GURUCUL
PREDICTIVE SECURITY ANALYTICS

Uncover Insider Threats

through Predictive Security Analytics

WHITE PAPER

Executive Summary

Gurucul pairs User Behavior Analytics with Identity and Access Intelligence to pinpoint threats from legitimate insiders and external intruders using compromised credentials

While most organizations put their focus on defending against and detecting cyber attacks, a more insidious threat is on the rise. Information security (InfoSec) professionals say that insider attacks are far more difficult to detect and prevent than external attacks, and insider threats have become more frequent in the past year.[1]

Call it the Edward Snowden Effect, if you like. Once the extent of this infamous National Security Administration (NSA) contractor's actions became apparent, many organizations were forced to ask themselves if a similar insider attack could happen to them. The answer, of course, is yes. Any organization that has information of value is vulnerable to threats from within. An incident in which confidential or sensitive information is stolen or damaged may not have the high profile national security ramifications as in the Snowden case, but significant damage can be done to the organization nonetheless.

There's a secondary kind of threat that relies on having an insider's access credentials and privileges, although it's not the employee himself doing the dirty work. In this scenario, an external intruder uses a compromised account to surreptitiously gain access to the internal systems. Phishing attacks and data breaches have made available massive lists of user credentials and passwords for sale on the Dark Web. For a small fee, anyone can obtain legitimate, active credentials to login to all sorts of enterprise networks and SaaS applications, and for all intents and purposes, the intruder appears to be the real user.

As a prime example the Securities Exchange Commission (SEC) filed a complaint in August 2015 against 32 named parties, noting the defendants made trades on illegally obtained information and reaped over $100 million in unlawful profits. The international fraudulent scheme involved the hacking of computer servers of at least two newswire services over five years to review confidential information of an estimated 100,000 press releases on publicly traded companies not yet released to the public. The defendants used the stolen nonpublic information to make trades and reap profits.

Workers inside the perimeter don't need to use malware and other such techniques to gain access to sensitive servers, databases and applications. They already have legitimate access to these systems, giving them the opportunity to steal information, corrupt essential computer systems, and disrupt business as usual.

1. In a survey of information security professionals conducted by Crowd Source Partners, 62% of respondents say that insider attacks are more difficult to detect and prevent than external attacks. Likewise, 62% say insider threats have become more frequent in the last 12 months, but only 34% of respondents expect additional budget to address the problem. Crowd Research Partners, "Insider Threat Spotlight Report," 2015.

Executive Summary

User behavior analytics (UBA) is often advocated as the best means to detect nefarious activity by internal actors. UBA involves keeping track of what users are doing – particularly those with elevated privileges such as system administrators, and workers with access to highly sensitive information like intellectual property (IP) or customer account data – and looking for behaviors that are outside the range of normal activities. Certainly analyzing users' behavior is important, but this method is insufficient on its own. A determined actor, especially one who knows the internal systems intimately, can do his damage without raising red flags over his behavior. Conversely, even sanctioned activities can be flagged as suspicious, creating too many false positive alerts that overwhelm InfoSec analysts who are already stretched thin.

A much more effective approach combines UBA with in-depth intelligence about a user's identity attributes and the privileges he has on the network. This approach involves analyzing the access rights and entitlements a person has; the activities he has been performing across multiple accounts, both now and in the past; and the typical activities that members of his peer groups are doing. It takes a combination of the right data sources, sophisticated machine learning and perceptive data science to pinpoint truly aberrant actions that are good indicators of misuse of assigned privileges.

This paper provides an overview of methodology that merges user behavior analytics with identity access intelligence to yield highly accurate indications of insider threats. Once these activities are highlighted, alerts can be raised to the security operations center for investigation, response and remediation.

The Challenge of **Insider Threats**

The Insider Misuse pattern shines a light on those in whom an organization has already placed trust—they are inside the perimeter defenses and given access to sensitive and valuable data, with the expectation that they will use it only for the intended purpose. Sadly, that's not always the way things work.

— Verizon 2015 Data Breach Investigations Report

Most organizations devote a majority of their IT security budget and other resources on the defense and detection of cyber attacks coming from outside the network perimeter, mainly through signatures and rules. Companies prepare for the eventuality of threats from sophisticated cyber actors such as criminal syndicates and rogue nation-states. Meanwhile, a more pernicious threat is already sitting inside the network and wearing an employee badge, at least figuratively. This actor can be a true insider – employee, contractor, business partner or vendor – or an external intruder who is mimicking an employee through subversion of a legitimate set of credentials.

