Detecting and Localizing End-to-End Performance Degradation for Cellular Data Services

Faraz Ahmed[†], Jeffrey Erman[‡], Zihui Ge[‡], Alex X. Liu[†], Jia Wang[‡], He Yan[‡],
[†]Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA
[‡]AT&T Labs - Research, Bedminster, NJ, USA
{farazah,alexliu}@cse.msu.edu, {erman,gezihui,jiawang,yanhe}@research.att.com

ABSTRACT

Nowadays mobile device (e.g., smartphone) users not only have a high expectation on the availability of the cellular data service, but also increasingly depend on the high end-to-end (E2E) performance of their applications. Since the E2E performance of individual application sessions may vary greatly, depending on factors such as the cellular network condition, the content provider, the type/model of the mobile devices, and the application software, detecting and localizing service performance degradations in a timely manner at large scale is of great value to cellular service providers. In this paper, we build a holistic measurement system that tracks session-level E2E performance metrics along with the service attributes for these factors. Using data collected from a major cellular service provider, we first model the expected E2E service performance with a regression based approach, detect performance degradation conditions based on the time series of fine-grained measurement data, and finally localize the service degradation using association-rule-mining techniques. Our deployment experience reveals that in 80% of the detected problem instances, performance degradation can be attributed to non-network-location specific factors, such as a common content provider, or a set of applications running on certain models of devices.

Categories and Subject Descriptors

C.2.3 [**Computer Communication Networks**]: Network Operations—*Network monitoring*

Keywords

Cellular Network, End-to-end, Performance

1. INTRODUCTION

Cellular data services are increasingly being used from smartphones to tablets in every aspect of everyday life such as emailing, web browsing, video streaming, and online shopping. Nowadays mobile users using smartphones and cellular data services have a high expectation not only on the service availability but also on their application performance, i.e.,

the end-to-end (E2E) service performance. The E2E performance may vary significantly depending on application, content provider, device type, geographic location, time of day and day of week [1–3]. Providing high E2E performance is critical for cellular service providers to best serve their customers and maintain their competitive edge. In this context, E2E performance means the performance that users experience for a specific user location, content provider, device type, and application type.

In this work, we study the detection and localization of E2E performance degradation (such as slow webpage loading and unsmooth video playback) at cellular service providers across four different administrative domains: cellular service providers, content providers, device manufacturers, and application developers. For detection, we want to proactively detect E2E performance degradations before users call the customer service of cellular service providers to complain. For localization, we want to find the causes (user location, content provider, device type, and/or application type) of the E2E performance degradation. Detecting and localizing E2E performance degradation is crucial for cellular service providers, content providers, device manufactures, and application developers to jointly troubleshoot root causes. When users experience E2E performance degradation, they have no clue who is responsible. For example, when a user at New York experiences performance degradation for email application A provided by a content provider C on his device D, he does not know whether it is the cellular provider, the content provider, the device, or his email client app, that is causing the problem. Often when such E2E performance degradation arises, users blankly ascribe all the faults to their cellular service providers and may call customer service centers to complain, which may cause both reputation damage and financial losses for the providers. Furthermore, when cellular service providers receive such calls, it is a lengthy, manual, labor intensive process for them to locate where the issue is, and the actual issue may not be the cellular service provider's problem. The localization findings allow cellular service providers to quickly mitigate service issues, effectively communicate with customers complaining about service problems, and engage content providers, device manufactures, or application developers and operators to jointly troubleshoot for root causes.

2. PROPOSED APPROACH

In this paper, we propose a holistic approach to detecting and localizing E2E performance degradation at cellular service providers across the four dimensions of user locations, content providers, device types, and application types. We
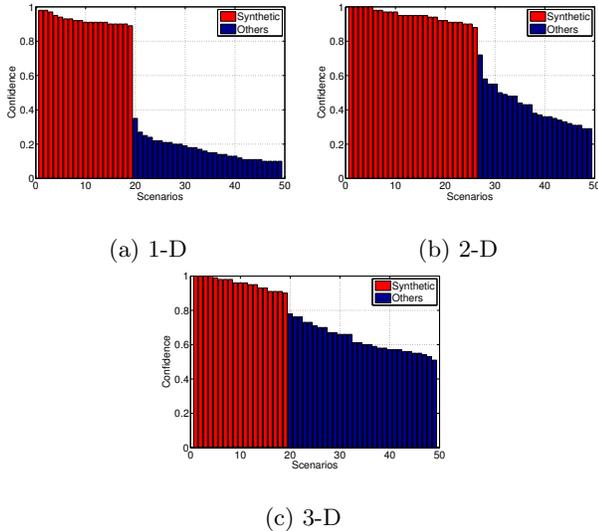
(a) 1-D   (b) 2-D

