MIXMODE'S ARTIFICIAL INTELLIGENCE: DYNAMIC LEARNING IN NETWORK SECURITY

By Dr. Igor Mezic, Chief Scientist and CTO for MixMode

Contents

MixMode's Platform for Artificial Intelligence in Network Security

There are by now many different platforms using Artificial Intelligence to provide information to network security analysts. The simplest type, in existence for decades, are rule-based systems, sometimes referred to as the "1st wave AI" systems. A rule is typically executed on an observation – be it of network traffic, of user actions. A typical example is the rule that triggers alerts when the size of a file exiting the local network is larger than some predetermined threshold, for example 30 MB. This network traffic based rule is called the "Large Outbound File Transfer" (LOFT).

Another example, triggered on user action, is the rule that informs the network analyst that there has been a large number of failed login attempts by a user, bigger than a predefined threshold, for example 10. Setting such thresholds requires a large effort by the analyst. For a LOFT threshold to be correct, the analyst would need to know in advance that a specific user very rarely sends files larger than 30 MB outside of the local network. That would require examining the traffic patterns and file sizes for each individual user on the network! To make things worse, the user pattern might change with a change of his/her responsibilities. The analyst would have to go back periodically, and reset the threshold to adapt it to a dynamically changing network dynamics.

Finally, if there is no context supplied with the information that a file is larger than 30MB, this approach might generate a large number of false positives – cases in which an alert is given, but the action is not malicious. This happens when the imposed threshold is too small. It can also lead to a large number of false negatives – cases in which an alert is not given, but the action in fact is malicious. This happens when the threshold is too large. Clearly, tuning the threshold so that it does not produce either false negatives or positives is again a great burden on the analysts time.

A more modern approach - but still 20-th century – is based on the mathematics of Machine Learning - a set of mathematical algorithms that enable detection of patterns in data. Artificial Intelligence systems based on Machine Learning are referred to as the "2nd wave AI".

There are many different types of algorithms used in machine learning, relying e.g. on Deep Neural Networks, Support Vector Machines, Bayesian Learning, and many other mathematical techniques. As opposed to a rule-based system, a Machine Learning-based AI utilizes historical data on the network and can, based on such historical data, set the aforementioned thresholds for the rules, as well as perform analysis of deviations from the previously observed behavior on the network. However, the large amount of data flowing on a typical corporate network requires a massive, computationally expensive learning effort that can last months or longer. And once the learning is finished, the dynamics of the network might have already changed, and the learning would need to start anew.

The above description, although simplified, establishes the two biggest obstacles to introducing an effective AI system into network security: the dynamically changing network environment, and the large amount of dynamically evolving, unlabeled data. The resolution of this problem is the introduction of the "3rd wave", dynamically learning, computationally efficient platform, that starts learning from the first 5 minutes it is deployed, does not require historical data, and is adapting actively to the dynamic changes in massive amounts of network data. MixMode's technology provides such a platform - based on mathematical algorithms invented in 21st century - and represents a huge step forward in AI for network security.

Supervised vs. Unsupervised Machine Learning

Supervised learning is a type of machine learning where labeled historical data are supplied as an input to an algorithm, that is then trained to recognize labeled patterns. The algorithm is required to recognize newly acquired data as being of a certain type that was already present in the historical data. A typical use of such algorithms is in image recognition. Deep Neural Networks are often capable of recognizing objects in new images based on similar objects in many labeled images that it has been trained on. In contrast, unsupervised learning does not require prior labels. It classifies objects it sees in historical data with its own, internal labels.

This is how humans in fact learn: even a baby will recognize a cat passing in front of it as something that moves, and thus deviates from a static background environment. Certain features are recognized – e.g. the object having 4 legs, whiskers, ears, eyes, and a long tail. But there is no label. The key to learning in this way is establishing a baseline (static background) and the deviation (motion).

In network security, Zero-Day threats, or even threats obtained as modifications of the well-known ones, do not come precisely labeled. Thus, reliance on an unsupervised learning AI system is a necessity. But even here, the dynamically changing network profile renders the use of off-the shelf algorithms, such as clustering algorithms, inefficient. And the issue of the large historical datasets is to be avoided.