The traditional security systems – firewall, SIEM, IDS, IPS, DLP, vulnerability assessment, sandboxing – are primarily tuned to detect breaches coming from the outside. They use rules to deny entry or passage of suspicious traffic and block known security problems based on a database of attack signatures. They also identify probes and open vulnerabilities to exploit around the network perimeter and internal systems. This is all well and good for preventing and detecting external attacks, but these techniques can't catch the malicious insider because he isn't using those attack methods. He doesn't have to. The organization has already given him (or rather, his digital identity) the privilege of accessing databases, servers, applications, data repositories, application code, and more—the systems that contain the organization's most sensitive or vital information. Moreover, the traditional security defenses typically don't watch the insider's activities or traffic because they are assumed to be legitimate for the person's work roles. This misplaced assumption is what enables the Edward Snowdens of the world to succeed in their malevolent missions.

Providing access via authentication and authorization like a key and then assuming all is well no longer works, much in the same sense as protecting resources with just signatures and rules. Moreover, the combination of on-premises and cloud apps in a hybrid environment increases the complexity of detecting insiders and compromised accounts.

Catching the malicious insider is a real challenge because activities alone don't explain the intent of what people do. Consider the system administrator who is found to be downloading the configuration data from the company's numerous firewalls over a period of days. This would certainly be viewed as a highly suspicious activity because the information is so critical to the enterprise's security posture. But the incident must be put into its proper context. As it turns out, this system administrator is collecting this data in advance of an upcoming security audit the organization must undergo. The IT department wants to ensure that the security devices are properly configured before the audit team assesses them and dings the department for misconfigured devices that result in non-compliance. In other words, the "suspicious activity" is actually part of the system administrator's job responsibilities, albeit one that is performed infrequently.

In an organization of 1,000, 5,000, or perhaps 30,000 employees, how can one write a set of rules for a system or device that sees this unusual activity of downloading sensitive configuration data and somehow know that it's okay in the context of preparing for the audit? Rules aren't effective in this scenario, and thus the false positive incident is flagged and assigned to an overworked security analyst for investigation. It's a waste of a limited and expensive human resource.

More information and better analysis of the situation can help determine if this is a genuine insider threat or not.

Using Identity as a **Threat Plane**

Over the next three years, leading UEBA platforms will become preferred systems for security operations and investigations at some of the organizations they serve. It will be — and in some cases already is — much easier to discover some security events and analyze individual offenders in UEBA than it is in many legacy security monitoring systems.

— Avivah Litan, Gartner, September 2015

Sifting through massive volumes of user activity, and even applying machine learning algorithms to understand appropriate user behavior patterns over time, is helpful to spot anomalous behavior, but insufficient to rule out time-wasting false positive alerts. Adding data and analysis of a user's identities and entitlements, and putting that information in context with the user's peer groups, vastly improves the process of pinpointing malicious insider activity.

For any given user, it's not uncommon for the person to have multiple digital identities for the various systems he logs into and applications he uses. And for each identity he might have multiple entitlements; for example, the right to upload and download data, to change or update records, to delete data or files, and so on. Altogether, these numerous identities and privileges create quite a threat plane—places where data or information can be stolen or damaged in some way.

Identities and entitlements are often in a state of excess access due to manual processes built upon legacy rules for identity management. This provides the insider or hijacked account user more room to roam than desired creating undue risk waiting for abuse. The perfect balance is the right data for the right user when they need it, and never access when they do not. Applying user behavior analytics for risk-ranked access is changing identity management to reduce excess access, manage high privilege access and detect orphan and dormant accounts as top level examples.

To really understand a user's identity, and to determine the risk of that identity as a threat plane, it's essential to collect relevant data from a variety of sources. Gurucul collects data from a number of internal and external sources, including: -

Identity Management Systems

Data is drawn from internal directory services, identity and access management platforms, human resources systems—wherever "people" and "account" information is kept. Gurucul has out-of-the-box connectors to ingest this data from other vendors' systems. Collecting this identity data allows us to understand who the people are within the organization, and what legitimate access rights have been assigned to them.