(c) 3-D

Figure 1: Synthetic anomaly detection

utilize anonymized flow level data collected from the core network of a major cellular service provider in the United States to compute TCP loss ratio and Round Trip Time (RTT) as E2E performance metrics. For each TCP flow, we capture information such as standard coordinated universal time (UTC), the serving RNC that describes the user access point (L), the handheld device type (D), the application type (A), and the content provider being accessed (P). First, we use training data to build regression based models that capture the normal performance of every *E2E instance*, which means the flows corresponding to a user location, content provider, device type, and application type. We use robust regression for this baseline modeling because it can minimize the impact of extreme outlier data points on the produced model. Specifically, we represent E2E performance in a four dimensional matrix that is called *E2E matrix* and is denoted $E_I = [1..L, 1..P, 1..D, 1..A]$. We build a baseline model for all E2E instances, identify the E2E instance groups that have significantly different performance, and then model these groups separately, leaving the rest of E2E instances still being modeled by the baseline model. Second, we use our models to detect performance degradation for each E2E instance on an hourly basis. For each E2E instance, if the real performance in the testing phase is significantly much worse than the expected performance based on our models, we label it as degrading. Third, after each E2E instance has been labeled as non-degrading or degrading, we use association rule mining to localize the source of performance degradation. For example, rule `Device − D, Email → degrading` shows that for a particular hour, for all locations and content providers, the cellular users of Device-D are experiencing significantly degraded performance for their email application.

3. EVALUATION

For evaluation purposes, we created three different synthetic sets of scenarios. The first set consists of 19 scenarios involving only one of the four dimensions, the second set involves 26 scenarios with two dimensions and the third set involves 19 scenarios with three dimensions. Each scenario corresponds to a one-hour anomaly associated with a group

of instances in the E2E performance matrix. Figure 1 shows that all 64 synthetic anomaly scenarios that we injected into the collected operational data are successfully detected and localized. For one dimensional scenarios, we obtain all scenarios as association rules with more than 90% confidence. For two dimensional scenarios we achieved 0 false positives for a minimum confidence threshold of 70% and for three dimensional scenarios we got 0 false positives for a minimum confidence threshold of 80%.

We implemented and deployed our detection and localization tool on an operational network and separately examined performance anomalies for packet loss ratio and RTT. We monitored 78 different user locations, 51 device types, 13 application types and 36 content providers for a period of two weeks. We did not observe performance anomalies in the user location dimension, which means that a single user location cannot be blamed for the anomalies occurring during the one week time period. We observe 16 and 27 one dimensional anomalous events for content providers and device types respectively, which is small when compared to anomalous events in the application domain. It is highly unlikely that the performance of a content provider will degrade across all user locations. This is because content providers utilize content distribution networks for serving content at different user locations. Therefore, a problematic content server can only affect the performance of a subset of user locations. For applications, we observe 180 single dimension anomalies. This could be attributed to application patches and bug fixes that are released periodically and frequently. Software bugs in these updates can cause performance issues across all user locations, device types and content providers.

For multi-dimensional cases, we detected a total of 8415 E2E instances as anomalous. In 83% of these instances, a content provider was involved, 86% involved device types, and 86% involved application types. Overall, 80% of these instances do not have location as one of the dimensions. In other words, 80% of the time the anomalies are not due to issues in the cellular network. This composition indicates that most of the time content providers, device types and applications are involved in the performance degradations. These results are extremely useful for network operators as they point network operators to the problematic dimensions and reduce the search space of root causes.

4. REFERENCES

[1] 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 *Proceedings of the 10th International Conference on MobiSys*, pages 225–238, 2012.

[2] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck. An in-depth study of lte: Effect of network protocol and application behavior on performance. In *Proceedings of the ACM SIGCOMM*, pages 363–374, 2013.

[3] M. Z. Shafiq, L. Ji, A. X. Liu, and J. Wang. Characterizing and modeling internet traffic dynamics of cellular devices. In *Proceedings of the ACM SIGMETRICS*, pages 305–316, San Jose, California, June 2011.