Non-cooperative Diagnosis of Submarine Cable Faults

Edmond W. W. Chan, Xiapu Luo, Waiting W. T. Fok,
Weichao Li, and Rocky K. C. Chang

Department of Computing, The Hong Kong Polytechnic University
{cswwchan|csxluo|cswtfok|csweicli|csrchang}@comp.polyu.edu.hk

Abstract. Submarine cable faults are not uncommon events in the Internet today. However, their impacts on end-to-end path quality have received almost no attention. In this paper, we report path-quality measurement results for a recent SEA-ME-WE 4 cable fault in 2010. Our measurement methodology captures the path-quality degradation due to the cable fault, in terms of delay, asymmetric packet losses, and correlation between loss and delay. We further leverage traceroute data to infer the root causes of the performance degradation.

1 Introduction

Submarine cables are critical elements of the Internet today, because they provide cross-country routes for transoceanic data and voice transmissions. The demand for high-capacity submarine cables has been increasing for the last few years. For instance, the recently deployed Trans-Pacific Unity submarine cable system can transmit data between Japan and the west coast of the United States up to 4.8 Terabits per second (Tbits/s). Dramatic capacity upgrades to the existing Asia-Europe cable systems and the emergence of five new submarine cable systems connecting the Middle East were also reported [11].

Data loss and substantial service interruption as a result of submarine cable faults conceivably entail huge economic cost. Although submarine cable systems are protected by various reliability technologies (e.g., [15, 16]), they still appear to be highly vulnerable according to numerous submarine cable faults reported in recent years (e.g., [1–3]). The worst one is the incident of massive cable cuts due to the Hengchun earthquake in 2006 [1]. Moreover, a submarine cable fault requires considerable time for tracing the fault location and repairing.

Besides the traffic on the faulty submarine cable, the Internet traffic that is not carried by the faulty cables can also be affected. A common quick-fix strategy for restoring the disrupted communication is to reroute the affected traffic to other submarine/terrestrial/satellite links. However, the side effect of such ad hoc traffic rerouting mechanism is introducing a high volume of traffic, and therefore substantial congestion, to the backup paths. However, the impact of submarine cable faults on the global Internet connectivity has not received attention from the research community. Therefore, very little is known about the Internet's vulnerability to the faults in terms of path-quality degradation, congestion on the backup paths, and speed of network recovery.

In this paper, we report the impacts of a recent SEA-ME-WE 4 cable fault incident [2] measured from our neighbor-cooperative measurement system [14]. In this system, a number of coordinated measurement nodes persistently monitor the performance of network paths to a set of web servers. The impacts of the cable fault are observed from the degradation in the path quality. To infer the root cause of the degradation, we leverage the forward-path Tcptraceroute gleaned from the measurement nodes to study the IP-level/AS-level route changes. Based on this dataset, we analyze how submarine cable faults affected the routes used by the network paths and the performance of these paths. We also evaluate the effectiveness of network operators' responses to the incident.

The paper is organized as follows. We first introduce our measurement methodology in §2. We then present our measurement findings on the impacts of the SEA-ME-WE 4 cable fault in §3. After discussing the related works in §4, we conclude the paper with future works in §5.

2 Measurement methodology

2.1 Measurement setup

We have been conducting end-to-end Internet path measurement from eight Hong Kong universities, labeled by UA–UH, since 1 January 2009. A measurement node is installed just behind the border router of each university to measure network paths to 44 non-cooperative web servers (without requiring software setup on the servers) in Hong Kong, Australia, China, Finland, France, Germany, Japan, Korea, New Zealand, Taiwan, the United Kingdom, and the United States. We use HTTP/OneProbe [17] for data-path quality measurement and Tcptraceroute for forward-path tracing. Our measurement produces 12-GB measurement data daily.

2.2 Measurement scheduling and traffic

To obtain comparable results, all the eight measurement nodes measure the same web server around the same time. We employ several measures to avoid congestion introduced by the measurement traffic. In particular, we divide the set of web servers into five groups and measure the groups in a round-robin fashion. The nodes launch HTTP/OneProbe to measure each group for one minute and then perform Tcptraceroute with the default configuration to the same group for another minute. For each path, HTTP/OneProbe dispatches a sequence of Poisson-modulated probe pairs to each web server with a probing frequency of 2 Hz and an IP packet size of 576 bytes, and each probe pair elicits at most two 576-byte response packets from the server. Therefore, the aggregated probe traffic sent to each server is less than 200 Kbits/s. Moreover, we use separate network interfaces for conducting the measurement and receiving the data.

