Forcepoint DLP

Risk-Based DLP Incident Ranking

Revision A

Contents

Introduction

Security systems can generate a large number of alerts, but only a small number are a genuine risk to the organization. Broken business processes, false positives, and minor breaches can create noise that make the task of identifying data theft activity challenging, not to mention increase operational costs.

To solve this security challenge, Forcepoint has developed an integrated analytics system that:

1) Correlates related incidents and alerts into meaningful DLP cases

2) Applies various statistical methods to assess baselines and identify anomalies

3) Utilizes artificial intelligence to recommend a data loss classification (e.g., data theft, broken business process, and unintentional leak) and provides the business context for each case (who, what, why, and when)

4) Assigns a data loss risk score to each case

The score represents the actual data loss risk and is designed to enable the security operations team to initiate an appropriate investigative response. The risk score is evaluated by algorithms that combine knowledge about the content, baseline information, and various observables and indicators regarding the data, the source, and the destination. These indicators are fused together using a framework called **Bayesian Belief Networks** that, eventually, allows the system to accurately assess the likelihood of data theft and other data loss classes.

This integrated security analytics feature provides capabilities that enable Forcepoint DLP customers to gain much better visibility and facilitate fast triage of DLP incidents. It also offers automated identification of broken business processes.

Future releases will build on these capabilities and address additional use cases such as allocating a lower risk score to personal communications and supporting automated policy efficacy tuning.

This paper discusses some of the analytical and statistical techniques used to deliver the security analytics capability within Forcepoint DLP.

Quantifying risk

Each person has an intuitive notion regarding risk; however, assigning a meaningful and consistent risk metric is difficult. Although some clear high-risk cases are easy to discern—such as a file with thousands of credit-card

numbers that was sent in the middle of the night to a dubious destination by an employee with a poor record—it is much harder to decide about cases with an ambiguous data classification, or incidents within the "gray area". These can stem from an employee's mistakes or broken business processes or from sophisticated insiders who attempt to make their activity look "normal".

Systematic approaches to risk quantification and management were first developed in the field of insurance and were based on the **expectation value of the loss**. Broadly speaking, this can be expressed as:

Risk = (Probability of "bad" events)•(Amount of loss associated with the events)

To this day, insurance underwriting is still based on this basic formula, which is also widely used for quantifying other risks, and it is, by large, the canonical measure for risk quantification.

The intimate acquaintance of content-aware DLP with sensitive content, whether that be intellectual property or regulated data sets, allows the system to assess the potential damages or losses associated with cases in which a certain type of content is stolen or otherwise exposed.

In general, the impact would be based on the classification and the size of the exposed data: an incident with a single credit card data number is much less severe than an incident with a hundred numbers, which is yet less severe than stealing credentials for a database with millions of sensitive records. In order to assess the risk, we need also to assess the probabilities of the various possible "scenario classes." Was it a deliberate data theft? In this case, the impact can be very large, and there is an urgent need to address the problem. Was it a broken business process, where information is exchanged in a non-secure manner? In this case, the risk is enduring and requires systematic, yet not urgent, action. Or was it was a one-time mistake?

On the other hand, false positives and events of low importance, such as personal communication, also convey cost, associated with the time that was spent and the attention that was devoted for their analysis, as well as the "opportunity cost", associated with missing high-impact events that got "lost in the shuffle". It is therefore also important to be able to identify those superfluous incidents, whenever possible.

In order to assess the probabilities, our researchers have developed an advanced tool based on a technology called **Bayesian Belief Networks**, that utilizes the various observables and indicators to assess the plausibility of the various scenarios by combining expert's knowledge, deep learning techniques, and statistical inference.

A key observation in this respect is that we need to see behind the single alert or incident. Before assessing the risks, the system first correlates related incidents into **cases** that aggregate various incidents based on key attributes such as the source, destination, and data types, as well as more subtle patterns that take into account various similarity measures between incidents.

