DETECTING SOPHISTICATED ONLINE ATTACKS WITH STREAMING ANALYTICS

RSA Silver Tail Operationalizes Big Data to Provide Real-Time Protection from Business Logic Abuse Threats

Security professionals are aware that cyber criminals have increasingly sophisticated weapons at their disposal for maneuvering through online commerce systems and stealing information. Traditional firewalls, IPS/IDS, and web application firewalls (WAFs) do little to help online businesses understand the behavior of website visitors. Instead, they narrowly focus on the network and server exploits only. Because of this gap in technology, cybercriminals are evolving to exploit a new attack vector known as business logic abuse, which results from criminals exploiting flaws in the business functionality of a website, such as shopping cart, logins and file downloads. These costly threats are growing significantly, making the job of IT professionals extremely complicated. Attackers generally use legitimate webpages to launch their schemes, once they have gained entry into a site.

One example of Business Logic Abuse seen during a recent attack on a leading eCommerce site involved the use of coupons on items previously on sale. In many instances, if a sale expires but the item is still in a shopping cart, the price is saved for that cart. In the case of this retailer, coupons couldn't be applied to sale items, but once the sale expired, the shopper could use a coupon. Hundreds of carts were filled with sale-priced items, at which point all the coupons were executed and the transactions pushed through at enormous discounts.

According to a recent survey of IT professionals from the Ponemon Institute and RSA Silver Tail[1], 90 percent of businesses report losing revenue to business logic abuse, and 88 percent of respondents say this type of attack is the most important security issue facing their companies. However, 74 percent of respondents also say they have difficulty telling the difference between "real" customers, and the criminals accessing their websites.

1 Ponemon Institute and Silver Tail, 2012 Web Session Intelligence & Security Report: Business Logic Abuse Edition, October 2012: http://buzz.silvertailsystems.com/PonemonStudyFullReport.html.

The Five Layers of Fraud Prevention

As Gartner explains in its recent report, *The Five Layers of Fraud Prevention and Using Them to Beat Malware*, "…no single layer of fraud prevention or authentication is enough to keep determined fraudsters out of enterprise systems. Multiple layers must be employed to defend against today's attacks and those that have yet to appear."[2]

RSA Silver Tail is unique in its ability to detect anomalous website behavior by utilizing layers 2, 3, and 5, all in one platform.

The Traditional Approach

Existing solutions for detecting and analyzing online criminal behavior usually identify either pre-authentication threats (infosec products) or post-authentication threats (fraud products) – but not both.

Prevention

Armed with knowledge of the types of attacks their sites may face, IT professionals can use firewalls or rewrite software to block the traffic they assume to be dangerous. However, the business logic abuses favored by today's cybercriminals are beyond the scope of preventive solutions. Fraudulent activity committed by abusing shopping cart logic, or by manipulating rebate or gift card actions, happens alongside legitimate traffic.

Detection

The challenge of detecting anomalous activity in real-time requires gathering various "big data" sources and correlating them to understand user behavior. However, current methods of detection fall short of this goal – individually, they examine only pieces of the behavior puzzle, not the entire picture. Web application firewalls (WAFs) can examine transaction signatures, but will only block traffic previously identified as a potential threat. Security information and event management (SIEM) solutions use limited data in log files to seek out behavior that seems anomalous. These solutions can only identify broad trends or rule violations; they can't correlate anomalies to individual user sessions.

