What cyber security can learn from algorithmic trading?

IT security teams are under more pressure than ever before. Criminals, hackers, competitors, users, business partners, regulators and the changing face of IT are individually and collectively creating pressure points. Some are directly attacking IT systems. Some are acting as an enabler both wittingly and unwittingly. Meanwhile government and regulators are placing additional requirements on how IT security must work.

All of this is combining to create a continuous stream of successful attacks and data breaches. Each incident has the potential to not only disrupt a company but cause severe brand damage and threaten the survival of businesses.

Algorithmic (algo) trading systems exhibit many of the same challenges as those facing IT security. There are very large volumes of data that must be captured, analyzed and then acted upon in real-time. Machines are programmed to act autonomously based on market activity. Everything happens quickly. The potential for failure is always present and the impact of failure can be business threatening. Compliance and legislative requirements are strict and uncompromising.

Today's financial markets businesses have fully embraced high speed autonomous machines and artificial intelligence to gain a competitive edge. However, they had to take a wholly different approach to IT in a number of areas to assure and safeguard the delivery of the business. One major difference is their use of network traffic as a trusted and effective way to monitor what the machines are actually doing. Deep analytics of network data can extract unique insights into user, machine and network activity in real-time. It allows them to detect and react to technology issues and anomalies quickly as any failure can cause massive financial losses in a matter of seconds. They also use the same approach to detect and react to cyberattacks. Both of these deliver lessons for IT security. In this article, we look at the challenge for IT security and a better way of doing User and Entity Behavioral Analytics.

Technology is part of the problem

IT security teams are unable to effectively evolve at the same pace as the threat they face is evolving. This leaves them trying to deal with the threat using the technology that they already know and have installed.

Deploying patches to operating systems and applications is typical of the problem. History has shown that patches can break business applications. When you are only responsible for 200-300 servers you might have a window where you can deploy a patch and deal with any problems. When you are patching thousands, or tens of thousands of servers and applications a single patch error can cause catastrophic downtime. This is why IT departments test patches. The problem is that hackers are using attacks that don't provide for that time.

IT security teams have deployed lots of agents and software to capture data from around the IT estate. On top of this they are now logging and capturing information from multiple layers of devices. All of this data is then looked at by the security analysts. To reduce the volume of data they have been using business intelligence (BI) and analytics solutions.

The problem they face is that they now have too much data to work with effectively. The challenge is how to refine the volume of data to identify real attacks from those that are simply annoyance and therefore signal noise. It also takes time and that gives attackers a window of opportunity.

The role of User and Entity Behavioral Analytics

One of the biggest technology shifts has been the introduction of user and entity behavioral analytics (UEBA). This approach monitors what an individual, an application or a machine is doing. It seeks to establish patterns of behavior that can be classified as normal and safe.

It creates patterns from activities such as when the user logs in, where they log in from, the speed with which they input their password, the way they access systems along with the data and systems that they access. The latter is interesting because it allows security teams to track how often a user accesses a data set and spot changes to that frequency.

This is no longer simply about the users logging into machines. The vast majority of connections are machine to machine, and modern hackers can use algorithms resident on compromised hosts that automatically act based on detection of triggered events. They use locked in credentials that rarely change. The reason for this is that those credentials get picked up in code and changing them can cause systems to stop working. This means that username, password, and login activity can only be considered to be part of the pattern. To complete the full picture, we need to associate the specific network activity with the specific machine and user activity. Today the security analyst has to do that manually jumping between SIEMs, firewalls, directories, DNS, PCAP and flow data records trying to match up context and time. It is a challenging and often time-consuming exercise.

Algorithmic trading environments use real-time network UEBA

Cybersecurity solutions tend to be focused on the relationship between the user and the machine. This information is often gleamed from the harvesting of the appropriate log files analyzed within the SIEM. Unfortunately, log files do not contain information relevant to machine activity on the network. The algo trading world however had to solve a similar problem. It needed to monitor and detect in real-time any trader, their algos, running on any machine causing certain patterns of trading activity on the network and shut it down if it breached certain rules, patterns or policies.