After constructing the various cases, the probabilities of the various scenario classes or possible explanations are assessed using special Bayesian Belief Networks that were developed and trained for these specific classes. The various explanations compete with each other, and, eventually, the product obtains the likelihood of each scenario, as illustrated in Figure 1:

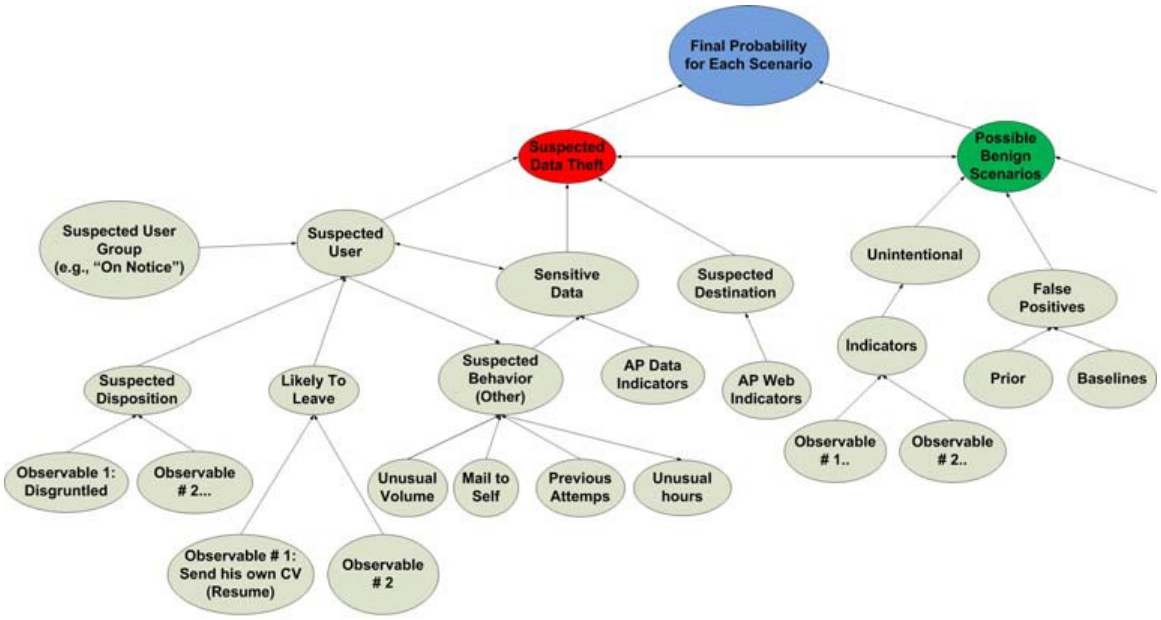
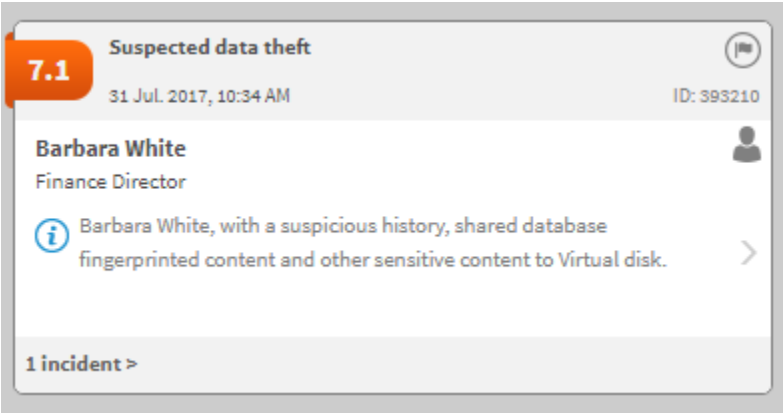


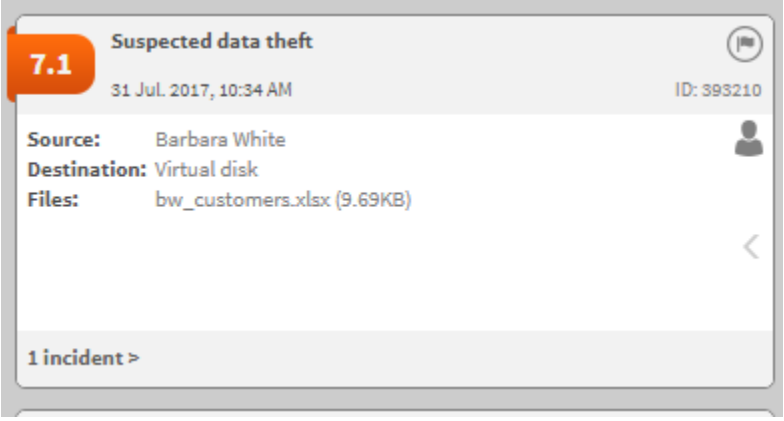
Fig. 1

Bayesian networks simplify the assessment of likelihood based on multiple observables and indicators using the notion of conditional independence in various hierarchical levels. The system can group a relatively small number of observables and indicators—such as the employee sent his own resume to another company's HR and sent an email that suggests a negative disposition to their boss—to assess the plausibility of various hypotheses — such as that the employee is likely to leave soon. This could indicate that an incident where he sent source-code to his own Gmail account is more likely to be data theft incident than a case in which he merely wanted to work on the code on his spare time.

