

Juggling the Jigsaw: Towards Automated Problem Inference from Network Trouble Tickets

Poster + Demo Submission for an NSDI'13 referred paper

Rahul Potharaju
Purdue University

Student Author

Navendu Jain
Microsoft Research

Cristina Nita-Rotaru
Purdue University

1 Introduction

This demo will present NetSieve [1], a system to do automated problem inference from network trouble tickets. Trouble tickets are diaries comprising fixed fields and free-form text written by operators while troubleshooting a problem. Unfortunately, while they carry valuable information for network management, analyzing them to do problem inference is extremely difficult—fixed fields are often inaccurate or incomplete, and the free-form text is mostly written in natural text. To achieve this, NetSieve combines statistical natural language processing (NLP), knowledge representation and ontology modeling to infer three key features from network trouble tickets:

1. Problem symptoms indicating what problem occurred,
2. Troubleshooting activities describing the diagnostic steps taken by an operator,
3. Resolution actions denoting the fix applied to mitigate the problem.

To cope with ambiguity in free-form text and to adapt to network changes, NetSieve leverages learning from human guidance to improve its inference accuracy.

2 NetSieve Overview

NetSieve operates in three phases. First, the knowledge building phase constructs a domain-specific knowledge base and ontology model using existing tickets and input from a domain-expert. This phase is executed either when NetSieve is bootstrapped or to re-train the system using expert feedback. Second, the operational phase uses the knowledge base to make problem inference from tickets. To export the inference output, NetSieve supports SQL and outputs results in a variety of data formats such as XML/JSON. Third, the incremental learning phase improves the accuracy of knowledge base using human guidance. To improve inference accuracy, it is important to continuously update the knowledge base to incorporate any new domain-specific terminologies. NetSieve provides an interface to allow a domain-expert to give feedback for improving the ontology model, synonyms, and whitelists. After each learn-

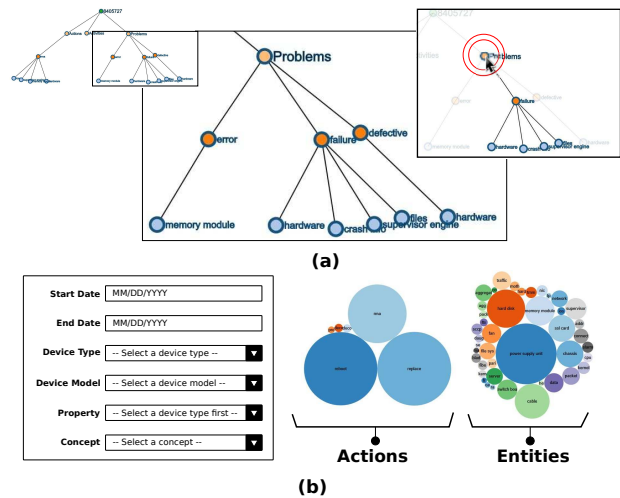


Figure 1: NetSieve Visual Interface

ing session, NetSieve performs problem inference using the updated knowledge base. Our evaluation shows that NetSieve provides 89%-100% accuracy and its inference output is useful to learn global problem trends.

3 Demonstration

The demo is divided into two parts: Part 1 introduces *Concept Tree* (Figure 1 (top)), a hierarchical tree-like visualization, to summarise the problems, activities and actions from a trouble ticket. Part 2 shows *Trait Charts* (Figure 1 (bottom)), bipartite bubble plots (where the size of a bubble is proportional to the frequency), that show the aggregate problem trends in terms of the top-k entities and actions inferred across tickets. For instance, what are the most failing hardware components across routers?, what are the major actions taken for repair? and what problems frequently occur after a configuration change? We will demonstrate how NetSieve enables answering these questions across a large ticket dataset.

References

- [1] POTHARAJU, R., JAIN, N., AND NITA-ROTARU, C. Juggling the jigsaw: Towards automated problem inference from network trouble tickets. In *Proceedings of Networked Systems Design and Implementation (NSDI)* (2013).