Baselines and Anomalies

Ideally, an AI system needs to be able to discover the underlying normal and abnormal patterns on the network automatically, and dynamically adapt to them. Bayesian methods, based on the 19th century Bayes' theorem, assume a certain prior and update it with acquisition of new data. However, There is quite a bit known about the type of dynamics that a typical network exhibits over time.

As a simple example, traffic volume will be smaller over the weekend than during workdays in a typical corporate environment. An AI system does need to take into account – learn – such regularities, including regularities in the "stochastic" part of the network behavior that depends on freewheeling human exchange over the network.

MixMode's AI does that using the theory of Koopman Mode Decomposition, invented by its Chief Scientist and CTO in 2005, and patented for network security use by MixMode. The methodology is adapted to the specifics of network data. It encodes the various spatial and temporal patterns of the data in the so-called Koopman Modes. These mathematical objects encode the regular patterns of dataflow – on a network, on CloudTrail, in alerting platforms, or on any other timestamped data source. When the AI system is deployed to the timestamped data of network flows, the key learned elements are network Koopman modes – patterns that represent common behavior on the network over a specific timescale.

In Figure 1 we show such a pattern. On the horizontal axis are integer labels for source IP's, while on the vertical axis we show integer labels for destination IP's.

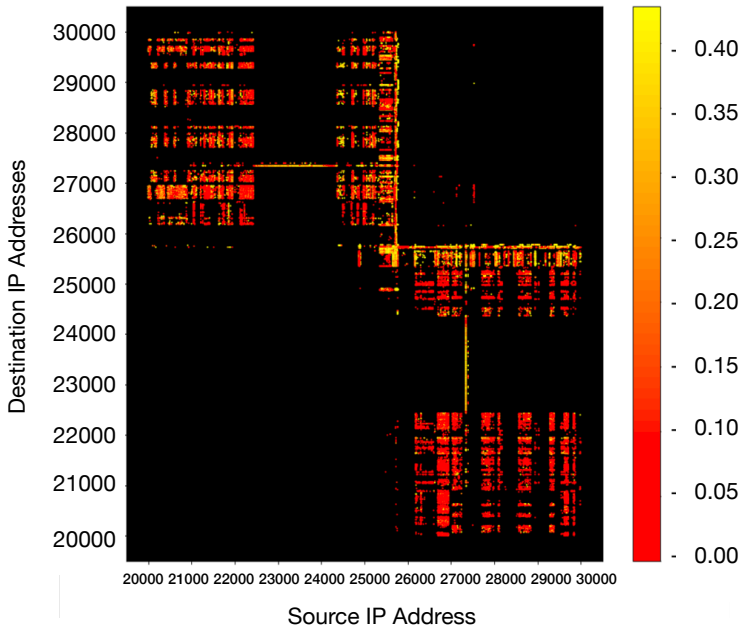


Figure 1: Patter of interaction on a network on the daily timescale.

The timescale represented in figure 1 is that of 24 hours, and thus the colors in it represent daily interaction. It is useful to look at the intersection of lines emanating from a specific source IP and destination IP. Colors on the vertical line stemming from a specific label, say 21000, on the horizontal axis indicate the level of data flow from that IP to any destination IP represented on the vertical axis. Colors on the horizontal line stemming from a specific label, say 22000, on the vertical axis indicate the level of daily flow to that IP from any source IP represented on the horizontal axis. For example, the source IP labeled 21000 and the destination IP labeled 22000 do not interact on a daily timescale as evidenced by the black color at the intersection of the vertical and horizontal lines stemming from them. However, if the color is yellow, the level of activity is high. The information that MixMode's AI analyzes consists of many such modes representing activity on many different timescales. But, even the single "heat map" shown in Figure 1 contains a lot of information about the network.

For example, it is clear from the Figure 1 that source IP's labeled roughly 21000-25000 only send information to destination IP's labeled roughly 25700-26000, and vice versa. This immediately indicates two different subnetworks of this network that can be easily identified by MixMode's AI. There is also some deviation from that pattern. A few scattered dots in the upper right and lower left corner of Figure 1 indicate there is some interaction within individual subnetwork. Such deviations can be investigated for anomalous time dependent behavior and brought to analyst's attention, sharpening the focus on individual IP behavior.