Investigation

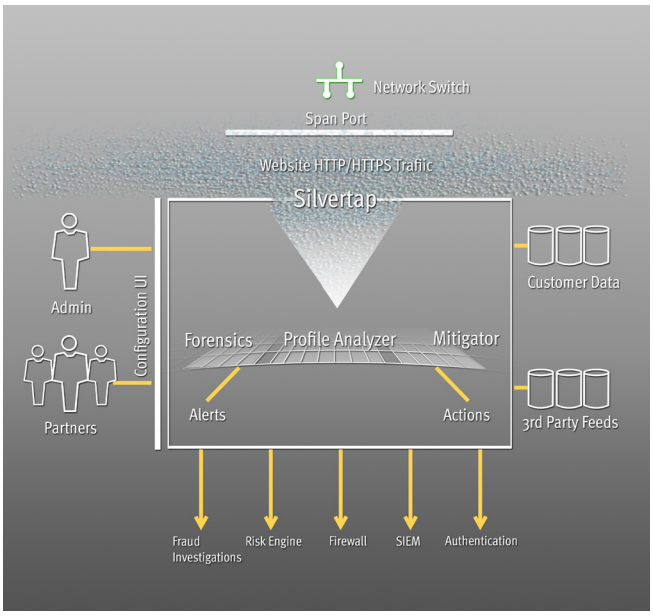
Flagging certain transactions for review is a commonly used tactic for detecting fraudulent activity – and it served its purpose in the days when cybercriminals used exploits such as brute force attacks or stolen credit cards. However, the data that information security teams need to identify suspect transactions is often scattered among multiple systems. For example, website logs are with web admins, IP level information is with the network team, and account information is with operations. And since the team can't see a user's specific session on the website, they can't compare the incidents they are flagging to the actual use activities, so the investigators gain no insights on behavior. A more centralized approach that allows atomic analysis is needed.

2 Gartner, "Gartner Says 15 Percent of Enterprises Will Adopt Layered Fraud Prevention Techniques by 2014," Oct 2, 2012: http://www.gartner.com/newsroom/id/1695014.

The RSA Silver Tail Approach

Sessionization of Data

One of the key differentiators of the RSA Silver Tail solution compared to other web analytics suites is its ability to "sessionize" a user's clickstream. What this means is that every click a user makes on a website to navigate from login to checkout/logout is grouped together so that the entire stream can be utilized for analysis on a user to crowd, user to user, and individual stream basis. This is achieved in real-time by a number of innovative components, beginning with Silver Tap.



Silver Tap is software installed on a server in the data center, which sniffs packets from a SPAN port configured to mirror traffic from a web server. Architecting the deployment in this manner allows Silver Tail to be completely separate from the web server and have zero impact on latency or increased risk for connectivity issues for the end user. Silver Tap then filters and reassembles the packets to extract the TCP payload. It then parses any of several protocols, including HTTP and HTTPS, in order to extract important attributes and create metrics about the traffic at all levels of the protocol stack. Once this is all completed, the resulting messages are placed on a message-bus for distribution to the other core elements of Silver Tail performing tasks like logging, scoring, and reporting.

During the configuration of Silver Tap, Silver Tail defines items such as the parameters that define a session cookie, the login and logout page URI, UID, SSL certificates and/or HSM information, and items such as PII that need to be hashed or truncated so as not to be captured for compliance reasons.

Silver Tap immediately starts to inspect all traffic once installed, scrubbing away not only the items that have been configured to be ignored, but also all data non-essential for its purpose, such as image or video content, to limit its footprint. The data then goes through an extremely efficient compression, indexing, and storage mechanism so that the Silver Tail solution can now perform all of its analysis, reporting, and alerting functionality.

The data being stored is also one of the key differentiators of the product, as it stores and intelligently breaks apart for easy search and analysis:

1. 6 TCP transport connection attributes including IPs, ports, and the timestamp down to 100s of microseconds from EPOCH
2. 5 TCP Request/Response attributes including the Range of TCP data packet ordinals relative to the TCP connection
3. 7 SSL/TLS Transport connection attributes including the SHA-1 fingerprint of the server certificate
4. 3 SSL/TLS Request/Response attributes including the total size in bytes of the TLS record

Top 10 Retailer Halts Fraud and Boosts Productivity with RSA Silver Tail

A top 10 online North American retailer, generating more than US$1 billion in annual revenue from online operations, was able to immediately begin identifying and halting fraudulent activity on its e-commerce website and improve productivity using the RSA Silver Tail Forensics solution.

