What have I learned from the IM submission?

Rocky K. C. Chang

January 25, 2013

MonoScope: Automated Network Faults Diagnosis Based on Active Measurements

- Submitted the paper.
- "Another is that we should make the claim more accurate. So far we have not automated the diagnosis yet. But we have shown is that it could be automated by obtaining data from different data-plane and control-plane sources." (My comment)
- Reviews received and rebuttal started
  - "Is this really an 'automatic' system? They claim that the system detects and diagnoses network faults ..., but it would appears that the system only aids, ... Is this true?"
  - "It is unclear which part of MonoScope is 'automatic'?"

The co-authors' contributions to the rebuttal process

- "The reviewers care more on whether our system is automatically (means that it can diagnose network fault without human intervention and in real-time) or manually (analyze the data after collecting a set of data). I think we need to state more clearly and response the questions directly."

- "If we didn't consider/fulfill the aforementioned real-time computational requirements, we may not need to claim that our analysis is in real-time, but in automatic manner. I think it's okay because the target of our work is to reduce the work needed to by conducted by admins through an automatic diagnosis platform."

- "Which part is automatic? Three components are automated including event detection, grouping simultaneous events, and estimating the scale."

Our final replies

- "Is the analysis done in real-time or off-line? MonoScope periodically retrieves data from a database fed with real-time data from our active measurement system, and performs analysis automatically. In the paper we used historical data for demonstrating the effectiveness of MonoScope.

- How much human analysis is required? Human analysis is needed when training the model and validating the result.

- Which part is automatic? 3 components are automated, including event detection, grouping simultaneous events, and estimating the scale."

Replies from the reviewers

- "After reading the rebuttal, I am satisfied by some of the answers given there. I strongly recommend that the authors modify their paper based on the reviewers' comments and their responses, since the same questions are likely to arise in the mind of an ordinary reader."

- "We are delighted to inform you that your paper … has been accepted to the regular Technical track of the IM 2013. Congratulations! All submitted papers underwent a rigorous review process with a minimum of 3, but usually 4 reviews for each paper."

Wait, are they really true?

- "Is the analysis done in real-time or off-line? MonoScope periodically retrieves data from a database fed with real-time data from our active measurement system, and performs analysis automatically. In the paper we used historical data for demonstrating the effectiveness of MonoScope.

- How much human analysis is required? Human analysis is needed when training the model and validating the result.

- Which part is automatic? 3 components are automated, including event detection, grouping simultaneous events, and estimating the scale."

MonoScope: Automat**ed** Network Faults Diagnosis Based on Active Measurements

- Network faults could occur anywhere in a network path, from local network, Internet Service Provider, network carriers, to the remote network. There are also many different types of faults, such as submarine cable outages and misconfiguration. Continuous active measurement at distributed end-systems can help diagnose network faults. In principle, a larger-scale deployment can increase the chance of detecting faults and finding the root cause. In practice, however, the resulted large volume of measurement data makes it very difficult to diagnose problems effectively.

- In this paper, we propose MonoScope to **automatically** detect and diagnose network faults based on large-scale active measurement from end-systems. MonoScope analyzes multiple network path metrics, including round-trip delay, packet loss and reordering rate, to **automatically detect** network events. These metrics are then synthesized to diagnose the origin of network faults and estimate the scale of events. We use three real, but different types of, network faults to illustrate the operation of MonoScope. We demonstrate that the **automated processes** can significantly reduce the effort of diagnosing network faults.

MonoScope: Automat**ing** Network Faults Diagnosis Based on Active Measurements

- Network faults, such as router failures, cable outages, and configuration errors, could seriously affect network performance. In this paper, we use **network faults** very loosely, including those that will yield unfavorable end-to-end network performance, such as packet reordering and suboptimal routes. Diagnosing network faults on end-to-end paths is a very challenging problem, because it generally involves other domains. Even if it can be done, the process is very time consuming, because multiple sources of data, which are scattered in different places, are needed for such diagnosis.

- In this paper, we consider **the problem of making the fault diagnosis as automatic as possible.** Based on coordinated active measurement from a set of end systems, we propose a **procedure** of detecting network faults and identifying their locations. Although this procedure **cannot be fully automated for the time being**, we show that some of the components **could be automated**, and we are automating them in a **preliminary** system called MonoScope. We demonstrate the efficacy of this procedure through several real-world cases that we have encountered in our four years of network monitoring experience.

What is the problem?

- Mix up "what could/can be done" with "what has been done."
- Failed to identify the key contribution at each stage of research
- Temptation of publication
  - to entice or allure to do something often regarded as unwise, wrong, or immoral.
  - to attract, appeal strongly to, or invite
- Temptation of getting ahead of others

What is the cure?

- (Constantly) Re-discover our motivation for doing research.
  - For ourselves or for others?
- (Constantly) Re-visit the research ethics ([Research Ethics and CS](#))
  - "The CS bears a great responsibility to the world to conduct and report their research in an ethical manner."
- Allow time to write and reflect.
- Take every single feedback seriously.

From now on, for the measurement papers

- A complete draft for a top journal/conference submission must be ready **at least a month** before the deadline.

- A complete draft for a not-so-top journal/conference submission must be ready **at least half a month** before the deadline.

- Each co-author is responsible for proof-reading the manuscript critically.

- I will stop any submission that contains unsubstantiated claims.

If we have no integrity, we have nothing left.