2.3 Metrics

Routing metrics To evaluate the routing behavior as a consequence of submarine cable faults, we continuously measure both IP routes and the corresponding

AS routes (by resolving IP hops into AS numbers) from the measurement nodes to the web servers. To quantify the IP-level route changes, we resort to the IP-level Jaccard distance defined in Eqn. (1) to measure the difference of a route measured at times $i-1$ and i, which are denoted by R_{i-1} and R_i. The Jaccard distance is computed by the number of dissimilar elements divided by the total number of distinct elements in R_{i-1} and R_i. Therefore, the IP-level Jaccard distance is zero for two identical IP routes, and one for two completely different IP routes. We similarly compute an AS-level Jaccard distance for AS routes based on Eqn. (1) to analyze the AS-level route changes.

$$J_\delta(R_{i-1}, R_i) = 1 - \frac{|R_{i-1} \cap R_i|}{|R_{i-1} \cup R_i|}. \tag{1}$$

Using Jaccard distance to characterize route changes is not new. Pathak et al. [18], for example, studied the AS routing asymmetry by computing the Jaccard similarity index between forward-path and reverse-path AS routes. Schwartz et al. [19] used the Levenshtein distance to quantify the difference between the dominant route and other non-dominant routes for a pair of source and destination. Since reordering of elements in the IP/AS routes after route changes is rare in our dataset, we simply use Jaccard distance which does not take into account the order of elements in each route, whereas the Levenshtein distance does.

Path performance metrics We employ HTTP/OneProbe to measure TCP data-path performance for each path between measurement node and web server. HTTP/OneProbe uses legitimate TCP data probe and response packets to measure RTTs and detect one-way (i.e., forward-path and reverse-path) packet losses. To evaluate the paths' congestion status, we also apply the loss-pair analysis [13] to correlate the one-way packet losses with the RTTs. Moreover, we correlate the route change metrics with the path performance metrics to analyze path-quality degradation due to submarine cable faults.

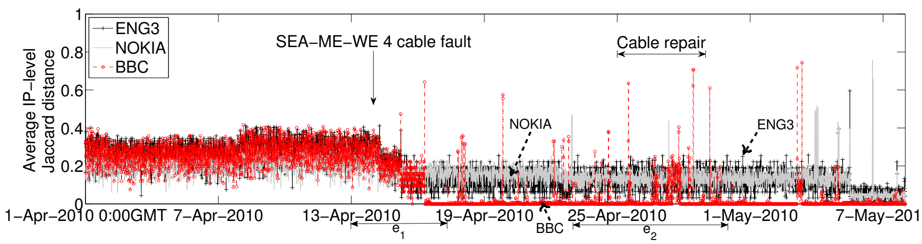
3 The SEA-ME-WE 4 cable fault

The South East Asia-Middle East-Western Europe 4 (SEA-ME-WE 4) submarine cable [8] is a major Internet backbone connecting Southeast Asia, the Indian subcontinent, the Middle East, and Europe. It involves 17 landing points and carries Internet traffic among 15 countries, including Egypt, France, India, Saudi Arabia, and Singapore. The SEA-ME-WE 4 cable has a data rate up to 1.28 Tbits/s [4] and is owned by a consortium of 16 companies, including the Tata Communications (or TATA).

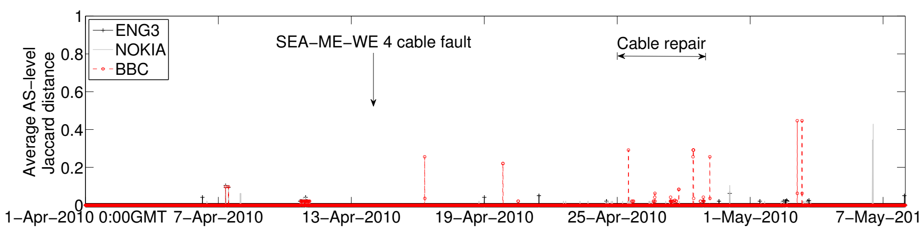
The SEA-ME-WE 4 cable encountered a shunt fault on the segment between Alexandria and Marseille on 14 April 2010 [2, 9], but the exact time was not reported. The shunt fault was caused by a short circuit when the submarine cable, whose insulation was damaged, came into contact with the sea water. Since the cable was not severed, it was still operable with limited capacity. The cable fault affected a number of countries whose global connectivity relied on the SEA-ME-WE 4 cable (e.g., [2, 12]). The repair was started on 25 April 2010, and it took four days to complete [2]. During the repair, the service for the westbound traffic to Europe was not available.

