Analytical modelling and advanced techniques for AML/CTF

"A model is always partial, but it offers resources for advancing knowledge"
Jean-Pierre Changeux[73]

This section describes some of the trends and more innovative industry practices based on analytical modelling and advanced techniques for the identification, management, control and oversight of ML/TF.

The context for the analytics approach to AML assessment

With the emergence of more restricting regulation, aiming for a quicker and better identification of risk, and new technologies available, financial institutions are moving along a new transformational journey regarding the implementation of adoption of advanced AML analytics[74]. The three primary tools used to detect ML include the customer risk rating, the Transaction Monitoring, and the Sanction Screening rules.

Customer Risk Rating

The customer risk rating is a model based on the risk drivers associated with the ML identification, such as customer's country, occupation and salary, banking products, etc.

Statistical models have become the mainstream practice for customer risk rating, by the application of different techniques to solve the anomaly detection issue. However, this problem is complex to identify or reproduce, and produces imbalanced samples.

The application of advanced data methods allows us to overcome these limitations, improves the customer risk rating accuracy and fosters its relevance along the AML program. The customer risk rating progressively evolve to a behavioral customer risk rating in which continuous data is updated and enriches the risk identification process[75]. Furthermore, models themselves are incorporating the benefit of using machine learning techniques. Supervised methods, such as random forest, are the first to be implemented to unveil hidden relationships between risk drivers in an augmented set of factors.

As the computational power, richness and depth of the data increases, these behavioral models can also incorporate triggers for potential transaction structuring, namely, collective strategies to launder money by multiple individuals through small amounts, to avoid detection by classical static detection strategies. The ability to build algorithms and strategies that run, not at an individual customer or customer plus transaction basis, but on ensembles of customers enables the identification of transaction structuring in a more proactive and effective way. These so-called graph algorithms[76, 77] leverage upon potential connections coming from different sources of information[78]. Moreover, the ability to build a comprehensive network representation of all clients brings the additional value of streamlining the process of alert investigation, amongst others.

Transaction monitoring and filtering

The most common approach to transaction monitoring and filtering consists of a rule-based system, in the style of a decision tree. Each rule is configured to identify a defined behavior masking potential ML activities of the customers and entities involved in the transaction[79]. These rules are generally identified as "scenarios". More complex rules and scenarios try to address the identification of nested accounts and more sophisticated

[73] Jean-Pierre Changeux (b.1936) is a French neuroscientist known for his research in various fields of biology, from the structure and function of proteins, to the early development of the nervous system, to cognitive functions.

[74] However, there is not uniformity in the degree of adoption of these advanced analytical techniques. While some financial entities are experimenting with innovative solutions, simple applications are more usual in the industry, and the reliance on analytic support is at its inception for others. Nevertheless, the present and future of the AML/CTF programs cannot be understood without looking at the new technologies and methodologies available.

[75] For example, incorporating information from transaction monitoring, payments screening or outlier analysis around channels, volumes, geolocation, etc.

[76] Soltani, Reza & Nguyen, Uyen & Yang, Yang & Faghani, Mohammad & Yagoub, Alaa & An, Aijun. (2016. 1-7. 10.1109/UEMCON.2016.7777919.

[77] Scalable Graph Learning for Anti-Money Laundering: A First Look; Weber, M; Chen, J.; Suzumura, T.; Pareja, A.; Ma, T.; Kanezashi, H., Kaler, T.; Leisersen C.E.; Schardl, Tao B.

[78] For example, closed circuits of transactionality – regular transfers-, to joint accounts ownership, single address, branch of choice or mostly visited branches or ATMs, geopositioning via mobile app, coincidence of merchants, etc.

[79] This suspicious behavior will be most likely based on outliers on location, transaction count or transactions amounts.

relationships between parties, but the basis of the outlier identification broadly remains at individual transaction level by looking at the data received during the transactional process. When an outlier is identified, an alert is triggered, which subsequently requires expert evaluation[80].

In this process, the initial set of rules is broken down into a deeper segmentation of behaviors in which the line of business, the level of transactional activity and the risk assessment of the customer determine the final behavioral outliers, i.e., the alerts that would be trigger.

Data analytics methods can be leveraged to detect more quality alerts, increasing true positives and reducing false negatives, i.e. more true alerts are identified without increasing the noise in the identification. Data analytics and machine learning techniques are implemented to optimize the segmentation providing more accurate identification of patterns thanks to the exploration of historical data[81].