Privileged Account Management Systems

Many enterprises use specialized tools to control and track the activities of powerful accounts, such as those belonging to system administrators, database administrators, security professionals, and so on. Moreover, these identities are often the target of spear phishing attacks. With all the things that these accounts can do, they are prime for privilege abuse and must be monitored closely. Gurucul collects this data in order to understand what the most privileged accounts are doing.

Directories

The most common is Active Directory (AD) for on-premises; this source may also include LDAP directories or other directory servers.

Log sources

These sources track every activity that goes on in an environment. The data can be collected from log aggregators, SIEMs, syslogs, databases, applications and individual end systems. By collecting this information, we get every bit of activity that is taking place throughout the enterprise, and who those activities are attributed to.

Using Identity as a **Threat Plane**

Chief information officers, chief information security officers and security managers should favor [User and Entity Behavioral Analytics] vendors who profile multiple entities including users and their peer groups, and devices, and who use machine learning to detect anomalies. These features enable more accurate detection of malicious or abusive users.

— Avivah Litan, Gartner, September 2015

Defense-in-depth systems

Systems such as DLP, anti-malware, IDS, IPS, firewalls, SIEM, etc. raise alerts when they find suspicious activity. We want to include those alerts in our analysis and correlate it to specific network identities.

Intelligence sources

This information typically comes from external sources that are tracking a broad scope of indicators of compromise and threat patterns. In addition, Gurucul has built our own out-of-the-box algorithms to quickly detect risk situations. We overlay these algorithms and libraries to the identity information to help detect internal threat activity that is taking place throughout the enterprise, and who those activities are attributed to.

Gurucul gets very fine-grained with the data. We build our machine learning algorithms to accommodate 254 different attributes around identity. What's more, the architecture is open to various structured and unstructured data sources from the cloud or from on-premises systems using a flexible meta data framework.

While it might seem that combining all these sources of data would yield a dataset that is too large and too complex to work with, Gurucul actually distills the data by normalizing it to a standard format and scrubbing it to ensure the data is clean. Then we feed it all into a big data application to derive access analytics and store historical data into a data lake based on Hadoop. Gurucul applies machine learning algorithms, including self-learning and training behavioral profile algorithms, which look at every new transaction and risk scores it. Using clustering and outlier machine learning makes suspicious behaviors stand out from other benign activities. And more importantly, we are training our analytics engines about users' behaviors over time.

But we don't stop there.

Baselining Behavior to **Dynamic Peer Groups**

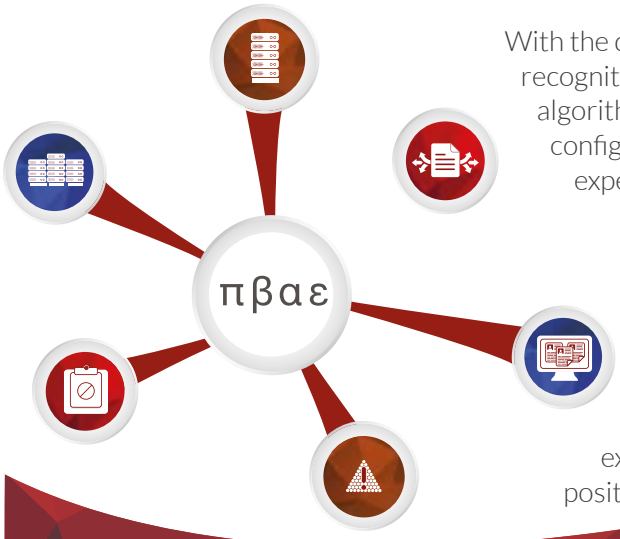
This is, hands-down, the most sophisticated example of behavioral analytics we have seen to date. While they are not the only player in this space, their product is well thought-out and it really works well.

— SC Magazine, July/August 2015

The directory services that enterprises use – Active Directory and similar products – tend to put people into static groups to facilitate access provisioning. People are grouped in various ways—by department, by job roles, by location, and so on. This information is somewhat useful for analyzing identities, privileges and activities. However, these groups might be poorly maintained and too plentiful to be useful. As time goes by, some enterprise directories actually grow to contain more groups than there are people in the organization. What's more, people are often left in old groups, even though they've moved on to other roles within the organization.