3.1 Impacts of the cable fault

Fig. 1 shows the time series of the average IP-level and AS-level Jaccard distances for the paths from UA–UH to two web servers (BBC and ENG3) in the United Kingdom and one web server (NOKIA) in Finland between 1 April 2010 0:00 and 8 May 2010 0:00 GMT, inclusively. As Fig. 1(a) shows, the IP-level Jaccard distance for the paths overlapped with one another at the beginning and then gradually declined starting from 14 April that coincides with the date of the cable fault incident. The BBC's Jaccard distance dropped to zero with intermittent surges after 16 April 7:30 GMT, whereas the ENG3's and NOKIA's distances fluctuated between 0.05 and 0.22, and experienced another drop on 5 May noon GMT. Moreover, Fig. 1(b) shows some significant AS-level route changes.



(a) IP-level Jaccard distances.



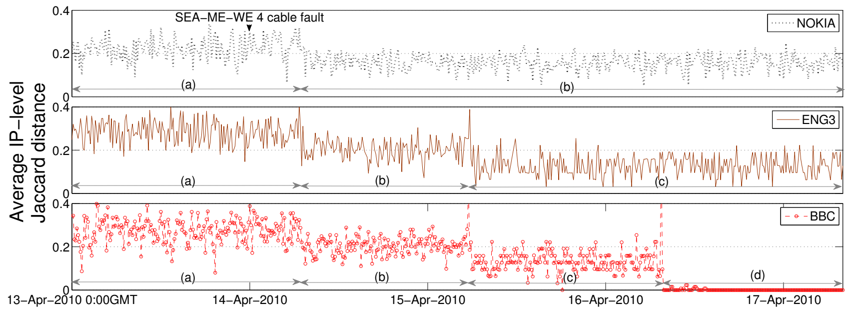
(b) AS-level Jaccard distances.

Fig. 1. Time series of the average IP-level and AS-level Jaccard distances for the paths to BBC, ENG3, and NOKIA.

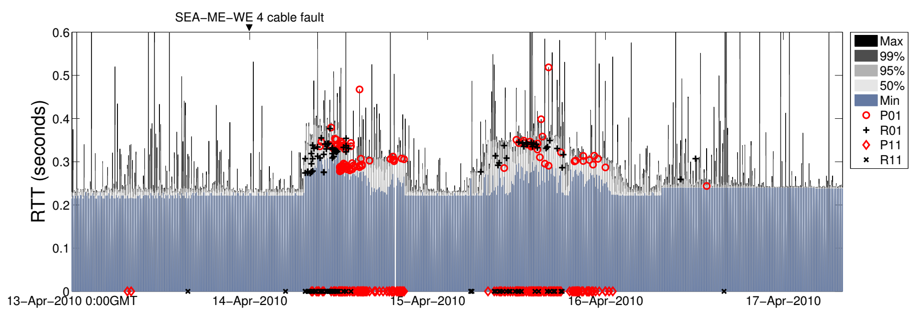
To probe deeper into how the paths to BBC, ENG3, and NOKIA evolved after the cable fault, we zoom into an episode e_1 which spans between 13 April 0:00 and 17 April 8:00 GMT in Fig. 1(a). Fig. 2(a) shows that their average IP-level Jaccard distances during e_1 exhibit staircase decreasing patterns, meaning that the paths became more similar after the cable fault. We can also distinguish at most four distinct phases labeled with (a)–(d) for the NOKIA, ENG3, and BBC paths which have two, three, and four phases, respectively.

A traceroute analysis reveals the subpaths corresponding to the four phases shown in Fig. 3. To generate the figures, we resolved the IP hops' locations based on their DNS names and grouped all the hops with the same location together.