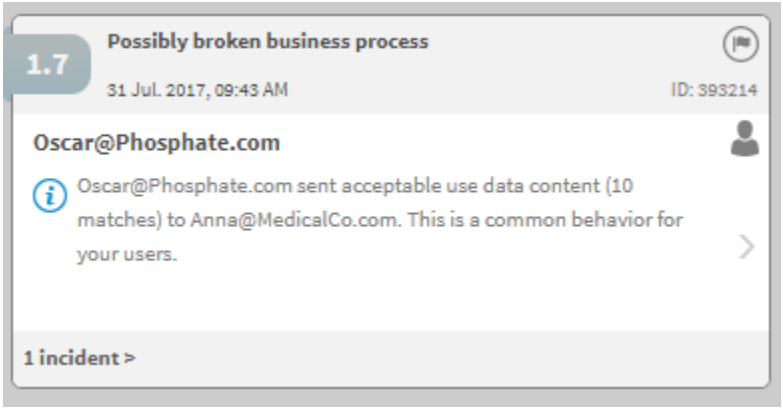
On the other hand, in some cases the content itself may provide strong indication that the incident is a data theft incident—for example, there is no conceivable reason to send the Security Account Manager (SAM) database to an external Gmail address.

The results are automatically explained on cards in the Forcepoint DLP incident risk ranking report:

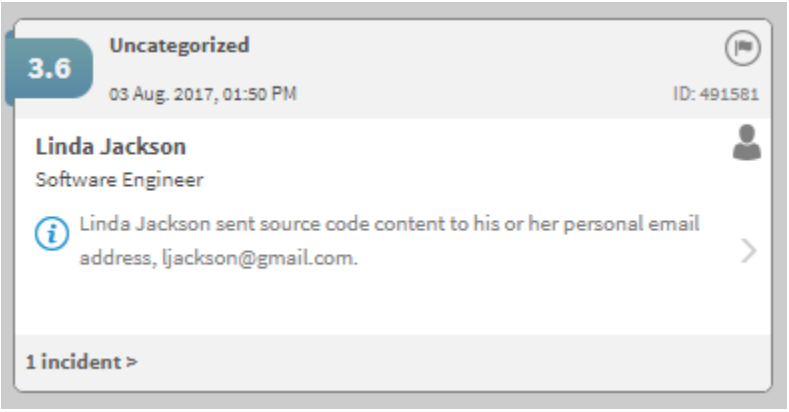




In some cases, the system identifies patterns of data breached that are more likely to be a broken business practice:



In other cases, it would be hard (or impossible) to identify the scenario. The system renders these as “uncategorized”:



Folding, chaining and grouping incidents

Grouping incidents is an effective way to summarize data and overcome the deluge of incidents. In principle, an incident group is a collection of incidents that can be meaningfully described. Forcepoint DLP defines four basic types of groups:

- *Basic cases and folding*
- *Incident chains and processes*
- *Superfluous incidents*
- *Behavioral baselines and anomalies*

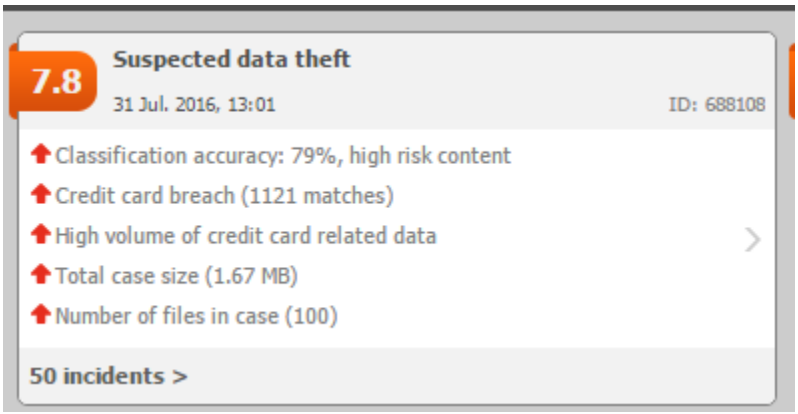
Related concepts

Basic cases and folding on page 6
Incident chains and processes on page 6
Superfluous incidents on page 7
Behavioral baselines and anomalies on page 7

