Rogue Behavior Detection

Identifying Behavioral Anomalies in Human Generated Data

White Paper

Executive Summary

This white paper describes the Numenta Rogue Behavior Detection application. This application utilizes Hierarchical Temporal Memory, a machine intelligence technology that captures the essential properties of the neocortex, to automatically learn and model an individual's behaviors and identify abnormal or irregular activities. Key points covered in this white paper are as follows:

- This application detects the patterns and anomalies in individual human behavior using human and machine generated data.
- Anomaly detection of human behavior is useful in IT security, regulatory compliance, financial risk assessment and device access control.
- Automated modeling of individual behaviors provides superior flexibility, scalability and precision; it also eliminates the upfront effort spent defining what is normal or abnormal for each individual.
- Individual behavioral models are developed by encoding both human and machine generated metrics into Sparse Distributed Representations (SDRs) that are then processed by the HTM Learning Algorithms.
- Numenta simulated scenarios where its employees recreated events worrisome to IT security experts. Examples of anomalies identified are provided.
- Parties interested in recreating or expanding upon these results can access the encoder and algorithms used in this application through NuPIC, Numenta's open source project.
- This application takes advantage of the same underlying Hierarchical Temporal Memory theory and Cortical Learning Algorithm code base that is used in other Numenta applications, demonstrating the generalizability of Numenta's technology to a variety of use cases.

Advances in Anomaly Detection

> **a·nom·a·ly** Something that deviates from what is standard, normal, or expected.

The ability to detect anomalies in real time can be incredibly valuable. Imagine being able to notice early warning signs for failure in a large turbine, see slight variations in a heartbeat indicating disease, or detect an unusual pattern of shopping cart failures on an ecommerce website. Anomalies are not always bad or indicative of a failure. For example, detecting a subtle change in consumer buying habits could provide an opportunity to discover a new trend. In today's world where the amount of data being collected is exploding, the opportunity for detecting anomalies is rapidly growing.

However, accurately detecting anomalies can be very difficult. First, what qualifies as an anomaly is always changing. Systems evolve over time as software is updated or as behaviors change.

Therefore, effective anomaly detection requires a system to learn continuously. Second, to detect anomalies early one can't wait for a metric to be obviously out of bounds. Early detection requires the ability to detect subtle changes in patterns that are not obvious or easily detected. Furthermore, because anomalies by their nature are unexpected, an effective detection system must be able to determine whether new events are anomalous without relying on preprogrammed thresholds.

At Numenta we have taken a fresh approach to this problem and have created what we believe is the world's most powerful anomaly detection technology. This approach is derived from our understanding of the neocortex, which is itself a powerful prediction and anomaly detection system. Our suite of applications takes advantage of this understanding to drive state-of-the-art breakthroughs in two dimensions: 1) how we utilize the processes of the brain to model data, and 2) how we detect anomalies based on that model. This paper will describe these advances by illustrating how Numenta's rogue behavior detection application models the data created by human actions to identify behavioral anomalies.

Behavioral Modeling and Rogue Behavior Detection

The Numenta Rogue Behavior Detection application detects the patterns and anomalies in individual human behavior. By focusing on the digital fingerprints created by human actions, this application provides new opportunities for monitoring the behaviors and compliance of enterprise personnel.

Gartner estimates worldwide IT security spending was $64B in 2013[1], with approximately $3B going to Security Information & Event Management applications[2] that monitor network hardware and applications to provide real time analysis of security alerts. On top of these expenditures, many companies in heavily regulated industries, such as energy, healthcare, intelligence and finance, are also faced with the costs of monitoring personnel for safety, privacy and risk compliance. The Numenta Rogue Behavior Detection application may be used to address these types of activity and compliance concerns in the following ways:

- Detect unusual employee access to intellectual property and internal systems
- Identify abnormal financial trading activities or asset allocations by individual traders
- Alert when employee behaviors or actions fall outside of typical patterns
- Detect the installation, activation, or usage of unapproved software
- Alert when employee computers or devices are used by unauthorized individuals

Existing products designed to provide IT security and compliance monitoring focus on network policing and credential management. The scope of activities that these types of services can adequately monitor is limited, and many enterprises find that no products exist to track other types of employee activities relevant to their specific situations. As a result, these companies are forced to rely on cumbersome employee self-reporting and infrequent third party auditing.