The node labeled with “Unresolved” is located in Hong Kong, and we could not resolve its DNS name. Phases (a)-(c) apply to all three web servers, and all the routes went through the London IX (LINX) via the FLAG network (AS15412). On the other hand, phases (d)-(e) apply only to the BBC paths, and TATA was the carrier. We will discuss phase (e) in the next section.



(a) Time series for the average IP-level Jaccard distance during e_1.



(b) RTT time series for the path between UB and BBC during e_1.

Fig. 2. Time series of the average IP-level Jaccard distance for the paths to NOKIA, ENG3, and BBC, and time series of RTT for UB⇌BBC during e_1.

Phases (a) and (b) Fig. 3(a) shows three subpaths inside the FLAG network in phase (a). Upon the onset of phase (b) on 14 April 7:00 GMT (the same day of the reported cable fault), the IP-level Jaccard distance started declining, a result of the missing subpath via Mumbai as shown in Fig. 3(b). We also plot the RTT time series in Fig. 2(b) for the path between UB and BBC (denoted as UB⇌BBC). The figure includes the RTTs obtained from P01s and R01s which are the respective loss pairs on UB → BBC (forward path) and BBC → UB (reverse path). A P01 (R01) is a packet pair in which *only* the first probe (response) packet is received by the destination, and the first packet’s RTT can be used to infer a congested router’s queueing delay upon packet loss on the forward (reverse) path [13]. Moreover, P11s (R11s) on the x-axis in Fig. 2(b)

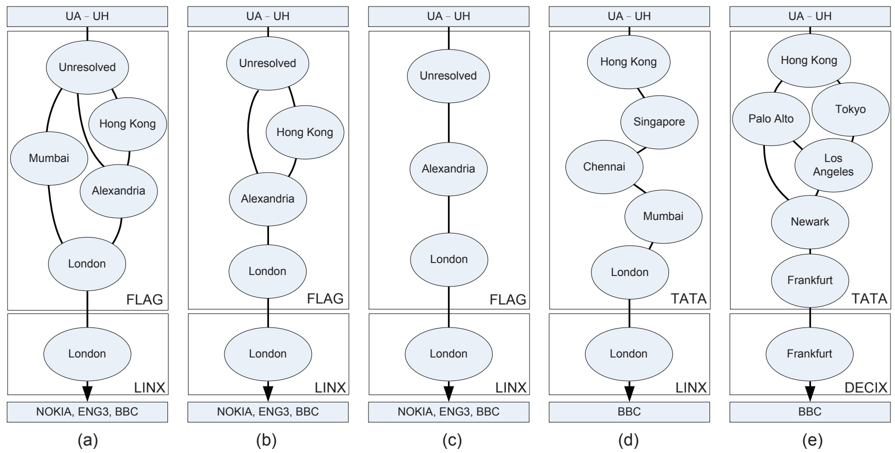


Fig. 3. Five sets of subpaths observed from the NOKIA, ENG3, and BBC paths.

show the RTTs when *both* packets in a probe (response) pair are lost. We align the x-axes in Figs. 2(a) and 2(b) to facilitate a clear comparison.

Fig. 2(b) shows that UB⇌BBC suffered from significant congestion in phase (b). We also observe similar results for the other BBC paths and the NOKIA and ENG3 paths which are not shown in this paper. Comparing with phase (a), phase (b) exhibits both RTT inflation and more loss pairs with the measured path queueing delay [13] between 34 ms and 228 ms. In many cases, both packets in a probe pair or a response pair were lost. The figure also shows a prolonged congestion period in the forward path, indicated by persistent probe packet losses. However, the path performance improved in the second half of the phase which corresponds to the non-working hours in the United Kingdom.

Phases (c) and (d) Fig. 2(a) shows a further reduction of the IP-level Jaccard distance for the ENG3 and BBC paths on 15 April 5:40 GMT (i.e., the onset of phase (c)), because only the subpath via Alexandria and London was retained in FLAG (as shown in Fig. 3(c)). Moreover, Fig. 2(b) shows more prolonged RTT inflation and packet losses during phase (c), which was probably caused by the reduced alternate routes.