Gurucul goes beyond static peer groups and allows the analytics engine to process all the source data described above, or one specific data source, against dynamic peer groups. These groups define people according to the types of activities they typically perform, as well as the types of identities and privileges they hold. Dynamic peer groups yield a much tighter clustering of behavior and much more accuracy in highlighting outlier activities in behavior patterns. Thus when someone is abusing the privileges of their digital identity, the behavior really stands out. At this point, we have tremendously reduced the chance of a false positive alert often seen with static peer group analysis.

Hijacked accounts that use legitimate credentials are normally associated with espionage or other types of advanced attacks. They are quite a problem for organizations because they can't be detected through traditional security mechanisms, yet they can do quite a bit of damage. With Gurucul's tight peer groups, an intruder has access to compromised account, assume the credentials, login and act like the legitimate user does amongst the multiple dynamic peer groups that user has been put in. All this just to remain undetected by machine learning clustering and outliner algorithms that model good behavior to detect unknown bad. That's a pretty big challenge and there's little chance of it happening.



With the dynamic peer groups, what we end up creating is pattern recognitions and normal baselines of user behavior. To make the algorithms even more effective during predictive analysis, they are configured to understand and learn time-based norms, which allow for expected changes over time. For example, a department manager takes a leave of absence for a few weeks. During that time an employee within the department temporarily assumes the manager's role and responsibilities. This person's digital identities now have new, company-sanctioned privileges for a short time. Ordinarily his new activities might raise an alarm, but Gurucul's dynamic peer groups and time-based behavioral baselines can accommodate these approved exceptions without driving up the risk scores or creating false positives.

Predicting and Isolating **Malicious Insider Behavior**

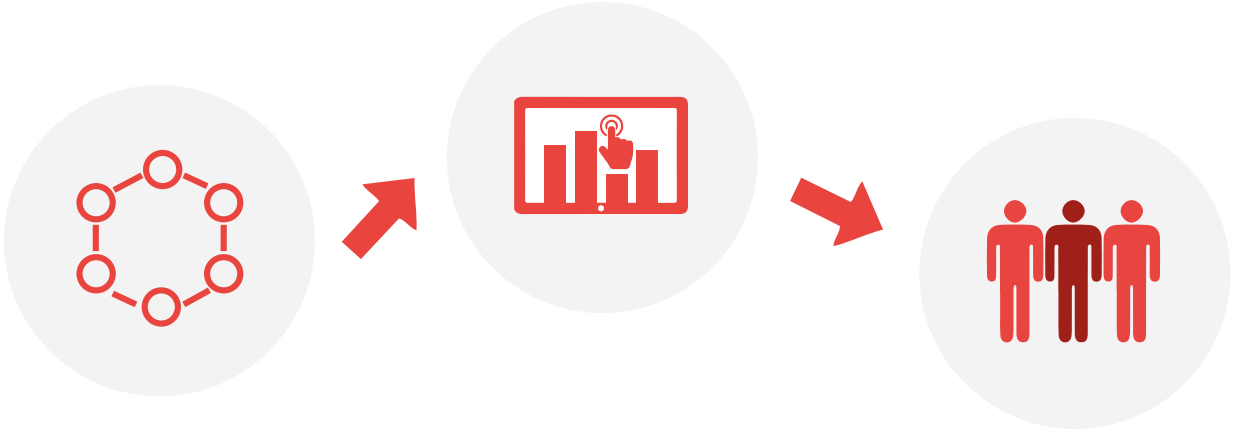
According to the CERT Insider Threat Center, a malicious insider threat to an organization is a current or former employee, contractor, or other business partner who has or had authorized access to an organization's network, system, or data and intentionally exceeded or misused that access in a manner that negatively affected the confidentiality, integrity, or availability of the organization's information or information systems.

Time-based norms are very important to the next step in the process. All of this behavioral baseline information is fed into predictive machine learning algorithms to do extensive risk modeling. The resulting normalized risk scores are highly accurate and allow the organization to pick up the malicious insider activities. Here are just a few scenarios that become very easy to see by the risk scores Gurucul produces:

- A sales employee is planning to leave the company to take a job with a competitor but she hasn't yet made her intentions known. Before she resigns she wants to gather as much competitive information as possible. This is information she would normally have access to for her job role, but her new actions of downloading more data than usual is seen as suspicious. When compared to her dynamic peer groups, she is shown to be the only sales person doing this volume of activity. The risk models give her identity a high score for the recognized anomalies, indicating her actions need to be investigated.