Their solution approached the problem from the perspective of the network by capturing all traffic that flows across it. By doing deep analysis and machine learning on the raw packet data it can extract the necessary information that relates network activity and patterns to the machine source and responsible user. This is what was needed to get the full picture. The surprising fact is that all of this information is possible to get from the packet data flowing in the network. It is just hard to get it. Additionally, the network provides a forensically verifiable source of data that can be trusted. Machine to machine communication can also be used as a trusted audit trail and a true record of what actually happened. This is critical when an analyst is presenting or defending their incident reports to an external regulator where major fines or reputational damage are on the line.

The key to network UEBA is being able to analyze the data correctly. This starts by ensuring that the capture process preserves the fidelity of the data. One of the keys to that fidelity is being able to trust the start and end point of the traffic. Rather than have the problem of stolen user credentials, the network traffic contains an absolute reference, either IP address or MAC address that identifies the start and end point of the traffic.

Once the data has been captured it has to be decoded and transformed so that the context of the actions taken by the machines can be easily understood. As part of that it is important that the activity can be related to the machine, user or application that initiated it.

The resulting system brings together every machine action with accurate and trustworthy timestamps. This ensures that sequence and causality can be determined. It also builds a data map between network activity, machine activity and the entity which initiated the action.

There are a number of ways that this data map can be used. It provides a forensic trace as to the actions and results that occurred. By putting them on a timeline it is possible to see if the actions were in isolation or whether they were part of a coordinated set of actions across the network. The flow of data also shows the impact on other machines and systems. Machine learning can be applied to the information to detect unusual activities and patterns. A key to the effectiveness of the machine learning algorithms is the quality of the underlying data.

Another benefit here is that the network UEBA approach is non-intrusive. Tapping the network has zero performance impact and it does not affect or make any changes to the data in transit. The network UEBA approach does not require special agents to be resident on the host machines which might slow down their actions. This is important in the trading world. It will also be important in the future cyber world where internet of things will interact autonomously with smart machines. It will not be possible to place agents on all these endpoints and the network will be the only possible way to track and detect activity.

What does this mean for cybersecurity?

Cybersecurity teams are overwhelmed by multiple data sources that do not contain the relevant, context and correlations to quickly detect, hunt and react to attack. This means data has to be constantly transformed, compared and verified. This not only introduces delays in the data analysis but also means that real-time interpretation and action is impossible.

The way that the data is captured also creates a challenge for forensic teams who will want the data in order to prove legally that an offense has occurred. This is part of cybersecurity that is often overlooked. The approach from algo trading is to preserve the integrity of the data and the precise timeline of every machine action which ensures that it meets court standards.

Cybersecurity teams also struggle to identify the secondary impact of an attack. For example, a machine is infected with a virus. Which machine does it infect next? How many machines does it infect? How does the infection spread across the network? The algo trading focus on the correlated data map exposes the way an action or infection moves through the network. This not only provides an accurate view of what has happened but also allows a cybersecurity team to predict the impact of another attack. This helps them build new defenses to mitigate or prevent the spread of a future infection.

One of the biggest benefits here is that the algo trading approach doesn't separate user action from machine action. It sits inside the network so as soon as there is activity on the network it is capturing it. This means that a hacker or insider threat, malware or rogue software, the actions are captured, recorded, replayable and available as a forensic record.

Conclusion

The use of UEBA by cybersecurity teams often limits their ability to see what is really happening on the network. While they are rightly focusing on rogue or unexpected behavior, they are capturing it in such a way that misses the critical relationship with activity on the network. More importantly, they will never be able to detect and act on threatening activity in real-time, using this approach. This is the future of cybersecurity as hackers turn to automated computing, data science and AI technologies to advance their agenda. Our ability to distinguish intentional cyberattack from unintentional technology mishap will blur. This means our need to see, act and protect in real-time will become imperative.

The algorithmic trading community has lived this transition over the past decade. It had to innovate new approaches and solutions to assure and protect its business in real-time. It discovered that the network was the best and most reliable place to watch over and detect what the machines were doing and to safeguard the business from all potential risks, both intentional and unintentional.