

Assumption Busters Workshop – Malicious Behaviors

Background: The U.S. Federal Cyber Research Community is conducting a series of four workshops designed to examine key assumptions that underlie current security architectures in cyberspace. These “Assumption Busters” meetings are designed to create the environment for the development of novel solutions that are based on a fundamentally different understanding of the problem and creates a stronger basis for moving forward on well-founded assumptions.

In the next few months, we will assess the problem for each assumption as well as any potential weaknesses that transcend the four categories: defense-in-depth, trust anchors, malicious behaviors, and cloud computing.

Introduction: The third “Assumption Buster” workshop was held on 20 June 2011 at the FDIC’s L. William Seidman Center in Arlington, VA to focus on the assumption that “abnormal behavior detection finds malicious actors.” Over 35 participants from academia, government, industry and the research community attended.

The first half of the day was devoted to presentations from the financial community (FDIC, Federal Reserve System, World Bank, SEC, and FICO) about their efforts to reduce losses due to fraud. Financial services companies have been fairly successful in establishing fraud detection analytics, based on abnormal behavior identification, which identify financial transactions that seem out of the norm for a particular financial services customer.

The afternoon session provided 12 selected attendees an opportunity to present highlights from submitted workshop papers. Each presenter gave a 5-10 minute summary of their current research activities followed by a brief question and answer period from all the attendees.

This paper identifies key themes that supported or challenged the assumption along with potential next steps to apply these techniques in the cyber domain. These discussions validated the assumption and identified near- and long-term directions for further research and exploration.

Workshop Assumption: “Abnormal behavior detection finds malicious actors”

Participants overwhelmingly supported the assumption that identification of abnormal and normal behavior are both key to identifying and stopping malicious actors. Financial fraud detection, for example, works on the assumption that malicious fiscal behavior is a subset of abnormal behavior. If the fraudulent user mimics the financial behavior of the authorized user, these methods do not work. Detection methods do not assume, however, that malicious behavior is automatically distinguishable from *unusual* behavior on the part of authorized users. The fraud detection algorithms use the financial services customer’s history to build a profile of “normal” transactions and develop thresholds for unusual behavior. Depending on the sophistication of the algorithms, the system might also include specific information on the financial participants (initiator/recipient), physical devices and locations involved in the transaction, or other parameters. The volume of transactions allows for reasonable thresholds to

be established. Fraud detection methods rely on strong models of normal behavior, and/or known criminal behavior characteristics. Typically, each transaction is scored and the financial institution approves or rejects the transaction.

The group spent considerable time understanding the approaches taken to identify abnormal behavior in the financial sector, in order to understand the extent to which these procedures might be more broadly applicable to cyber security.

Highlights of Discussions from the Financial Sector

The FDIC sees credit card and ATM attacks increasing. Fraud is evolving and becoming more sophisticated. Precursors to malicious activity are frequently observed 24 hours prior to attack and phishing incidents. Insider threat is potentially as damaging as attacks from the outside. Insiders are defined as someone with privileged access.

The Federal Reserve is not overly proscriptive in response. Banks and financial institutions (and other organizations) must identify their own risks and define their own solutions. The Federal Reserve works to ensure that institutions have a good risk management plan. Financial organizations can outsource security but cannot outsource responsibility, therefore due diligence is needed when selecting and managing service provider relationship.

Financial institutions are not required to report all cases of fraud. While there is no central repository for reporting industry fraud loss data, institutions share information about trends and losses via industry associations such as the American Bankers Association (ABA), BITS Financial Services Roundtable, and the Financial Services Information Sharing and Analysis Center (FS-ISAC).

The Federal Reserve believes that single factor identification is an inadequate control mechanism for high risk transaction and encourages multi-factor identification. The White House (Howard Schmidt) is working on an identification capability for multiple uses.