Basic cases and folding

A basic case comprises one or more incidents that, from user's perspective, should be referred to as a single transaction—for example, copying a directory that contains sensitive data within multiple files to removable media, or uploading a single file to cloud storage and the file being split into multiple data chunks by the web application. In these instances, all these incidents are folded into a single **case**.

The risk for the case is evaluated by first assessing the total impact of all the incidents in the case and the probabilities for various scenarios (data theft case, false positive etc.). The following card summarizes a case with 50 incidents involving credit card data:



Incident chains and processes

At the next level, the system looks at multiple incidents that together highlight a story. Chain-like cases are a sequence of incidents from the same source that highlight is as illustrated in fig. 2:

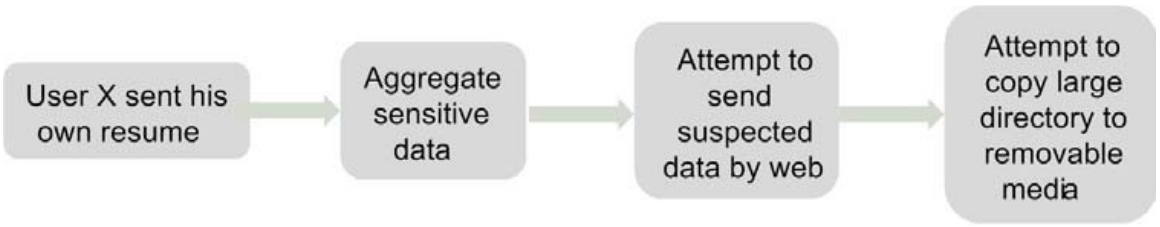


Fig. 2

This sequence of incidents constitutes a case that can be characterized as a chain. The context provided by previous incidents highlight the intention of the subsequent incidents, a data theft attempt in this case.

Other cases involve incidents that were created as part of a process, such as a sequence of events generated by an individual, a group of users or a machine that is used to achieve a certain goal or related to a certain theme, legitimate or illegitimate. Notable examples for such processes are business processes, in particular broken business processes where sensitive data is rendered unprotected, and deliberate data theft activity.

Superfluous incidents

Groups can include incidents that should not have been there in the first place—for example, false positives and personal communication. While 100 percent accurate automatic identification of false positives is virtually impossible, the system can assess the probability of false positives or the confidence level that an incident is a true positive. To do so, it uses:

- Statistical methods
- Deviations from baselines
- Prior information about the precision of the classifiers and rules in the various DLP sensitivity levels (“Wide”, “Default” and “Narrow”)

Behavioral baselines and anomalies

Baselines provide reference regarding normal (versus legitimate) behavior. Baselines are time-dependent and can be associated with sources, destinations, channels, and content in various levels of granularity. For example, the system can consider specific users, user groups, or the entire organization baselines for standard working hours, combinations of channels, rules, number of matches, destinations, and transaction sizes, as well as anomalies or deviations from the baseline that are statistically significant. In order to obtain statistically significant results, the number of elements in the group should be large enough, which may require lowering the resolutions. As a rule of thumb, the minimal group that constitutes a baseline should comprise at least 30 elements.

While anomalies provide an important set of indicators, most of the behavioral anomalies are benign, as people often change their behavior. For example, when you start working on a new topic, with new suppliers or customers, or when you travel to places you’ve never been to before, you create anomalies that may or may not become the new normal. Incorporating baselines and anomalies within a powerful probabilistic framework, such as Bayesian Belief Networks, allows digesting the relevant information from these indicators without creating the deluge of false positives typical to products that alert on each anomaly.

Conclusion

Advanced analytical capabilities are essential for obtaining an effective DLP solution. Such capabilities should allow going behind single alerts to digest various indicators regarding the source, destination, content, and environment and to fuse them together to provide an accurate risk metric. Forcepoint DLP, which combines Bayesian Belief Networks, incident grouping, anomaly detection, and impact assessment, allows you to gain much better visibility and facilitate fast triage of DLP incidents.

Published 26 July 2022