ART: Learning Operation Tree Patterns for Cloud Remediation

Hongyi Huang, Wenfei Wu
Tsinghua University

**Background & Motivation**

As IT service grows in a larger scale and more complex, operation trees (OTs) are raised to take place of IT operators in repeatable remediation tasks.

Operation trees are
- tree-organized
- configured with certain rules (edges) that determine the workflows and actions (nodes) that execute workflows on infrastructures
- in always-running states. The following is an OT example.

![KVM hot migration operation tree]

Even though some tools offer graphical interfaces to accelerate designing trees, they barely give deep insights into design patterns, which never ceases redundant construction and configuration.

**Problem**

IT operators have great difficulty in designing new operation trees targeting thousands of uncovered faulty scenarios.

**Goal and Solution**

Our goal is to leverage existing experience for reusability to assist constructing new workflows that handle more cases or alarms.

We devise ART, a scheme to learn instinct and structural patterns from existing operation trees and automatically generate trees in demand in an advisory capacity.

**How to Learn Patterns**

- Structural Pattern:
  - Common attributes of nodes
  - Similar relationship of nodes (adjacency)
  - Extraction from the attributes of nodes
    - apply NLP methods to extract keywords that summarize the nodes
  - Apply Tree Edit Distance (i.e., APTED) to measure relationship of nodes
    - assign add, delete and rename operations to measure structural similarity between trees

**Tree Pattern Mining**

- redefine operations in APTED to better characterize the structural similarity between trees
- apply enhanced-APTED to each pair of de facto trees to capture structural patterns
- exploited similar trees are classified into several clusters based on TED value
- representative trees (namely patterns) are elected in per-cluster manner and stored into the database
- assign topics (keywords) as indexes to patterns for future retrieval
- record occurrences of pattern workflows as support

**Operation Tree Generation**

If operators is building up a new tree from scratch
- retrieve patterns based on the keywords that are extracted from description of the new tree
- recommend those matched patterns of high support.

**Operation Tree Completion**

If a pending tree has partial structures completed
- decompose these structures
- find a match against any prefix of patterns by APTED