[1] http://www.gartner.com/newsroom/id/2512215
[2] http://www.marketsandmarkets.com/PressReleases/security-information-event-management.asp

This application provides new opportunities for enterprises to monitor behaviors and activities that are important to their specific applications. First, this application automatically builds behavioral models unique to each individual being monitored. In comparison to rule-based and role-based monitoring, this automation eliminates the upfront effort spent classifying employees into groups and deciding what is normal or abnormal for each group. Additionally, this automatic modeling provides a level of scalability and precision important to implementation across large organizations.

Second, this application ingests both human and machine generated data in order to build a comprehensive understanding of the individuals being monitored. While many products offer methods for analyzing machine data, this application's ability to analyze human generated data provides an additional level of insight into employee activity patterns. This enables companies to monitor additional classes of activities and identify new types of irregularities. Taken together, organizations now have a more robust ability to detect real-time irregularities in the specific activities and behaviors that they find important.

Modeling, Prediction and Anomaly Detection

Numenta's Rogue Behavior Detection application excels at rapidly learning an individual's behavioral patterns and then detecting activities that deviate from his or her norms. To do this, streams of behavioral indicators are encoded into Sparse Distributed Representations (SDRs). SDRs are the language of the brain and enable several useful attributes such as generalizability across data stream type, strong resistance to noise, and the attachment of semantic meaning to data points. Sequences of SDRs are then analyzed by the HTM Learning Algorithms, themselves simulations of a small slices of the neocortex. More information about Sparse Distributed Representations and the HTM Learning Algorithms are available in the technical pages of Numenta's website or through NuPIC, Numenta's open source project that contains descriptions and implementations of the algorithms used in our applications.

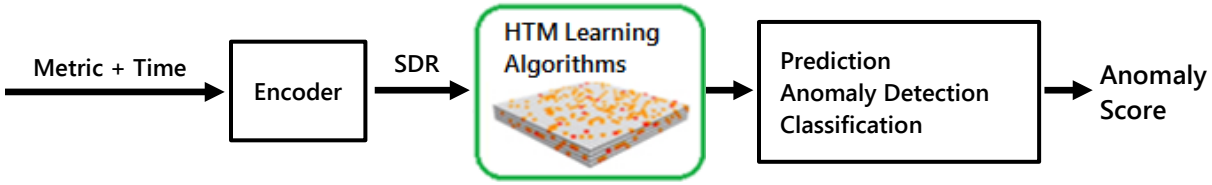


Figure 1 - Process for modeling human and machine generated behaviors and identifying anomalies

The range of data streams that can be analyzed is only limited by an enterprise's ability to collect data. Streams that represent continuous scalar values (e.g. CPU utilization, network usage and file activity) can be fed directly into the encoder for processing and analysis. Discrete events such as keystroke counts, screenshots taken, and instances of data being moved to external devices must first be aggregated and preprocessed into continuous scalar values before encoding. An example of this type of preprocessing is shown below in Figure 2.

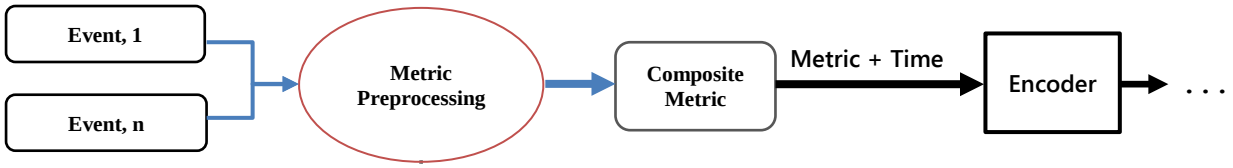


Figure 2 - Preprocessing of infrequent events

Rogue Behavior Detection in Practice

The Numenta Rogue Behavior Detection application makes it easy to monitor people across the various departments of an organization and identify those with the most abnormal behaviors. To demonstrate, this application monitored Numenta employees as they simulated scenarios worrisome to IT security experts. The following figures show the results of these tests, represented through the perspective of a fictional financial institution.