On 16 April 7:30 GMT, the beginning of phase (d), the service provider for UA–UH changed the upstream from FLAG to TATA (AS6453) only for the BBC paths. As a result, the IP-level Jaccard distance shown in Fig. 2(a) dropped to almost zero. We also observe a spike from the AS-level Jaccard distance for the BBC paths at the similar time in Fig. 1(b). Notice that this change significantly improved the performance for the BBC paths. In particular, Fig. 2(b) shows that UB⇌BBC enjoyed relatively stable RTTs and insignificant packet losses (and similarly for the other BBC paths), whereas the NOKIA and ENG3 paths still suffered from severe congestion in this phase.

Discussion Fig. 4 shows all the submarine cables available to FLAG and TATA for connecting the IP hops' locations in Figs. 3(a)–3(d). We generate the figures by inspecting the cables and landing points in the cable maps of FLAG [6] and TATA [5]. Fig. 4(b) shows that TATA uses only the SEA-ME-WE 4 cable to reach Singapore, Chennai, and Mumbai, but these segments were not affected by the shunt fault occurred in the Mediterranean segment [9]. Moreover, TATA uses different cables between Mumbai and London. On the other hand, Fig. 4(a) shows that FLAG does not use the SEA-ME-WE 4 cable for forwarding traffic from Hong Kong to the three web servers.

Based on Fig. 4, a plausible explanation for the congestion in the FLAG network in phase (b) is taking on rerouted traffic from the SEA-ME-WE 4 cable after the cable fault. Both FEA and SEA-ME-WE 4 (and SEA-ME-WE 3) are the major submarine cables connecting between Europe and Asia. Fig. 4(b) shows that TATA could use FEA to reach BBC when the SEA-ME-WE 4 segment in the Mediterranean region was not available. Therefore, the congestion was introduced as a secondary effect of the cable fault. On the other hand, the path quality for the BBC paths improved after switching to TATA in phase (d). Unlike FLAG, TATA has access to three submarine cables between Suez and Alexandria. There are also two cables between Alexandria and London. Moreover, the reduced path diversity from phase (a) to phase (c) could also be responsible for the congestion in the FLAG network, although the reason for the reduction is unknown to us.

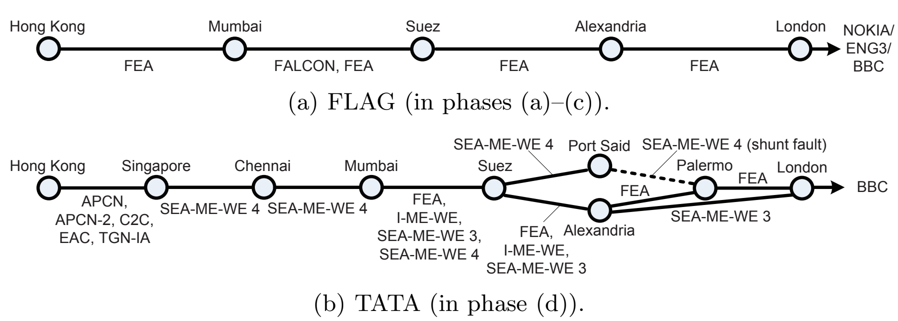
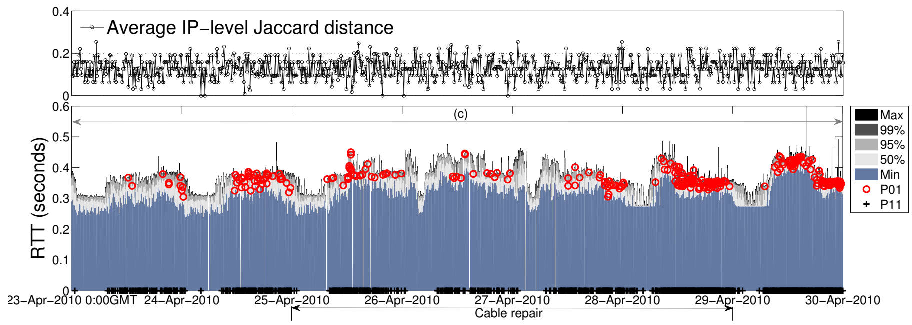