Over half of the phishing attacks are against financial institutions, and target both the institution and its customers. For example, hijackers steal bank credentials, impersonate the commercial entity, initiate unauthorized transactions, and send money to a transfer entity, known as a mule. Credentials are often hijacked through a combination of social engineering and use of malware. Financial institutions need to monitor the last two activities - transactions and fund transfers --- for abnormal behavior/malicious activity.

From an SEC perspective, there are primarily two types of abnormal behaviors of interest -- bad data and fake trades. Their focus is on detection and mitigation. Emphasis is placed on gathering information to look for abnormal behavior on the perimeter and within the network (corporate and production). The SEC's goal is to learn the source of the attacks and the vulnerabilities of the systems. Also, emphasis is placed on survivable or self-healing systems.

FICO, formerly Fair Isaac, provides credit ratings and scores individual financial transactions in terms of risk and/or fraud. FICO applies analytics to provide timely information on credit and

debit card transactions, e-payments, medical claims, and other financial transactions. Their process uses predictive analysis (including profiling and adaptive models) to determine a “trust” score, based on proprietary factors, for an activity. FICO starts with a rule base and evaluates an activity to look for things that don't make sense. There is a combination of profiles, models, and rules to do the scoring. The monitoring can be manual or automated, supervised or unsupervised, but there is no personal data exchanged with FICO.

FICO's transaction scoring is provided to the financial institution adjudicating the activity. In most cases point-of-sale (POS) transactions are approved even if the score indicated a high potential for fraud. This is because of the concern for false positives. Financial institutions do not want to “spook” a real customer. Financial institutions adjudicate potential bad transactions after the fact. Each institution must determine the level of risk they will assume in accordance with their business processes to allow or disallow the transaction.

FICO indicates that there is a need for several types of models so that the bad actors cannot learn the rules of the fraud detection algorithms. Also, the fraud detection system must adapt over time.

With respect to cyberspace activities, FICO believes that, using their approach, they could identify bad actors and compromised devices using variations of their current rules, models, and profiles.

Bottom Line:

Abnormal behavior is a good indicator of malicious actors. However, normal behavior is dynamic and can appear abnormal, so detection of abnormal behavior must be dynamic. But various patterns of normal or typical behavior must be established in order to understand and identify abnormal and malicious behavior. False positives can be a challenge, and malicious actors hide within known bounds of normal behavior; false negatives should also be sampled for malicious behavior. Single factor identification may be insufficient to identifying malicious behaviors.

Ideas that Support the Assumption:

Participants in the financial services sector do a good job of detecting suspicious activities. As techniques have improved, the number and values of fraudulent transactions have decreased. However, hackers are constantly adapting to their techniques to circumvent the latest detection and alerting mechanisms, so the systems require constant adjustment.

Abnormal behavior can find unsophisticated hackers. Normal behavior changes dynamically and some effort must be made to look at these transactions too.

If financial fraud is detected, the intention is usually pretty clear, in key elements: the link between account and attacker is apparent, and the data collected is usually very focused and relevant to the attack. Within cyberspace, attacks are potentially less clear in terms of purpose, actors involved, and the outcome.

The current financial system can perform risk authentication (scoring) on every financial transaction. FICO and other scoring services enable institutions to consider whether the transaction is risky for their customer and, if so, the decision to approve or deny the transaction rests with the financial institution.

One analytic approach for the identification of malicious behavior involves the detection of protocols thru encrypted channels to determine what is normal. Given hackers' sophisticated attempts to disguise activity, one can use Markov chains and models to determine two users through the same path in a network. This can allow identification of bad actors using encrypted channels to disguise malicious behavior.

Another approach involved the review of information usage patterns to determine between malicious and non-malicious activity. For example, consider how employees use information within an organization, including how that use changes over time (including, for examples, new roles and interests, possible malicious behavior, etc.). This allows for the development and application of a basic threat model based on these findings. The model can be altered as more data/information is obtained.