The dashboard depicted in Figure 3 shows two views; one representing the past 24 hours (day view) and another representing the past eight days (week view). In each view, the application orders the departments being monitored within an organization based on how anomalous the individuals are within each department. For each department, the four individuals with the most anomalous behaviors will be shown on the dashboard. One can then choose to view all the employees within a particular department by tapping the dropdown button on the right side of the department heading, or directly view the details of an individual employee's behavioral anomaly score by tapping on that individual. Return to the home dashboard by tapping on the application's icon in the upper left of the screen.

The colored bars below each person's name and job title represent the level of anomaly that was measured for each person at different points in time. The color and height of the bar have the same meaning. Small, green bars represent low levels of measured anomalies. Progressively taller yellow and red bars represent increasing levels on anomaly.

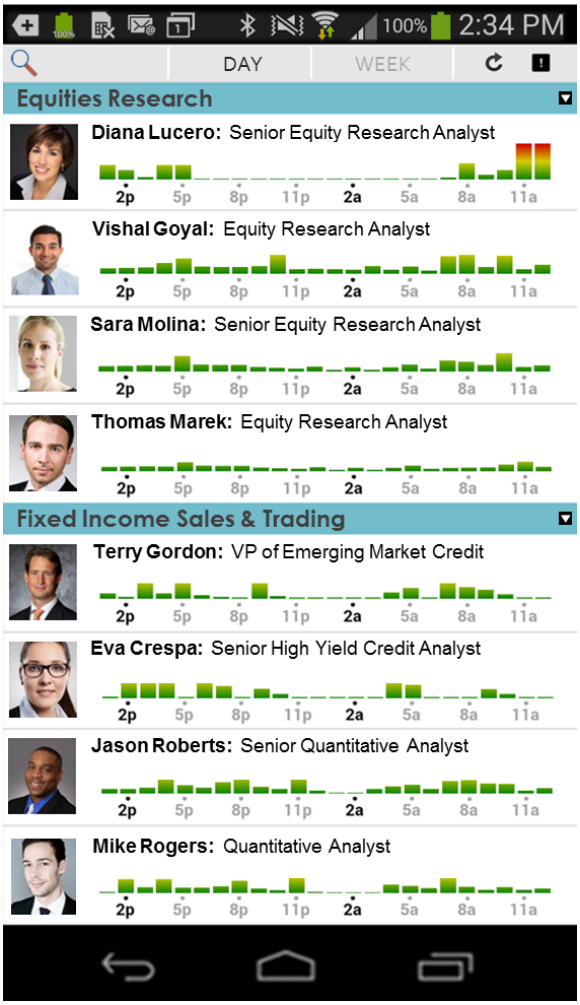


Figure 3 - Application dashboard showing overview of employee anomaly scores sorted by department

With a quick glance at this dashboard above in Figure 3, one sees two tall red anomaly bars indicating that Senior Equity Research Analyst, Diana Lucero has been exhibiting abnormal behavior. To drill down and view the specific metrics that contributed to these anomaly scores, one can tap on the behavioral overview panel for this specific employee (Figure 4). In this view, one sees that this employee's File Activity and CPU Percent were the metrics that contributed to the anomaly alert.

