fidelity information about mobile phone operation and network traffic. Figure 1 shows an overview of Haystack's architecture. Our solution combines three key approaches. First, Haystack runs completely in user-space without requiring root permissions, capturing user traffic *locally* on the device. To do so, Haystack takes advantage of Android's VPN permission to route transmitted packets through a process running in user-space. Second, by running locally on the phone Haystack can directly observe crucial app context (*e.g.,* determining the identity of the particular app generating the traffic), device status (*e.g.,* location, screen, and radio state), user-related information (*e.g.,* accounts, messages, recent calls, and contact lists), and the network traffic associated with user activities. Third, Haystack optionally intercepts encrypted traffic via a local TLS proxy. This enables on-device packet inspection to better understand application behaviors (e.g., exfiltration of personal data). The resulting combination of network traffic, app activity, and detailed metadata allows Haystack to provide unprecedented visibility to characterize mobile traffic and performance, assess app security, and detect privacy leaks.

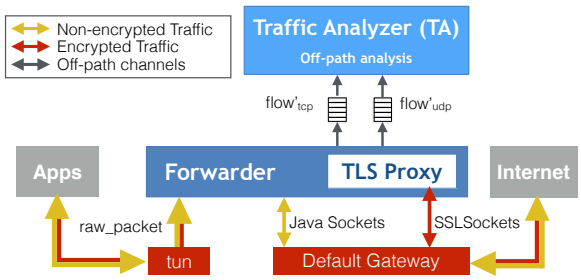

Figure 1: Haystack architecture design, highlighting the different system components and the data forwarding process for an outgoing packet created by a local application. The solid lines represent the actual forwarding path for traffic generated by mobile apps even if encrypted, whereas the dashed lines represent the off-line path used for privacy and performance analysis.

Our results counter the assumption that client-side traffic analysis would be prohibitively resource intensive on mobile devices. We demonstrate that user-space packet forwarding is feasible on Android with 3–9% power overhead and less than 5% CPU overhead. We stress-test Haystack using network-intensive applications and find that it achieves throughput well above that required for modern mobile applications (26–55 Mbps). We have deployed Haystack as an app in the Android market [1] and will demo the app during the poster session.

1. REFERENCES

[1] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin. A first look at traffic on smartphones. In *ACM IMC*, 2010.

[2] P. Gill, V. Erramilli, A. Chaintreau, B. Krishnamurthy, K. Papagiannaki, and P. Rodriguez. Follow the money: Understanding economics of online aggregation and advertising. In *ACM IMC*, 2013.

[3] A. Rao, J. Sherry, A. Legout, A. Krishnamurthy, W. Dabbous, and D. Choffnes. Meddle: Middleboxes for Increased Transparency and Control of Mobile Traffic. In *ACM CoNEXT Student Workshop*, 2012.

[4] N. Vallina-Rodriguez, J. Shah, A. Finamore, Y. Grunenberger, K. Papagiannaki, H. Haddadi, and J. Crowcroft. Breaking for commercials: characterizing mobile advertising. In *ACM IMC*, 2012.

[1] Available for download at: `https://play.google.com/store/apps/details?id=edu.berkeley.icsi.haystack`