Nevertheless, financial entities actively looking into incorporating advanced methods in their AML/CTF program might decide to focus on alert prioritization. The rule approach generates large amounts of alerts even when proper tuning of the scenario thresholds is implemented, and segmentation has been optimized. To address this, many banks implement supervised learning methods to rank alerts in terms of productivity[82]. The key aspect that determines the success of this approach is the utilization of differential metrics, beyond the expected and immovable variables available at transaction level.

The most disruptive approach to AML risk identification consists of abandoning the traditional individual rules approach to unveil hidden relationship with advanced analytics. However, few financial institutions are exploring the utilization of alternative methodologies. Some of these are:

- Graph analytics, which are taking their space in the identification of network relationships and are increasingly determinant of ML activities in the interconnected financial world.

- Clustering techniques, which help to identify outliers without assuming specific behaviors; therefore, capturing more frequently potential new illicit activities.

Advancing towards a non-rule-based approach does not automatically imply abandoning previously identified good optimization practices. In fact, reliance on advanced analytics to improve the customer segmentation, combined with network and outliers' detection, together with the utilization of alert prioritization could be seen as an integral solution.

Sanction Screening

The Sanction Screening engines compare individuals or companies against designated sanction list using fuzzy matching techniques. The most straightforward approaches are based on a wide range of transformations applied to the "names" (name order change, initials, transliteration, common vocal or consonant mistakes, etc.). The transformed names are standardized as strings and compared with the names in the sanction list, also standardized following the same rules. The

[80] See Scalable Graph Learning for Anti-Money Laundering: A First Look; Weber, Chen, Suzumura, Pareja, Ma, Kanezashi, Kaler, Leisersen Schardl, Tao.

[81] Data driven threshold tuning allows to optimize the buckets of increasing productivity along the variables used in the scenarios (more true positives) while providing measures of the potential risk not identified (limiting the false negatives). These common approaches rely on the existing rule-based engines.

[82] This approach may be seen as an imitation of the level 1 analyst review of alerts; however, this could be a more complex identification to address and not all the entities succeed in this effort.

An example of National Risk Assessment

The UK government regularly publishes a national risk assessment[1], which informs on the risks faced at a national level in Financial Crime. This national risk assessment provides references on the most common ML/TF techniques used and their level of implementation in the country and is an important reference for the institutions themselves in their risk assessment.

A firm must perform a Financial Crime risk assessment and use this to inform the design of their AML controls. The national risk assessment therefore serves as a strong foundation to build this assessment from, with the firm taking extra steps to understand, more specifically, the risks they face.

This would take into account, but not limited to, their portfolio of clients and the products they have - personal current accounts serve as a means of tax evasion for many small businesses as well as introducing exposure to many other money laundering techniques due to their ability for rapid fund transfers and accepting cash transactions. Additionally, a review of historical criminal activity can help understand any additional typologies faced by the bank.

Cash transactions, in and out of accounts, serves as an easy way for money launderers to break transaction trails. Whilst the use of cash in money laundering is widespread and is included in many of the strategies used, the controls around cash risks are usually the simplest largely due to the little information available for cash transactions.

Money mules are third parties that are either wittingly or unwittingly used to make additional cash transactions and fund transfers that mask transaction trails. This can be used in conjunction with other strategies, e.g. purchasing high value, resalable assets, to almost completely remove suspicions of the source of funds, where the temporary accounts could be those of a mule network. This is difficult to detect using traditional methods as no single account, and no single customer, may ever be used for large volumes of the transactions used in any stage of this process.

Similarly, cash-intensive businesses serve as another challenge for traditional detection methods. Businesses such as beauty salons, newsagents and car washes are used by money launders to document cash made from criminal activities as legitimate business proceeds so that large volumes of the criminal network's illicit funds can be centralized into one account. This proves difficult to detect as the business's cash income may seem consistent with its own history as well as the income of its peers, and therefore there may be no suspicions raised by the cash transactions of the business. These businesses, however, are commonly also linked to human trafficking and modern slavery, which include their own transactional behaviors that may be easier to detect. As with the usage of money mules, these typologies will commonly involve a network of seemingly unrelated third parties. These third parties may be the facilitators or even the victims of these crimes and therefore there are specific behaviors one would expect to see. Transactions in multiple different cities, especially in cities with transport hubs, heavy usage of fast-food restaurants, multiple transactions in the same hotel on the same day, multiple payments to mobile providers, fund transfers between accounts with similar behaviors, and international transactions especially cash and fund transfers are all strong indicators of these typologies. If these parties can be linked to the cash-intensive business, then the full network could be uncovered.