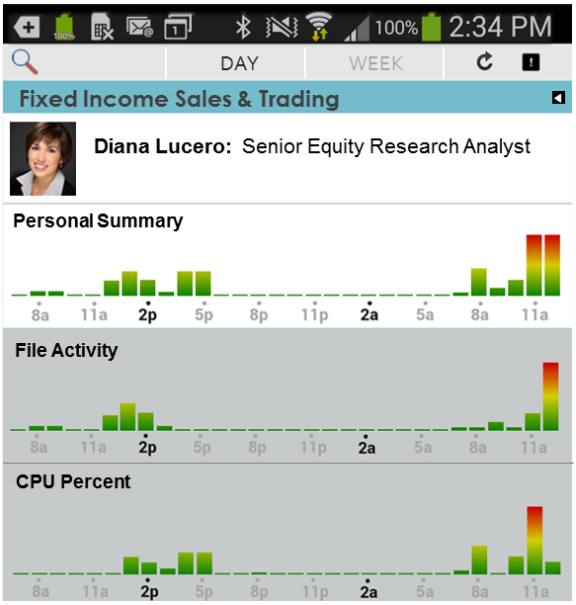


Figure 4 – Individual employee anomaly score perspective

Once a manager or security expert knows where to look, they may then choose to drill down into the underlying data for each metric of interest. For example, tapping on the CPU Percent panel brings up the chart on the left in Figure 5 showing that a spike in this employee's CPU utilization generated the first red anomaly bar. Likewise, tapping on the File Activity panel brings up the chart on the right that shows that the second anomaly bar was generated by a spike in bytes saved at 11:00 am.

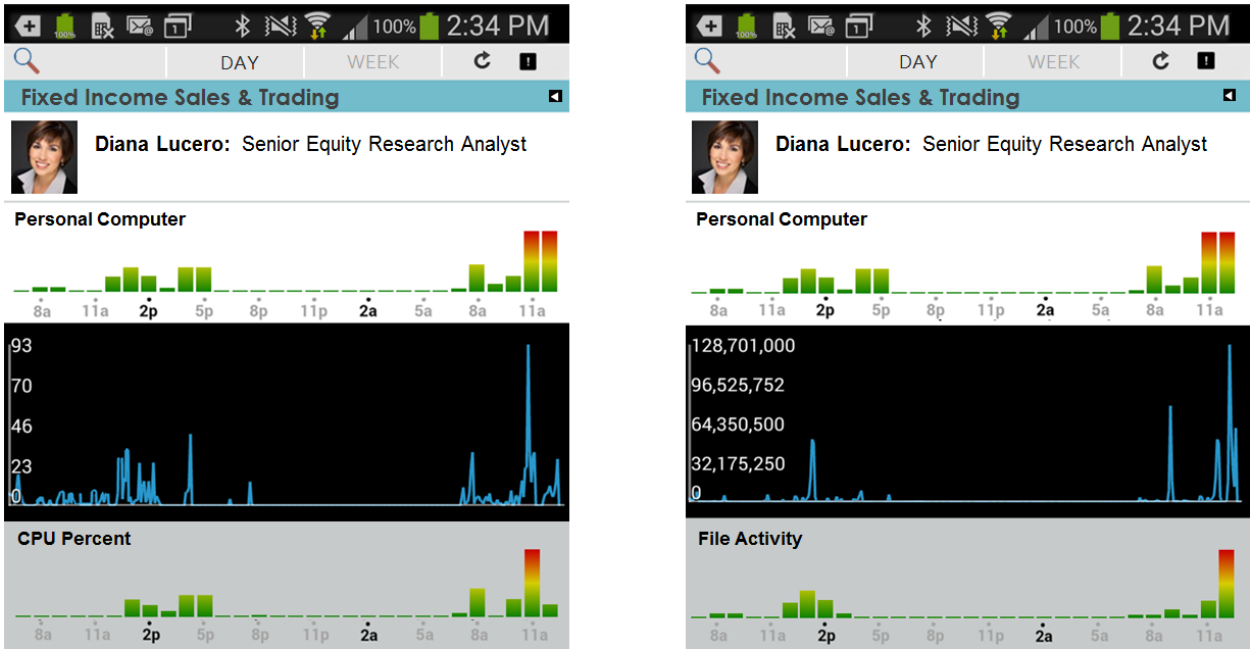


Figure 5 – Drilling down into the underlying metric data

In this case, further investigation into this employee's activities found that the anomaly alerts were generated by the creation of a large .zip file containing intellectual property that had been saved on the internal network. By being alerted to this suspicious behavior in real time, management would have the opportunity to take action before the misappropriated intellectual property could be misused.

