Cybersecurity Analytics
Leverage Your Data Lake To Deliver Analytics on Cyber Intrusion Behavior

With the rapid adoption of mobile, cloud, big data and the internet of things (IoT), cybersecurity tools need to detect potential cyber attacks faster. To do this, you need to blend a large volume and variety of data from the applications, network and server. This creates a big data problem that is hard to solve because bringing data together from different formats, at scale, is tough.

**Cybersecurity Problems Are Data Problems**
Because traditional methods of handling cybersecurity involve custom scripting and integrating multiple data sources, it's actually very time-intensive to blend multiple data sources at scale. Data scientists often spend 80% of their time cleaning the data to get it ready to analyze, and only 20% of their time analyzing the data. Often, data scientists are comfortable with scripting – but is custom coding the best use of their time?

**Finding Anomalies Is Time-Consuming and Urgent**
Once you know what a security threat looks like, it's easy to block. But how do you know what’s a threat in the first place? That’s hard, and that’s where data scientists come in to look for patterns.

The average analyst takes 205 days to detect an attack due to the huge volume and variety of data to go through. And it’s not just big data, but all kinds of messy data from all kinds of sources, including log files, web access data, and more, that are difficult to blend. It takes time to clean and blend the data.

**What Can You Do?**
As you evaluate your options, ask whether they provide:
1. An easy to use tool where you can drag-and-drop data sources.
2. Ways for analysts to analyze intrusion behavior faster.
3. How to handle “unknown unknowns” beyond blocking, enforcing and remediating threats.
4. Cost-benefit analysis: Find what is most effective for analytics on Hadoop.

**Hitachi Vantara’s Pentaho Platform Heightens Cybersecurity**
This platform enables end users like forensic analysts, cybersecurity analysts and data scientists to detect cyber threats faster by:
- Visualizing behavior patterns that could lead to the next cyber attack.
- Automating the orchestration of data flows to prepare, blend, report and alert security data with corporate data to better understand the source of intrusion or threat.
- Empowering data scientists with predictive analytics through integrated machine learning tools like R, Python and Weka.
- Reducing complexity by bringing all security data into a single platform, without the deep cost to deploy, maintain and operate other cybersecurity system.
Pentaho Capabilities

Use Pentaho to:

- Enable your data scientists and cybersecurity analysts to focus on analyzing potential intrusion behavior rather than spending valuable time on data prep.
- Reduce the high volume of hand coding data parsing and transformation processes in Hadoop to enable high-scale analytics across all potential security risk points.
- Maximize your data lake investment through quickly blending a complex variety of data to deliver trusted security intelligence.
- Our cybersecurity analytics solution is based on Pentaho’s enterprise-ready reporting, visualization and easy-to-use business analytics. Pentaho ensures that analysts have exactly the information they need to improve security practices within the enterprise.

Benefits

- **Productivity:** Pentaho Data Integration (PDI) delivers a streamlined and automated Hadoop orchestration that requires fewer staff to maintain and execute complex data ingestion and transformation; make your existing team more productive.
- **Reduce Complexity:** Avoid hand-coding data integration for unstructured and semistructured data. Nontechnical users can use PDI to easily incorporate new data sources as business needs evolve. No is coding needed.
- **Robust Analytics Option:** With Pentaho, embedding self-service analytics into applications like SIEM becomes easy. Delivering trusted data to analysts in a governed fashion for broader security behavior analysis ensures the right decisions and actions are taken to improve cybersecurity across an organization.
- **Customizable:** 100% Java-based platform leveraging the latest in open source developments, plus robust APIs, means you can customize functionality when you need it.

Case Studies

**Top Energy Company**

**Challenge**

Cyber attack detection is a complex big data problem requiring the parsing and analysis of huge log volumes coming from different sources and locations. Additionally, the data come in many formats, making it difficult to blend the data and find patterns.

**Solution**

Parsers that took weeks to develop through complex scripting were recreated in Pentaho in a single afternoon. Pentaho’s natural language processing capabilities and R algorithms were deployed at scale on Hadoop.

**Benefits**

Enabled client to detect cyberattacks faster – with a 10 times reduction in development time for processing and analyzing log data.

**Top Telecom Company**

**Challenge**

Detect anomalies in network traffic patterns that identify suspicious activity. For example, recognize an unexpected spike in outbound connections from a system with sensitive data.

**Solution**

Orchestrated Kafka and Apache Bloom, and used Pentaho transformation logic to identify anomalies in near real time.

**Benefits**

Greatly reduced time and complexity of network traffic anomaly detection by orchestrating leading big data technologies and algorithms.

About Hitachi Vantara

Hitachi Vantara, a wholly owned subsidiary of Hitachi, Ltd., helps data-driven leaders find and use the value in their data to innovate intelligently and reach outcomes that matter for business and society. We combine technology, intellectual property and industry knowledge to deliver data-managing solutions that help enterprises improve their customers’ experiences, develop new revenue streams, and lower business costs. Only Hitachi Vantara elevates your innovation advantage by combining deep information technology (IT), operational technology (OT) and domain expertise. We work with organizations everywhere to drive data to meaningful outcomes. Visit us at HitachiVantara.com.

Hitachi Vantara

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
HitachiVantara.com | community.HitachiVantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
HitachiVantara.com/contact

HITACHI is a registered trademark of Hitachi, Ltd. Pentaho is a trademark or registered trademark of Hitachi Vantara Corporation. All other trademarks, service marks and company names are properties of their respective owners.

P-033-B BTD April 2019