TO UNDERSTAND ANOMALIES..
...YOU MUST FIRST UNDERSTAND YOURSELF

2024-06-28

Wes Hardaker <hardaker@isi.edu>
Network operators are plagued with odd anomalies

What are these???

What is this????
Let’s think about this

How can we do *intelligent* analysis of anomalies?

How do we know *what is normal traffic* v.s. *what is new*?
To find what is abnormal, you need to know what is normal

- Doctors use **bilateral comparison**
  - When your foot hurts... compare your left to your right
  - Is one more swollen than the other?
  - Is one larger?
  - Differently colored?
  - Imagine if we didn’t have a left and a right!
  - Your doctor would have to memorize what normal was

- Network anomalies should be treated with similar analysis
An example DDoS attack field values
Spot the difference!
What value is the problem?

Normal Traffic

Attack Traffic Sample
A smaller anomaly – spot the difference!

Normal Traffic

Anomaly Traffic Sample
Let’s subtract – what is: right – left?
This is right overlapped with left – spot the difference!
If you look carefully...
you can see discoloration
This is the real *right* - *left*

When we remove “normal”
we are *left with just the abnormal*
Bilateral Comparisons
A specific problem space

Important considerations

- Major shifts in traffic are from a single cause
  - Something has clearly changed. *What?*

- “More of the same” ramp ups are not an “anomaly”
  - AKA diurnal patterns are “normal”
Terminology: left = “normal” right = “anomaly”

First profile based on a “normal” time: a “left” sample

Leaving the anomaly as a “delta”

Then find what is new in an anomaly: a “right” sample
INTRODUCING “TRAFFIC-TAFFY”

The toolset that helps automate bilateral traffic analysis
Introducing traffic-taffy

- **Approach:**
  - Perform deep packet inspection of a **base-line**
  - Perform deep packet inspection of an **anomaly**
  - Compare levels for **each value of each protocol field**
  - **Sort, Filter and Report** based on findings

- **Some of the tools:**
  - taffy-dissect: enumerates field counts in a pcap file
  - taffy-compare: compares one file/time-range against another
  - taffy-graph: graphs enumerated fields in pcap files

- **Easy to install:** `pip install traffic-taffy`

Warning: Beta Software
Let’s go back to studying traffic anomalies

Using bilateral comparisons!
Case Study 1: three large bumps seen at b.root-servers.net

Dataset:
- Three 5x spikes
- At 1 anycast site
- What are they?
- Can we find the root cause?

Graph produced with taffy-graph
taffy-compare: find differences between points in time

Taffy-compare:
- Takes PCAP data from two points in time
- Uses the *left* side as a “normal” profile
- Identifies major shifts in the *right* side

Output: colorized results to the console
- Total counts per protocol field
  - Left and right
- Percentage of traffic for each field value
- Deltas between both values and percentages
- All filterable by threshold values
taffy-compare: find differences between points in time
taffy-compare example: colored console differences

Important note: I did not pick these fields to study – the tool did!
Warning... this is easy to recreate
Try it!

• Please test it!
  – https://traffic-taffy.readthedocs.io/
    • Multiple (longer) video presentations
    • The DNS-OARC video has greater usage detail
  – pip install traffic-taffy

• Thank you to the Comcast innovation fund for sponsoring this work!

Wes Hardaker <hardaker@isi.edu>