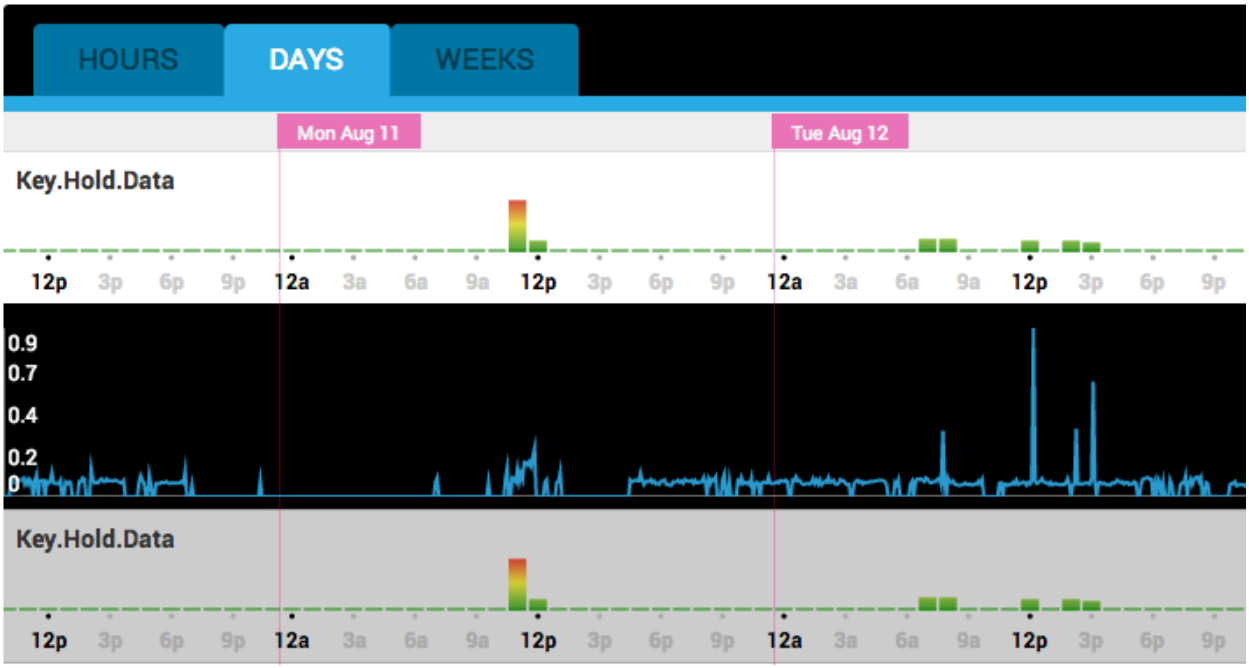
Another experiment run by Numenta shows the nuance with which this application is able to separate true anomalies from the normal fluctuations in employee behaviors and work patterns. Figure 6 shows the anomaly score generated by an employee who installed an unauthorized program onto their computer.

Figure 6 - Downloading an unauthorized program

Most interesting about this example is how a clear anomaly was flagged even though the underlying data does not look especially anomalous to the naked eye. In the chart above, the Memory Percent metric has a series of peaks and valleys during the time period shown. Although this metric did record its highest value at the time that the anomaly was flagged, the height of this spike was only slightly greater than the heights of the other spikes shown. However, because this application had already developed a sophisticated model that predicted the behaviors of this individual user, it was able to identify this particular activity as a significant behavioral irregularity. As a result, the anomaly score generated for this spike was nearly three times greater than the score for any other event in this period.

A final example shows how the Numenta Rogue Behavior Detection application is able to identify when a computer terminal or device is accessed by an unexpected user. In the example shown below in Figure 7 (depicted from the Rogue Behavior Detection web client), keystroke dynamics were being collected to model the specific patterns with which individuals interacted with their devices. As

can be shown in this figure, no anomalous usage patterns were identified by this application before Monday, August 11th. However, when a new user began using this laptop at 11:00 am on Monday, August 11th, this application immediately flagged a clear anomaly based on abnormalities to the keystroke hold patterns it observed.



