datastream.io

Mentat Innovations

NETWORK INTELLIGENCE AT SCALE

NETFLOW ON LARGE ENTERPRISE NETWORKS
APPLICATION LAYER DATA
LOCATION FROM MOBILE VPNS
UNSTRUCTURED CONTENT FROM DEEP PACKET INSPECTION OR SYSTEM LOGS

GLOBAL REAL-TIME INTELLIGENCE OF LARGE ENTERPRISE WAN AND VPN:

- NETFLOW LAYER
- APPLICATION LAYER
- MOBILITY AND LOCATION
- DEEP PACKET INSPECTION

FEATURES:

- SYSTEM LOGS

PRINCIPLES

- Your anomaly detector is only as good as your model of normal behaviour.
- Fixed rules can quickly become obsolete. Streaming learning ensures rules evolve with the data automatically.
- Normal behaviour is complex, heterogeneous dynamic and often erratic.
- Scaling with velocity needs:
  - **real-time information extraction and data reduction on-the-fly**
  - **algorithms and mathematics that allow for fast updates without compromising on accuracy**
  - **self-tuning behaviour that does not require human data scientists to supervise models**
- Anomalous behaviour is usually not a single occurrence, but:
  - **a chain of events**
  - **at multiple timescales (per second, per minute, per hour, per day, per week, ..)**
  - **at multiple layers**
  - **spanning multiple connected devices**
  - **occurring as isolated anomalies, bursts of anomalous activity, or persistent changes affecting not only traffic volume, but also IP destinations, applications, protocols, etc.**
- Actionable insights rely on interpretable output and human-readable alerts.

FEATURES

- **Semi-structured data:** Vectorise unstructured input to construct features (e.g., system logs).
- **High-dimensional learning:** anomalies often manifest themselves in unlikely combinations of features, rather than unlikely values in a single feature.
- **Built-in feature extraction** automatically focuses on the most informative feature combinations, generalising beyond known signatures to uncover hidden patterns, and enabling zero-day threat analytics.
- **Predictive analytics** underlie the entire toolbox as they are internally used by the learning algorithms. This means that forecasting abilities are also built-in
- **Peer group analysis:** behaviour that seems unlikely at the individual device level might be easier to explain by looking at a device's peer group. We design peer groups by a combination of device features as well as streaming clustering technology which identifies similarities across vast numbers of devices with respect to thousands of different attributes.
  Interpretability: although we use powerful machine learning to discover interesting patterns in the data, we are always able to justify why an event was flagged as anomalous in the form of visualisations and decision rules in human-readable forms:

> *"This device triggered a **red** alert because it sent X amount of traffic to IP A where the request contained the keyword B from location C. The probability of this behaviour being 'normal" is 1% for this user, and only 3% for its peer group."*

- **Self-tuning, unsupervised operation:** a critical bottleneck in the attempt to scale pattern recognition technology with velocity is the need for a human data scientist to tune the models. Using the latest available techniques in stochastic learning we design robust algorithms that tune themselves on-the-fly using the accuracy of t heir predictions about the future as their guide.
  No arbitrary parameters are fed to the user to tune using time-consuming trial-and-error.
- **Scale with asynchronous multiple input sources**
- **Sequence-of-events analysis:** using sequential analysis and advanced streaming decision tree technology we are able to anticipate the next request of a mobile device or the features of the next packet in an IP transaction, and understand anomalies in event chains.
- **Graph analysis:** features of interest sometimes appear at the network level, rather than the device level: unlikely paths through an enterprise network, sudden appearance of hubs in the network at large; sudden changes in the connectivity of a network of IPs.

Cloud Offering : Offered as SaaS solution on AWS consuming directly data off your S3 repositories or via custom data consumers.

Mentat Innovations
Imperial College Incubator
Level 1 - Bessemer Building
London SW7 2AZ - UK

+44 203 6373330

www.ment.at
info@ment.at