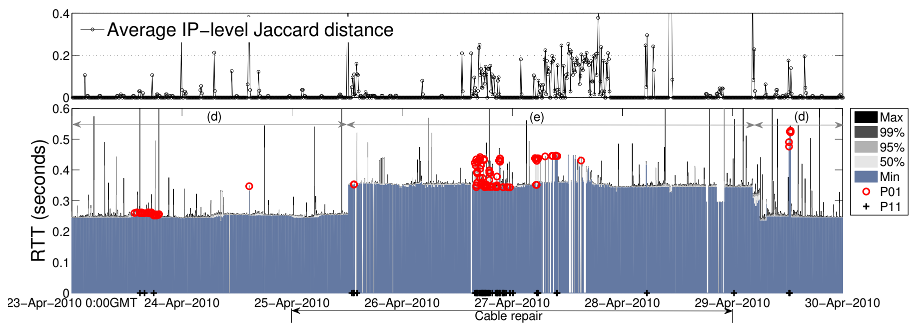
Fig. 4. The submarine cables available to FLAG and TATA for connecting the IP hops' locations in Figs. 3(a)–3(d).

3.2 Impacts of the cable repair

In this section, we analyze the impact of the four-day (25–28 April 2010) repair of the SEA-ME-WE 4 cable on the routing behavior and path performance. Figs. 5(a) and 5(b) show the respective time series of the average IP-level Jaccard distance (at the top of each figure) for the ENG3 and BBC paths and the RTTs for UB⇌ENG3 and UB⇌BBC between 23–30 April which is labeled as the second episode (e_2) in Fig. 1(a). We do not include the time series for the NOKIA paths, because they are similar to ENG3's. Moreover, the path performance for the other measuring nodes to ENG3 (BBC) also resembles the performance given in Fig. 5(a) (5(b)). To correlate the forward-path routing behavior with the path performance, each figure only shows the loss pairs and both-packet-loss events (i.e., `P01` and `P11`) observed from the forward paths. Note that the ENG3 paths remained in phase (c) during the entire period, whereas the BBC paths switched from phase (d) to phase (e), which involves a significant route change, and then back to phase (d).



(a) ENG3.



(b) BBC.

Fig. 5. Time series of the average IP-level Jaccard distance for the paths to ENG3 and BBC, and the time series of RTT for UB⇌ENG3 and UB⇌BBC during e_2.

Fig. 5(a) shows that the ENG3 (and also NOKIA) paths suffered from diurnal congestion in e_2. Traceroutes show that the NOKIA and ENG3 paths still went through the FLAG subpaths shown in Fig. 3(b) and 3(c), respectively, for the entire episode. As a result, both the RTT and Jaccard distance time series exhibit similar patterns as in phases (b)–(c) of Fig. 2. The paths also encountered more severe congestion since 25 April when the SEA-ME-WE 4 cable's repair process began. It is thus likely that the FLAG subpaths were further utilized by other affected parties as alternate routes during the repair process. However, FLAG's network operators did not seem to respond to the degraded path performance until they switched to eastbound routes on 5 May noon GMT (which is shown in Fig. 1(a)), and the path performance was subsequently improved.

Fig. 5(b), on the other hand, shows that the BBC paths were quite good during e_2 except for sporadic packet losses and routing instability. In particular, at the beginning of the episode, the paths still went through TATA which routed the BBC traffic via the subpath given in Fig. 3(d). Probably due to the interruption caused by the repair work [2], TATA rerouted the traffic to another set of subpaths with longer RTTs on 25 April 13:30 GMT, and we refer to this period as phase (e). We can also see a positive correlation between fluctuation in the IP-level Jaccard distance and significant forward-path packet losses during phase (e). TATA finally restored the subpath in Fig. 3(d) on 29 April 6:00 GMT (which is close to the completion time of the repair), and therefore the path performance returned to the level observed from the beginning of the episode.

4 Related work

RIPE NCC [10] reported a longitudinal study of two cable cuts in the Mediterranean region in 2008 based on its routing information (RIS), test traffic measurements (TTM), and DNS monitoring (DNSMON) services. The study showed that the affected networks involved frequent rerouting in BGP, significant network congestion, and increased latencies. In our study, we mainly use end-to-end path measurement and IP traceroute to study the impacts on the paths under our monitoring. Based on a set of measuring points, we are able to infer that the path congestion was due to the secondary effect of the cable fault, which has not been reported in previous studies.

Renesys [7] also reported a few studies on the impacts of submarine cable faults based on BGP routes and RTTs (measured by traceroute) obtained from its data collection infrastructure. Comparing with their analysis on the same SEA-ME-WE 4 cable fault [12], our methodology uses TCP data packets to measure the data-path performance, instead of ICMP packets that can be processed by different paths in the routers and thus produce biased measurement. Therefore, our measurement observed quite stable RTTs for the paths via the TATA network, whereas Renesys observed significant RTT fluctuation from the TATA network in the similar time period. Besides, our analysis also obtains useful packet loss information that was not considered in their analysis.