Capturing Dynamics

MixMode's AI computes such patterns of interaction over many different timescales, and contrasts the pattern over the next short interval of 5 minutes with what was seen previously. If the patterns deviate, an assessment of the security risk implied in the deviation is computed and presented to the user. Even if the threat is Zero-Day, the unsupervised nature of MixMode's dynamic learning algoritms is able to recognize it.In addition, if the risk is low, the deluge of intel and notices presented to the user is minimized, eliminating false positives.

Thus, MixMode's third wave AI makes its decisions on Zero-Day threats and False Positives based on the intuitively transparent concept of interaction of network elements over variety of timescales - in the same way a human would - but utilizes its massive computational powers to do it efficiently.

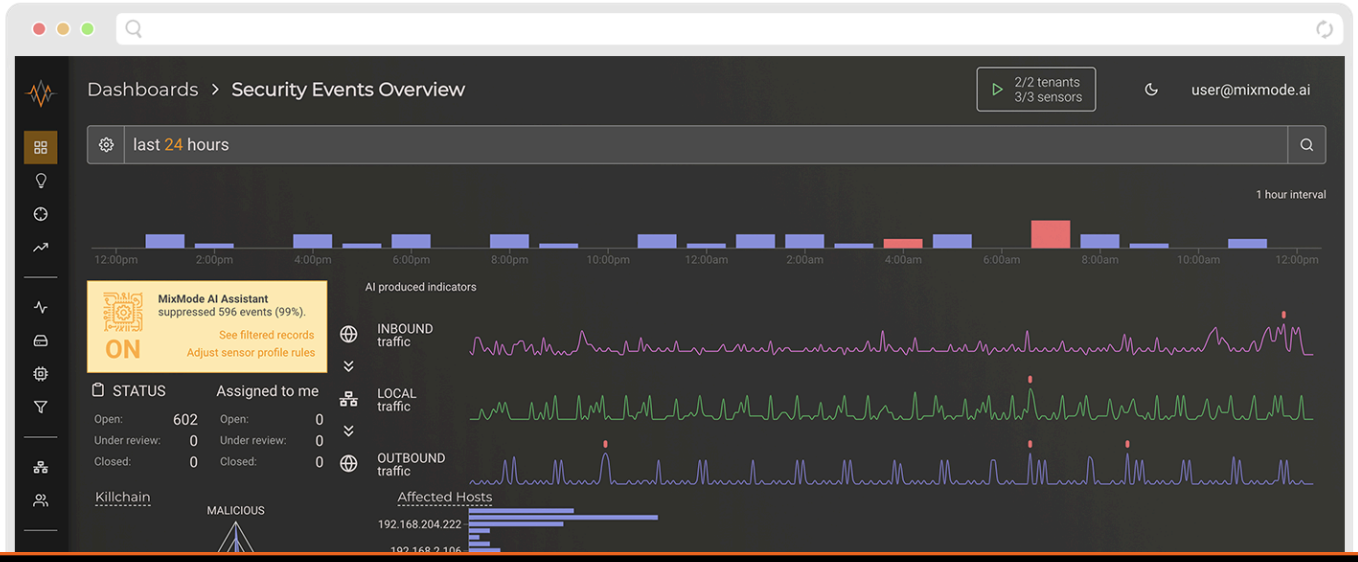
Capturing Correlations

Correlations of activities – in time, space, and across data sources are of critical importance to network security analysts, and part of the radical innovation that MixMode is introducing into the security space. For example, a network analyst would expect that observation of an intrusion into the system, could be followed by lateral movement on local hosts, and attempt to exfiltration of the data from the local network. These three actions are linked, and in the case of a smash, grab and run action – correlated in time.

Figure 2 shows such a time-correlated event on the network detected by MixMode's AI, and clearly correlated in time over inbound, local and outbound trendlines. Since the times between actions on different subnetworks do not have to be precisely defined, it is difficult for machine learning algorithms to classify such correlation. But, in MixMode' AI case, the classification is quite natural as the time dynamics is naturally incorporated into the guts of the algorithm.

Correlations can also exist in space. The heat map clearly shows that activities between different spatial – subnetwork – domains. Deviations from such patterns – small pockets of activity shown in yellow and red amongst otherwise black areas of the map – can be autonomously flagged for investigation. It can be seen that such deviations can be at the level of a few individual IP's. This is another powerful feature of MixMode's AI approach: it points out to the network analyst anomalous behavior, tracing it down to an individual IP.