International transactions are another high-risk transaction identified in the national risk assessment. These are seen in a variety of machine Learning techniques, as well as presenting a risk in other aspects of Financial Crime. This is seen in human trafficking, which is estimated to be one of the largest generators of criminal proceeds globally. Human trafficking requires sending money abroad to members of the associated organized crime gang in the countries associated with the trafficking. This may be as cash withdrawn in the UK and physically moved abroad or via money mules in a similar way as the behavior associated with cash deposits previously described.

Terrorist financing is identified as a high-risk typology within the UK. The raising and moving of funds are not considered a primary goal of terrorists, especially since the majority of recent terrorist attacks have been low-budget and low-sophistication, frequently planned, funded and done by an individual. Terrorist financing is commonly used for moving funds abroad through relatively simple methods such as physically moving cash abroad or employing MSBs. Therefore, detecting terrorist financing requires a collection of key indicators in the same way as required for the usage of cash-intensive businesses in money laundering.

The risk associated with crypto assets grows year-over-year as crypto assets become and more common and easily accessed, but the controls around them remain relatively new with the UK introducing regulations around the usage of crypto assets for money laundering only in January of 2020. Organized criminal gangs use crypto assets for money laundering by first purchasing the crypto assets with their illicit funds, potentially after an initial stage of layering, before selling the assets to provide a legal source of their funds. Additionally, crypto assets can easily be moved across borders allowing criminals to move significant funds internationally with significant ease in comparison to fiat currencies.

This serves as an example of new emerging risks in Financial Crime presenting another challenge for firms to develop and action new controls on a regular basis to keep up with the changes and developments found by money launderers.

[1]HM Treasury: National risk assessment of money laundering and terrorist financing 2020. December 2020.

matching rules or logics measure the degree of separation between the two strings. The engine may return a score of the matching, or an alert based on a pre-defined rule of matching, however, the underlying rationale is the same, i.e., the two strings are similar enough to grant an expert review.

As in the case of Transaction Monitoring, these rules produce a large number of false positives[83]. Furthermore, the potential for optimization based on tuning is lower than in the case of Transaction Monitoring.

For this reason, entities are exploring alternative methods to improve the quality of identification based on translation and transliteration technologies, and the application of NLP techniques to improve the name matching. The improvement in the analytical methods for Sanction Screening run in parallel with the exploration of these techniques in the identification of negative news.

The next steps into analytical approaches to AML/CTF assessment

The application of innovative methods and technologies does not stop at the ones highlighted above. Extended natural language processing and deep learning, blockchain applications, electronic verification of identity, voice and speech recognition, biometrics, or geolocation are other technologies that may contribute to the identification of illicit activities.

Underlying to all these potential approaches, several trends in AML/CTF analytics can be found:

- Deeper analysis of existing data both at the transaction moment, and from the moment the customer and their relationships are implemented. Some of the analytical options outlined above become powerless if differential data is not available and incorporated into the analysis.
- Supplementary data from the internal sources and the different dimensions of the AML/CTF program (i.e., customer risk rating, due diligence, sanction identification, transactions) and external sources (public data on PEP, ownership relationships, reputational sources, open searches) is required to create a holistic approach to the AML/CTF risk identification.
- Technologies and methods can be as complex as innovation allows, however dimensioning the most adequate ones to the nature of the business and risk assessment of the institution is critical to optimize the use of technological and human resources while ensuring regulatory compliance.

[83] Engines can be more or less complex in the incorporation of innovative transformations applied to names, or incorporate more quality sanction sources improved with PEP information, however, they all exhibit the same weaknesses.

Supervisors and regulators are in general reluctant to sudden changes and favor well-stablished methodologies before fully embracing revolutionary changes. However, for those institutions willing to embark in a full transformation program of AML analytics, a number of advances have taken place in recent years[84]: from specific developments of fuzzy matching applications or PEP screening in joint collaborations, to the constitution of innovation centers and sandboxes.

In the journey towards more sophisticated risk identification, interpretability and appropriate risk control remain at the core of the regulator's concerns (and of the institutions).

The use of advanced analytics in the AML/CTF program is linked to the consideration of the implemented rules as models and are therefore subject to the identification, monitoring and control practices that entities have deployed under the Model Risk Management (MRM) function. While the distinction for the customer risk rating is clear, as it fulfills all the conditions typically established in the model risk management (MRM) framework to be a model or at least a user tool that should be monitored, AML/CTF engines have not been initially seen as models. The assimilation of the AML rule engines in the model risk management discipline has not uniformly happened across jurisdictions and main players want to avoid the burden on an incremental scrutiny of the AML programs[85].