Insiders, who by definition, have knowledge of corporate systems, tend to assess ROI on attacks and exploitations based on financial payoff, ease of defeating defenses, vulnerabilities of applications, probability of detection, and other factors. Knowing these "values" can enable the defenders to identify potential malicious behavior, especially where fraud or theft is taking place at the "edge" of normal behavior.

Ideas that Challenge the Assumption:

Bad actors sometimes mimic normal behavior. Hackers constantly push the limits of detection to determine what the system considers "normal" so malicious activities won't trigger alerts. Normal behavior needs to be defined and modified in definition over time. Within cyberspace, we will need a range of classification of information, actions and behaviors.

Since normal behavior changes dynamically, one needs to monitor both normal and abnormal transactions. Malicious or unusual behavior is different from normal behavior, so it needs to be identified and distinguished from one another and the norm. Malicious behaviors also morph over time.

No matter how quickly humans can write and implement a rule, malicious actors will find an alternative. Machine and self-learning systems are needed, and adaptability is critical to the success of the system. Anomaly detection needs to help explain what is actually happening; however, it is a cognitive process that needs to be well defined. As more anomalies are detected, the hypotheses need to be re-stated and re-tested.

With regard to cyberspace, one question is whether there are so many vulnerabilities in our networks that this approach is limited in its utility. Would a hacker have the advantage of economy of scale? Can malicious actors just hide in the network noise?

Areas for Further Research and Exploration:

The following areas appear to be strong candidates for further research and exploration:

--**Classification and Modeling of Normal and Abnormal Behaviors:** we need to continue definitions, classification, and modeling of factors that could potentially lead to identification of malicious behaviors – and potentially other problems – in cyberspace. These will need to remain dynamic, as the threat from malicious actors will adapt quickly over time. These definitions must also account for distinctions between malicious behaviors and simply unusual ones.

--**Relevance of Other Disciplines to the Problem:** we leaned heavily on the outstanding efforts of the financial sector and their attempts to identify fraud and other malicious behaviors in this workshop. The financial sector believes that their approaches may be useful based upon current rules, models, and profiles. There are potentially other disciplines – behavioral sciences for one – that may be able to offer other approaches to this problem. One discussant cited 13 behavioral antecedents (e.g., deception and technical expertise) to malicious behaviors that could be identified.

--**Definition and Delineation of Human and Machine Factors:** effective identification of malicious behavior in cyberspace will invariably involve decisions made by humans and those identified by machines as sources of concern. Given the speed and the volume of cyber transactions, some delineation between these and especially a focus on self-learning and/or adapting machines is necessary. The use of fuzzy logic and rule-based systems have potential here for identifying malicious behavior.

--**New Analytic Approaches to Malicious Behavior in Cyberspace:** while there are some useful approaches identified in the workshop, others surfaced here with specific application to cyberspace:

Automation of “normal” - Can we analyze traffic through a website, and only that website to find the normal transactions for that day versus the non-normal activity? Use this information to generate automated computer rules based on normal activity for that period of time. Let the computer adapt to the traffic. This allows for near real-time adjustments to shifts in transaction activity. This approach will identify that activity is anomalous but it's difficult to determine if it is malicious.

Automation of applied behavior analysis - Rather than just looking at behavior itself, need to look at three components: the before (antecedent), the during (actual), and the after (outcomes) of the behavior. Not use a signature, but build a behavioral assessment engine. Might look at the human activities for cyber activities. This means that the system would need to look at cyber information packets and convert to behavioral images.

Insider threat - Need to develop mathematical models to assess the risk of insider threat. Insiders place emphasis on exploitation interests on applications. Such factors as value of the application/data, ease of overcoming the defensive measures, how easy to steal/cheat, accessibility to get to the target and flee, etc. Defenders can take all these factors into

consideration and develop an instantaneous probability for an activity to identify potential malicious behavior.

--Active Defense Against Malicious Behaviors: one idea proposed was the use of “white viruses” on networks to counter malicious cyberspace behaviors, whether for development of network hygiene or proliferation of patches.