5 Conclusion and future work

In this paper, we employed non-cooperative path measurement to study the impacts of a recent submarine cable fault on the Internet connectivity and end-to-end path performance. With only eight measurement nodes, we showed that the non-cooperative methods (HTTP/OneProbe and traceroute) could facilitate an in-depth impact analysis of a cable fault occurred thousands miles away. In particular, our analysis revealed that a cable fault could significantly impact on Internet traffic on other non-faulty paths. Moreover, network operators did not always take immediate action to resolve the performance degradation problem as a result of the cable fault.

As an ongoing work, we will report our impact analysis of other submarine cable faults, such as a SEACOM cable fault in Africa in July 2010. We will also devise new algorithms based on non-cooperative path measurement to promptly identify and respond to path-quality degradation as a result of cable faults.

Acknowledgments

We thank the four reviewers for their comments and Aleksandar Kuzmanovic for shepherding our paper. This work is partially supported by a grant (ref. no. ITS/355/09) from the Innovation Technology Fund in Hong Kong and a grant (ref. no. H-ZL17) from the Joint Universities Computer Centre of Hong Kong.

References

1. Asia communications hit by quake. http://news.bbc.co.uk/2/hi/asia-pacific/6211451.stm.
2. Beyond the SEACOM network disruptions. http://www.seacomblog.com/team-seacom/2010/05/beyond-seacom-network-disruptions-april-2010.
3. Cable outage - 5th July 2010. http://www.seacom.mu/news/news_details.asp?iID=142.
4. Fujitsu completes construction of SEA-ME-WE 4 submarine cable network. http://www.fujitsu.com/global/news/pr/archives/month/2005/20051213-01.html.
5. Global footprint map - Tata Communications. http://tatacommunications.com/map/gfp.html.
6. Reliance Globalcom. http://www.relianceglobalcom.com/RGCOM_CoverageMap.html.
7. Renesys - presentations & reports. http://www.renesys.com/tech/presentations/.
8. SEA-ME-WE 4. http://www.seamewe4.com/.
9. SEA-ME-WE 4 maintenance finally over. http://sushantwagle.wordpress.com/2010/04/30/se-me-we-4-maintenance-finally-over/.
10. Mediterranean fibre cable cut - a RIPE NCC analysis. http://www.ripe.net/projects/reports/2008cable-cut/index.html, Apr. 2008.
11. Middle East bandwidth pricing report. http://www.telegeography.com/product-info/pricingdb/download/bpr-2009-10.pdf, Oct. 2009.
12. Bahrain's Internet ecosystem revisited. http://renesys.com/tech/reports/Renesys-Bahrain-July2010.pdf, Jul. 2010.
13. E. Chan, X. Luo, W. Li, W. Fok, and R. Chang. Measurement of loss pairs in network paths. In *Proc. ACM/USENIX IMC*, 2010.
14. R. Chang, W. Fok, W. Li, E. Chan, and X. Luo. Neighbor-Cooperative measurement of network path quality. In *Proc. IEEE Globecom*, 2010.
15. S. Harasawa, M. Sumitani, and K. Ohta. Reliability technology for submarine repeaters. *Fujitsu Scientific & Technical Journal (FSTJ): Quality Assurance*, 44(2):148–155, 2008.
16. K. Hodge, J. Vinson, N. Haigh, and I. Knight. Reliability of optical fibres: impact on cable design. In *IEE Colloquium on Reliability of Fibre Optic Cable Systems*, pages 1/1 –1/6, Jan. 1994.
17. X. Luo, E. Chan, and R. Chang. Design and implementation of TCP data probes for reliable and metric-rich network path monitoring. In *Proc. USENIX Annual Tech. Conf.*, 2009.
18. A. Pathak, H. Pucha, Y. Zhang, Y. Hu, and Z. Mao. A measurement study of Internet delay asymmetry. In *Proc. PAM*, 2008.
19. Y. Schwartz, Y. Shavitt, and U. Weinsberg. A measurement study of the origins of end-to-end delay variations. In *Proc. PAM*, 2010.