Figure 2: MixMode's AI



The AI does not rely on backroom analysts curating its suggestions on anomalies. But also, it does not shut down the activity on such detected anomalies, preferring to have a human in the loop making the decision. The “hands on the wheel” principle of 2nd level autonomy at work!

Correlations useful for a network analyst also exist across different streams of data. It is typical that different sensors and datatypes available on the network will provide different “observations” of an ongoing security event. MixMode’s AI is capable of taking in a large variety of such data into its Multistream platform. AI analytics deployed on a combination of SIM, Cloudtrail, Bro and other sensor data yield information characteristic of an attack and shine the light on the malicious activity from a variety of angles.

The essential feature that enables MixMode to provide this service is the use of a single underlying algorithmic approach, in large contrast to the majority of network security AI platforms that typically build customized algorithms for the various common types of security events. MixMode’s algorithm does not mathematically operate differently in for example detecting beaconing vs. detecting an unknown Zero-Day intrusion.

False Positives and False Negatives

We have discussed the issue of false positives and false negatives at the beginning of the article. It is a known feature of AI systems that it is hard to build an algorithm that at the same time enables Zero-Day threat detection (and thus minimizes false negatives), and at the same time features few false positives. It is again the ability of MixMode’s Koopman Mode Decomposition based AI to capture dynamic behavior on the network that is at the core of enabling such performance. In figure 3 we show the time trace of the total volume of traffic on a network, shown in blue, and occurrence of several alerts from the Bro sensor, such as LOFT, shown in green. It is clear to the human eye that the patterns of larger volumes and larger number of alerts are correlated.

The human intelligence would immediately conclude that most of these alerts are in fact false positives – it is simply that the total traffic has increased and thus the file sizes in that total traffic are larger as well.

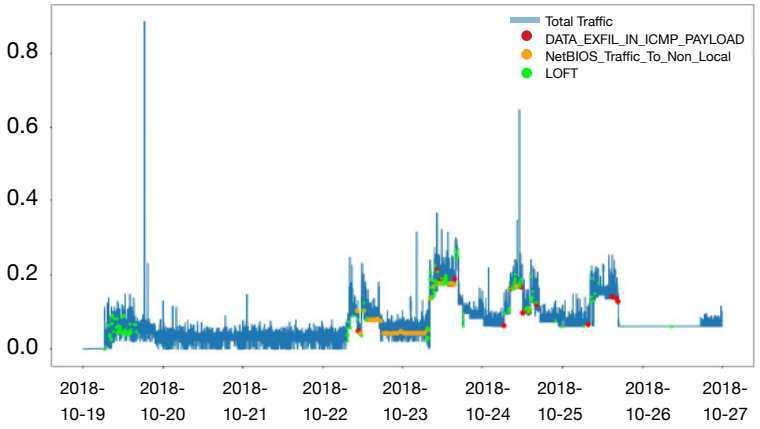


Figure 3: Total traffic in range and timestamps of logged alarms with nearest traffic datapoints

The exception is one green dot at the far right, where the total volume of traffic is low and non-fluctuating, and the alert occurs during nighttime. MixMode’s AI detects that alert as worth investigating, just like a human analyst would, based entirely in its unsupervised, uncurated learning algorithm.

Adversarial AI

A great question for any AI platform in network security is: does it protect from an adversarial AI system? Recall that MixMode’s AI relies on learning the baseline patterns – normal, coherent, and normal, random (caused by free-willed human actions) and detecting anomalous as a difference of current behavior with those normal patterns in an unsupervised manner. Thus, for an adversarial AI to beat it, it would have to learn precisely what MixMode’s AI knows. But in that process, MixMode’s AI would likely detect the learning behavior as anomalous, and report it. The resilience to adversarial AI is a built in feature to MixMode’s 3rd wave platform.

About the Author

Dr. Igor Mezic is the Chief Scientist and CTO for MixMode. He has spent his career developing highly complex algorithms and artificial intelligence for data analytics. He graduated with a doctorate from CalTech, holds 5 patents, and is a professor of mechanical engineering at the University of California, Santa Barbara.