According to a Total Economic Impact report recently prepared by Forrester Research[3], the retailer's security team was spending too much time on manual processes to identify malicious activity. Its existing anti-fraud systems couldn't detect fraud when IP addresses were being processed through a proxy or when domain owners were being masked, forcing the security team to laboriously collect and reconstruct log files.

Within six hours of the deployment of the RSA Silver Tail platform, the retailer detected malicious behavior on its websites, including site probing, targeted directory attacks, and site scraping for pricing – which traditional fraud systems were not able to identify.

In addition – for the first time – the retailer's security team gained visibility into entire web sessions and could see every click by every IP address or user on its websites. The team has been able to avoid lost revenue from criminal activity such as fraudulent gift card wins, phony rebates, and fake products posted to the retailer's websites. In addition, the retailer realized a 72 percent ROI over a three-year period, and recognized over $4 million in benefits in the first year alone.

5. HTTP attributes:
    a. HTTP message ordinal (1-based, listed twice) relative to the start of the TCP cxn, e.g., 4
    b. Method (GET, POST, HEAD, etc.)
    c. Status code
    d. HTTP protocol version
    e. URL
    f. HTTP headers
    g. GET and POST arguments
6. Multiple custom extracted attributes configurable by the administrator

This allows fraud analysts to run analysis or rules on any piece of data pertaining to the traffic on their website, and for the automated Scoring Engine to run its bleeding-edge Streaming Analytics (defined later in this paper) upon it as well, in real-time.

Threat Scoring

The RSA Silver Tail scoring engine builds statistical models based on the crowd behavior, which represents the good behavior. Each use session is then compared to this model and analyzed as to whether it fits the good behavior or falls under anomalous behavior. This real-time Streaming Analytics is performed without any latency. Further, this scoring engine is able to store an in-depth traffic profile composed of hundreds of attributes in memory so that the following scores can be calculated for each click.

– **Velocity:** The attacks of automated agents are commonly characterized by unusually fast page transitions. These rates of page transitions, or velocity, are among the scores used to identify potentially fraudulent user behavior. Measuring the speed of each transition across the population of all visitors to the site, and then measuring the divergence of a particular visitor from those norms, creates this score.

– **Behavior:** This score identifies unusually frequent activity, which is often correlated with computer-based attacks or directed attacks that do not follow the usual website access pattern. It is a measure of how a session's navigation sequence is different from all other sessions. The navigation sequences are measured in terms of the frequency of page transitions. The frequency of each page transition is compared to the norm for the frequency of the same page transition across all sessions, and a score is created to quantify how far from the norm a particular session is.

– **Parameter:** Parameters submitted through GET or POST are analyzed for rarity. Parameters with highly rare are given high scores. If a parameter is submitted that is normally not present, the engine is able to detect these anomalies and score based upon them.

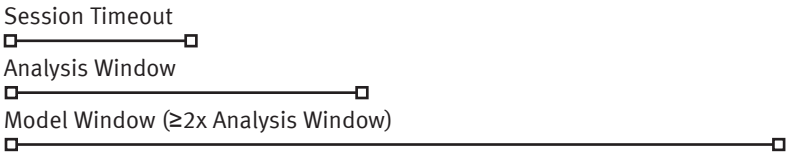
– **Profile:** These scores follow the same logic as the behavior scores. However, these are then compared against the user's history over a period of time. This period of time is mostly defined by the admin's choice of storage.

– **Man in the Browser (MiB):** MiB attacks are performed by malicious code operating within the user's browser or on the user's machine, and operate within the time span of a user session. This score is calculated by considering the pattern of page transitions, the geographical location of all of the actions associated with the user, and the speed of actions observed from the user, alongside multiple other elements of the transaction.

– **Man in the Middle (MitM):** MitM attacks are characterized by situations in which the attacker has gained access to, and is exploiting, the web session of another user. For example, while a user has logged into his/her banking website, a malicious user gains access to the session from another system. This cyber criminal performs cash transfers, while the user is completely unaware of any of these actions. Such attacks are detected by concurrent overlapped access to a site typically from different IP addresses by nominally one user, over a minimum number of page accesses.