- A set of login credentials belonging to a customer support employee of a financial company in the North America region has been stolen in a phishing attack. The thief logs into the on-premises support application account with all of the authority of the genuine user. The intruder accesses many of the same contact and support account records as the legitimate user normally would do; however with more frequency and data downloads for a normal customer support analyst. This activity is quickly spotted as something outside the norm, and the company is able to suspend the account until an investigation is conducted.

- A contract worker has been working for a certain company for many years. He works at home and performs his job via VPN for remote access. For all practical purposes, he acts like an actual company employee, with the access rights and privileges of an employee. And then his contract is terminated. He is de-provisioned in the HR systems but his login credentials for Salesforce.com are overlooked for termination. He still has access and continues to login to siphon off customer information to provide to a competitor. Gurucul can spot this behavior through analysis of the current HR data correlated to the contractor's current activities.



Insider Threat Detection and Deterrence **The Self Audit**

There is one extra step that Gurucul uniquely provides for insider threat and detection and deterrence: the self-audit. Users are provided a self-audit much like a credit card statement to view their own risk-ranked anomalous activities, identities, access, devices and other key data points in an easy to use web portal. Developed with a customer CISO and now gaining popularity with other CISOs, it co-opts users into a collaborative relationship to monitor and protect their identities. When users detect an anomaly the false positive rate is very low and the context provided is richer and faster than IT can provide. The visibility of what data sources are monitored and analyzed against dynamic peer groups also acts as a deterrent against insider threat.

Conclusion

By 2017, at least 80% of companies that adopt UBA will achieve at least a 5-to-1 ROI within one year of implementation.

— Avivah Litan, Mark Nicolett, Gartner, Market Guide to UBA

The traditional approach to security for an enterprise environment is to utilize an indicator of compromise (IoC)—a "known bad" pattern or signature of activity. Tools take an IoC and look for signs of it in every applicable step of the kill chain. There are limitations to this process; most notably, not all behavior with a malicious intent has a known bad signature to it. Many malicious behaviors, in fact, are unknown, and this is especially true for the insider or external intruder abusing the privileges of legitimate access credentials.

The most effective way to pinpoint the presence of insider threats, without creating a lot of false positive alerts, is to overlay user activities with user identity intelligence, cluster identities into dynamic peer groups, create time-based behavioral baselines, and continuously learn what is acceptable behavior in order to spot the unacceptable behavior. It takes a combination of the right data sources, sophisticated machine learning and perceptive data science to pinpoint truly aberrant actions that are good indicators of misuse of assigned privileges.

The methodology Gurucul provides includes three key parts. First is identity access intelligence (IAI) to risk-rank access to reduce excess access often seen with legacy identity access rules and manual procedures. This tightens up the identity access threat plane; in one case a customer reduced 83% of accounts and entitlements. Second is user behavior analytics (UBA) based on dynamic peer groups for improved clustering and outlier analysis and the understanding of time-based norms to reduce false positives using self-learning and training predictive machine learning algorithms. Third is to bring users into a collaborative relationship with IT security to protect their identities via self-audits to review risk-ranked anomalous behavior and access analytics. The combination of these three creates a solution for insider threat detection and deterrence.

We encourage you to learn more about what SC Magazine calls " the most sophisticated example of behavioral analytics we have seen to date" by visiting the Gurucul website. If you're ready to move ahead with a proof of concept, contact us at info@gurucul.com.

About Gurucul

Gurucul is changing the way enterprises protect themselves against fraud, insider threats and external intruders both on premise and in the cloud. The company's user behavior analytics and identity access intelligence technology uses machine learning and predictive anomaly detection algorithms to reduce the attack surface for accounts, unnecessary access rights and privileges, and identify, predict and prevent breaches. Gurucul technology is used globally by organizations to detect insider fraud, IP theft, external attacks and more. Gurucul is based in Los Angeles. To learn more, visit us at www.gurucul.com and follow Gurucul on LinkedIn, Facebook and Twitter (@Gurucul).