NETWORKWORLD Reprint

THE CONNECTED ENTERPRISE ≡ JANUARY 24, 2015

Prelert's Anomaly Detective: Finding Black Swans in Big Data

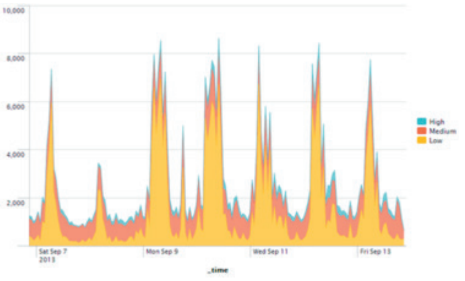
Anomalies are easy to find when they're big and frequent but what do you do when they're small and rare?

BY MARK GIBBS, Contributing Editor

Think about your systems and services ... what is normal behavior for them? You can probably characterize normal for, say, a network backbone carrying some mix of traffic with a bandwidth utilization between X and Y and so on. But now consider the opposite problem: What is abnormal?

Abnormal conditions can sometimes be easily detected, for example, if bandwidth utilization goes to 100% and stays there or SMTP traffic doubles. When the abnormal is characterized by large deviations from expected behavior the diagnosis is often (but not always) easier than when transient abnormal deviations occur because the latter get easily lost in the detail. So, which is more dangerous in almost any context? The answer: Transient abnormalities. Whether it's network security, e-commerce, or car engine performance transient conditions that aren't detected and corrected can easily lead to widespread systemic failure.

So, say hackers have infiltrated your network and FTPing customer account data to an external site. Noticing that one server amongst several hundred in your organization that doesn't usually establish FTP sessions is doing so once every ten minutes would be a needle in a haystack situation given the sheer volume of all other transactions. To find this you'd need to have characterized your network traffic and have a tool capable of detecting relatively anomalous behavior in huge amounts of network data. This is definitively a Big Data problem and, moreover, it's a realtime Big Data problem because if you don't detect and address anomalous behavior such as hacking in realtime you could suffer a loss that could have been either reduced in scope or even completely prevented.



Prelert
Responsys' IDS reported between 100,000 to 150,000 events per day.

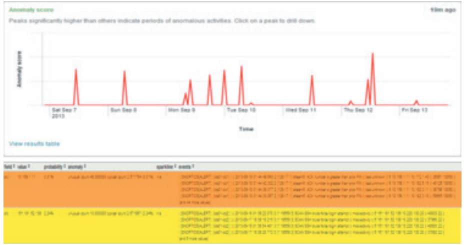
Anomaly detection in realtime is exactly what Anomaly Detective from Prelert can do. Presented with huge amounts of otherwise indigestible Big Data, Prelert Anomaly Detective applies analytical and statistical techniques to identify behavior that is outside the norm.

For example, quoting from a Prelert case study, Responsys, a relationship-based digital marketing company, routinely saw their Intrusion Detection System report between 100,000 to 150,000 events per day. These included between 2,000 to 6,000 high severity alerts, "an impossibly large quantity for their security team to sift through."

Instead of relying on writing rules and thresholds (which despite their best efforts resulted in thousands of false alerts), Anomaly Detective helped the team easily monitor for red flags in their IDS logs with a single command:

```
index=ids sourcetype=snort | eval
dest:protocl:port=dest+":"+dest_port
| prelertautodetect population_
rare(signature_name), population_
distinct_count(category), population_rare_
count(dest:protocol:port) by src
```

This command is saying to look at the IDS index from Snort (the software they use to aggregate their data), and create a new field that is the destination protocol and destination port. Run this through Prelertautodetect and identify the source as the population, and for each individual source that Snort sees, look for rare signature names for that particular source. Then look for an unusual distinct count of categories for that source, and when that source talks to a destination protocol and port look for connections that are rare that the host has never triggered before, or where that host is generating an unusual amount of events for that particular destination.


Anomaly score
10m ago
Peaks significantly higher than others indicate periods of anomalous activities. Click on a peak to drill down.
View results table

Prelert
Prelert Anomaly Detective's filtering of Responsys' IDS data to identify anomalies.

Anomaly Detective helped Responsys turn their daily IDS alert dashboard from showing thousands of alerts each day into [1 to 3] IDS alerts per day.

With such an enormously reduced level of alerts, analyzing the causes and taking timely and effective action becomes possible and that's what really matters when active network infiltration is underway.

Anomaly Detective can also be used with any other Big Data such as ecommerce web site metrics to identify success and or failure edge cases and sales funnel issues.

There are two versions of Prelert's Anomaly Detective; the Anomaly Detective app that works with Splunk, an operational intelligence platform, and the Anomaly Detective Engine & API (for NoSQL, Hadoop and other datastores).

Prelert offers a Free Anomaly Detective license for up to 0.5GB of daily indexed data which includes comparative searches, autodetection of anomalies, categorization of unstructured data. Premium plans, which add real-time analysis and support, cost $75 per month for 1GB of daily indexed data and $225 per month for 5GB. Data volumes over 10GB are by quotation.

prelert®

www.prelert.com
sales@prelert.com
1-888-PRELERT

© Copyright 2015 by Network World, Inc., Southborough, MA 01772-9108 • Reprinted from Network World.com • Trademark is owned by International Data Group, Inc.
#C35667 Reprinted by The YGS Group, 800.290.5460. For more information visit www.theYGSgroup.com/content.