3 Forrester Research, The Total Economic Impact of Silver Tail, August 2012.

Once this pattern has been identified, the score is subjected to a series of score discounts that include whether the accesses report the same user-agent, whether the transactions are coming from geographically separated IP addresses, and whether the accesses include a login page. These score discounts help to identify and downplay less-suspicious and less-consequential IP-address changes. Transactions that continue to have high scores after being subjected to these discounts are considered likely Man-in-the-Middle attacks.



The Decaying Data Model

The baseline data used while generating scores is organized into Data Models. A data model consists mainly of frequencies of tracked items, e.g., the number of transactions containing a certain attribute with the value “John Doe.” A stepwise first-order recursive linear filter configured by two parameters controls the duration of the model frame.

During the configuration, the following attributes can be set.

Shelf life should be set for the frequencies and other numeric model values, which represent the amount of time data should be collected and used to update the current profile.

Decay factor is then configured, which works by setting an amount of time for a mathematical degradation of the data’s strength for comparison over time. This is done so that the most recent data is given the most strength. The best practice for defining both attributes depends on the computational resources available and the dynamic characteristics of the site. Once a counter’s decay has reached 0, its record is pruned to prevent unbounded growth of obsolete information.

The Analysis Window

The analysis window is a sliding window of time representing the portion of traffic to be analyzed and compared with the current data model for scoring threats. In real-time usage, the analysis window ends at the present moment. Session scores are calculated as the sum of all transaction scores within the analysis window. To ensure that we are comparing the counters from within a session to each individual session, this window should be configured as close to the average session time of a user on the website as possible.

The model frame and the analysis window are largely independent. The model frame determines the population data that is used to calculate transaction scores within the analysis window. Changes to the data model due to the decay of data in the model frame will have little effect on the analysis window calculations. However, there are requirements for the time periods in that the model frame must be larger than the analysis window, which should always be larger than the sessions themselves.

The duration of the model frame must be at least 2x that of the analysis window to satisfy the Nyquist sampling criterion. As a best practice, the model frame should be long enough to average out the hourly, daily, and weekly business cycles of the website, which means it should be an integral number of weeks.

Traditional Solution vs. RSA Silver Tail		
Requirement	**Traditional Solution**	**RSA Silver Tail**
Visibility	Limited view	Holistic view of traffic
Individual Behavior Analysis	Not available	Profile Analyzer Behavior Modeling Crowd -> User Scoring User -> Historical User Scoring
Improved Workflow	Multiple teams Multiple data sources Incomplete data	One tool Holistic view of traffic
End-User Experience	Complex file download	Zero impact until malicious behavior identified
Dynamic Web Sites	Disparate teams unable to update approaches	Dynamic modeling is self-learning and adaptive
Simple Installation	Code install on page Customer software New box in customer path	SPAN Port No customer impact

Conclusion

Catching cybercriminals in the act requires IT departments to look deeper into their web traffic, to examine many more sources of information about web visitors, and to view entire web sessions to determine what website behavior is typical for their website and what is not. Traditional approaches to detecting and preventing fraud don't paint a complete picture of website activity and don't connect the dots between various sources of data about online activity. RSA Silver Tail provides this visibility and empowers its customers to stay one more step ahead of criminals looking to exploit their online presence.

About RSA

RSA is the premier provider of security, risk and compliance solutions, helping the world's leading organizations succeed by solving their most complex and sensitive security challenges. These challenges include managing organizational risk, safeguarding mobile access and collaboration, proving compliance, and securing virtual and cloud environments.

Combining business-critical controls in identity assurance, data loss prevention, encryption and tokenization, fraud protection and SIEM with industry leading eGRC capabilities and consulting services, RSA brings trust and visibility to millions of user identities, the transactions that they perform and the data that is generated.

STS WP 0513

www.emc.com/rsa

EMC²

RSA