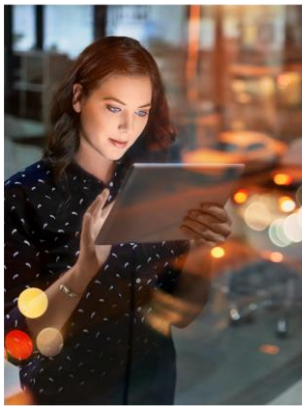
Machine Learning + Graph Analytics – Recipe for Transformational AML and ATF Programs

Solution to Financial Crime & Compliance Management

By: Garima Chaudhary, Oracle Financial Crime and Compliance Management Specialist

ORACLE

In an already stringent regulatory climate, banks and financial institutions face strict new hurdles such as new industry players (FinTech) and political uncertainty around anti-money laundering (AML) & anti-terrorist financing (ATF) compliance. With no sign of relief, the play of innovation in financial crime was never this critical.



"As FIs continue to transform their program to keep up with the ever-changing regulatory landscape, efficient monitoring will be a key part. Transformation capabilities, such as Machine Learning, will drive down their operational costs, while reducing risk and providing efficiency and agility"

Garima Chaudhary
Oracle Financial Crime and Compliance Management Specialist

Despite their potential, the adoption of Artificial Intelligence (AI) and Machine Learning (ML) within financial crime has been relatively slow. This is due to the limited understanding of AI and ML itself, limited understanding of how AI and ML within compliance programs, and to the fact that regulators and compliance officers are often concerned that AI and ML are "black boxes" whose inner workings are not clearly understood. Regulators typically require compliance officers to understand and validate not just the outputs, but also how the outcomes from models are derived. Additionally, increase in compliance scope is shifting the profiles of senior investigators, which is expected to be more like data scientists now. That will lead to the recruitment of such roles: more technical, which will automatically increase the cost of compliance. Despite some of the concerns, we already see movement and application of these technologies. Recent joint statement (Joint Statement on Innovative Efforts to Combat Money Laundering and Terrorist Financing) by major regulatory bodies in United States emphasizes the use of innovative technologies to combat financial crime is an indication that AI & ML usages will continue to nurture.

High false positive rate in transaction monitoring being a key factor for increased cost of compliance, as one of the area where AI and ML has been applied to increase the monitoring effectiveness and reduce false positives rate. In this white paper we will discuss how advance AI and ML technologies can be applied to achieve next-gen financial crime program.

TRADITIONAL APPROACH: RISK INDICATOR VS OVERALL PATTERN

ML is applied on red flags (or Event) generated by traditional deterministic rules-based detection systems. ML algorithms leverages historical data for training the ML models and predicts the likelihood of suspicious activity when new Events/red flags are generated. Based on the model outcome score & risk tolerance of organization the red flags are either suppressed or can result into cases. These Cases are further investigated by users. The final investigation decision about a Case (productive vs

non-productive) is fed back to the model for machine to learn. This approach leverages individual red flags both for feedback & ML.

Although, organizations were able to reduce the number of false positives by a magnitude, up to 30 percent, there are some major gaps which this approach possesses.

- **Lack Complete Pattern:** The red flags generated by deterministic detection systems are merely an indicator of money laundering or terrorist financing activity and not the complete pattern. And the only way to understand overall behavior is by combining the series of Events/customer activities. For example, if one structuring Event was generated for a customer, then by investigating just one Event, user may not be able to get the holistic view of pattern thus not making the right decision about customer behavior. Though, if the investigator is presented with a series of structuring Events then that increased the understanding of overall customer behavior, thus leads to more accurate decision. This means, a ML model which factors individual event is not considering the complete money movement pattern, consequently making inaccurate decisions.
- **Lack Customer Holistic View:** In an approach, when individual Events are factored for ML, the information associated with those events are leveraged by the model only. Therefore, the machine is not just missing the entire money movement pattern but also, the complete customer network. Which means, the design of ML models allows for a very small percentage of customer information therefore, lacks customer holistic view.

Traditional way of applying ML models focuses purely on reducing the number of false positive a firm may beget yet makes the entire program very inefficient.

MASS SURVEILLANCE: CASE AS A GRAPH

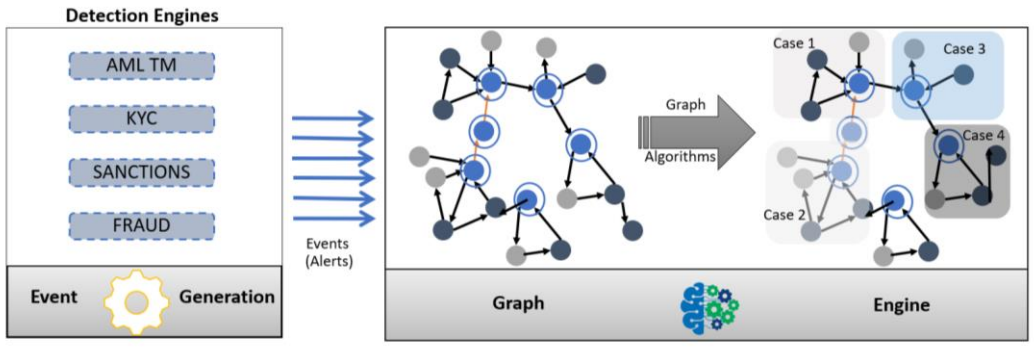


Figure 1. Cases as Graphs & Machine Learning

Step 1 – Detection Engines: During the event ingestion process the events should go through basic data checks and validation to ensure they can correctly process through the optimization layer. In case the event doesn't meet required data standards, it should flow into the exception queue. Event enrichment would aid scoring the event better, whereas case enrichment would facilitate holistic investigation.