Once this anomaly had been flagged, a variety of preventative or investigatory actions could be taken by the monitoring enterprise to ensure the security of IP accessible with this potentially compromised device. For instance, this laptop could automatically prompt the user to input a password or biometric scan to ensure proper access rights. Alternatively, this anomaly could prompt a security administrator to quarantine this device from accessing certain areas of the network until further investigation has been completed.

Conclusion

The Numenta Rogue Behavior Detection application provides new methods for detecting rogue behaviors through the use of human behavioral monitoring and anomaly detection. By modeling each individual's behaviors based on the actual sequences of their actions and activities, highly personalized and scalable monitoring can be achieved that has the ability to alert managers and compliance officers to abnormal and potentially suspicious employee behaviors in real-time. Additionally, the ability to continually learn new behaviors while remembering older patterns eliminates the painful setup and maintenance tasks associated with rule-based or role-based monitoring alternatives. This gives internal data security and compliance personnel the flexibility to offer their employees and contractors the network access they need, while minimizing the headaches associated with data protection.

Although the Numenta Rogue Behavior Detection application makes it easy to monitor the behavioral anomalies and compliance of individuals within your organization, the science behind this technology is extremely sophisticated. This application takes advantage of the same underlying Hierarchical Temporal Memory theory and HTM Learning Algorithm code base that also underpins Numenta's other applications (Figure 8). If you are interested in recreating the results shown above or developing new applications that take advantage of these advances in behavioral modeling, this application's source code is available in NuPIC – Numenta's open source project that contains candid descriptions of our algorithms and software modules.

We also invite you to visit our website to learn more about the science that underpins all of our applications and algorithms. By applying years of research in neuroscience and computer science, we believe that our approach to anomaly detection and pattern recognition represents a significant step forward for the monitoring of anything that generates continuous data.

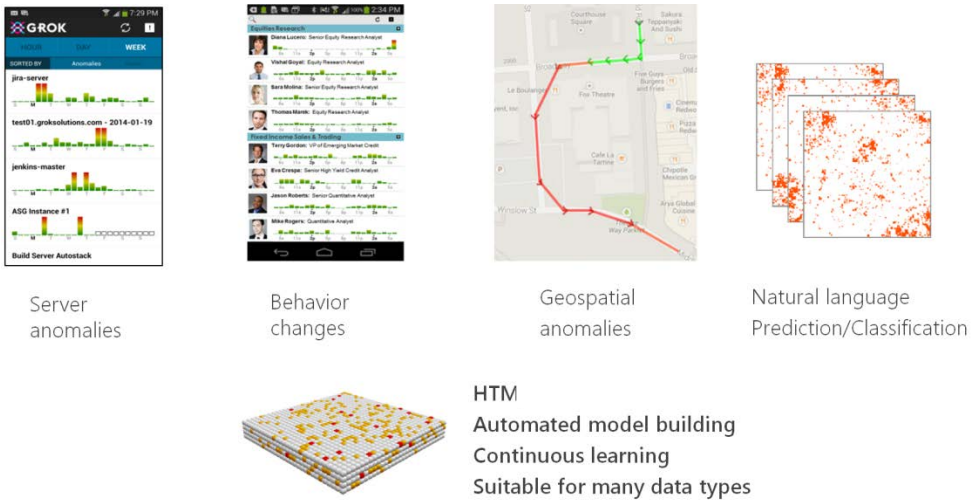


Figure 7 - The same HTM code base underpins Numenta's diverse suite of applications

About Numenta

Numenta was founded in 2005 to lead the new era of machine intelligence. Numenta builds solutions that help companies automatically and intelligently act on data. Its biologically inspired machine learning technology is based on a theory of the neocortex first described in co-founder Jeff Hawkins' book, *On Intelligence*. This technology is ideal for large-scale analysis of continuously streaming data sets and excels at modeling and predicting patterns in data. The application described above is just one of a suite of products and applications the utilize Numenta's flexible and generalizable HTM Learning Algorithms to provide solutions that encompass the fields of machine generated data, human behavioral modeling, geo-location processing, semantic understanding and sensory-motor control. In addition, Numenta has created NuPIC (Numenta Platform for Intelligent Computing) as an open source project. Numenta is based in Redwood City, California.