Step 2 – Graph: During consolidation all monitoring events should be consolidated as Graph.

Step 3 - Scoring & Correlation: Provision to add or subtract scores from the events in a pre-case while correlating before creating a case.

When a Case (or Event) is created for investigation, the investigators are expected to understand all entities related to the focused entity, as well as money movements between all of them in the network. The information may be stored in a tabular format however, the analysis must be happening in a format that is understandable by a human. This involves picturing all involved entities, their relationship with one another and money movement flow between them. If we draw this entire process it will turn out as s Graph. Therefore, it is fair to say that a Case is nothing but a Graph. By definition - “a Graph is a collection of points, called vertices (or nodes), and line segments connecting those points, called edges”. Like a Graph, a Case contains nodes which is nothing but involved entities (customer, account, address, external entity, financial institution) and edges which is relationship between those entities (customer to account, customer to address, originator, beneficiary).

Considering that Case as a Graph opens a completely new avenue for detection as well as for investigation. From detection standpoint instead of generating individual red flag as a Case, the model should be transformed to produce entity network (a Graph), by correlating all involved parties. Once correlated, apply individual Event scores & correlation scores to influence overall Case score, and then decide the generation of a Case.

In a traditional monitoring environment ML is applied on individual red flags/Event. This fails to factor into consideration the overall network behavior. If the monitoring system is based on how people have been able to beat the system in the past, it will fail to find new methods and techniques to cheat the system. This limitation can be overcome by using advanced AI/ML techniques and models by correlating various Events as a comprehensive network (a Graph) during detection.

MACHINE LEARNING: CASE SIMILARITY

Transforming Case as a Graph opens new avenues of ML or deep learning, such as Graph Similarity. Graph Similarity involves determining the degree of similarity between these two Graphs. Intuitively, the same node in both graphs would be similar if, its neighbors are similar (and its connectivity, in terms of edge, to its neighbors). Again, its neighbors are similar if their neighborhoods are similar, and so on. This intuition guides the possibility of using belief propagation (BP) as a method for measuring Graph Similarity, precisely because of the nature of the algorithm and its dependence on neighborhood structure.

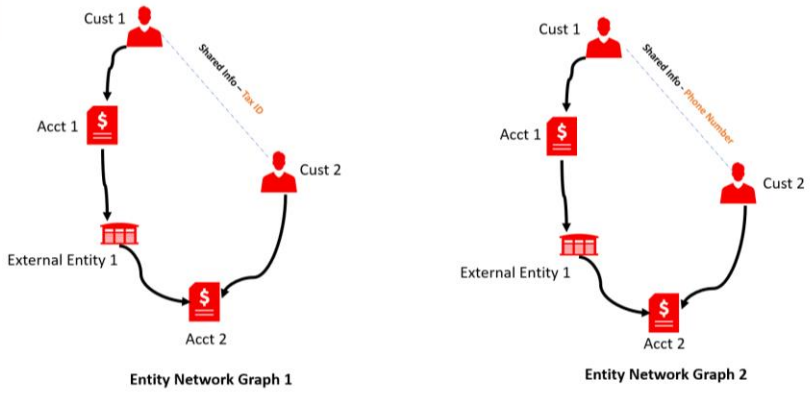


Figure 2. Case Similarity

If we compare above graphs, in entity network graph 1, customer 1 is sending money to customer 2's account through an external entity. Customer 1 and customer 2 has common tax ID. In entity network graph 2, again customer 1 is sending money to customer 2's account through an external entity. Customer 1 and customer 2 share common phone number. Which means both the entity network graphs are similar. This is a very simple example and algorithms can be applied to learn from historical "similar" Graphs to influence new Case outcome. For example, using the iterative method framework, in which two Graph elements are similar if their neighborhoods are similar, a computer can immediately suggest a straightforward way to build the score of a node: a node in one Graph like a node in another Graph their respective source and terminal nodes are similar. This definition of node similarity introduces a coupling between edge and node scores, called Coupled node-edge scoring. Exploring the information about various graphs contained in the similarity matrices is an ongoing task and this is the area where ML or deep learning can be applied.

With Graph Similarity approach, the ML model is not limited to specific point in time risk indicators (such as Transaction Volume, Transaction Count, Risk Level, Number of Parties) but expands to overall money movement pattern and customer holistic network. This is a more efficient way of comparing new vs previous Cases leading to a better outcome.

GETTING AHEAD OF THE CURVE

Traditional way of ML detection system is engineered to detect anomalies rather than to detect patterns. The combination of Case as a Graph and ML while detecting the probability of suspicious activity will transform AML & ATF program. Graph analytics & Graph Similarity has numerous key applications in diverse fields (such as social networks, image processing, biological networks, chemical compounds, and computer vision), and therefore there have been many algorithms and similarity measures already available which financial crime industry will be able to leverage.

To learn more about how Oracle addresses this topic, contact us here.

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com.
Outside North America, find your local office at oracle.com/contact.

blogs.oracle.com/financialservices facebook.com/OracleFS twitter.com/oraclefs

Integrated Cloud Applications & Platform Services