Nevertheless, the machine learning technologies to improve risk identification are broadening the conception of what is meant by a model subject to MRM. Despite their willingness to foster their application to AML/CTF programs, supervisors make clear the need for ensuring a proper degree of understanding and interpretability of the methodologies implemented and outputs obtained. Black box models are to be avoided. Machine learning models may suffer from a lack of transparency in the feature selection and explainability, model performance evaluation, etc. Appropriate documentation, testing of the model, interpretability modules; the basic principles of a robust MRM framework will support the adequacy of these models for the AML/CTF use.

Use Case: Enhancing suspicious pattern detection through network analysis

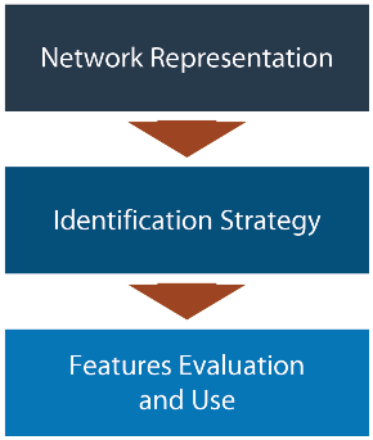
As explained before, one of the primary tools to detect ML is transaction screening and filtering. Whithin this tool, the commonly approach to identifying suspicious ML behavior by fixed patterns of transactional movements, previously described, is being progressively enriched by combining more powerful analytics. The utilization of network analysis for example has been proven to help in the characterization of ML patterns with unique metrics or features. The enriched features may be used in rule-based approaches or in more complex techniques such as machine learning or fuzzy algorithms.

Three relevant stages are at the core of the integration of network analysis in the ML detection: (i) collection of relevant data and construction of the graph representing the relationships between the entities involved; (ii) definition of the identification strategy that would allow to identify the cluster of entities and relationships which are suspicious; and (iii) characterization of those clusters by appropriate metrics which will be used as features of the ML detection models (see figure 4).

Stage 1: Network representation

A network allows to examine complex relationships between related entities, either through links of internal data, such as transactions, or external data, such as addresses and ownerships

Figure 4. Stages for detection by network analysis.



[84] In words of the recent paper issued by FATF, "new technologies have the potential to make AML and counter terrorist financing measures (CTF) faster, cheaper and more effective". Additionally, the FATF enumerates the multiples initiatives of worldwide supervisors and entities that constitute the forefront of the industry evolution See: Opportunities and challenges of new technologies for AML/CTF, available in https://www.fatf-gafi.org/media/fatf/documents/reports/Opportunities-Challenges-of-New-Technologies-for-AML-CTF.pdf .

[85] A join statement by the FRS, FDIC and OCC addressed industry questions on how the MRM guidance should be applied to BSA/AML compliance models. The supervisors consider that not all the systems are required to be classified as models, and the bank itself may categorize models as they see fit. Most importantly, they stated that the banks are not required to have duplicative process or conduct duplicative testing activities to comply with BSA regulations. Although providing certain degree of maneuvering to financial institutions, the statement reinforces the view of bank addressing the risks associated to the AML systems (models or not).

(see figure 5). The construction of a network requires to compute sufficient granular data points that could connect the entities with different objects such as companies, addresses, digital uses, etc. and consider the strength of these relations (e.g., transactional connection). This network can be structured as a graph (both directed or undirected, and weighted or unweighted). The network constructed and the information contained within it will determine the suitability of certain techniques (for example, a weighted undirected graph could be treated in the following steps using clustering techniques, such as spectral clustering).

Stage 2: Identification strategy

An identification strategy is needed to uncover potential money laundering or other illicit activities' patterns within the identified network. There are different strategies that can be used, for example:

- Heuristic approaches based on proximity to confirmed suspicious cases or entities
- Probabilistic approaches and pattern recognition
- Community detection approach, leveraging on machine learning techniques

When implementing the community detection approach, the different communities need to be discovered. A community is a subgraph in the network with a higher number and more intensive relationships among the members of the community compared to random, uninformative subgraphs (see figure 6). Community detection is a useful approach to detect and characterize the targeted structures, which may require the use of algorithms such as k-means, hierarchical clustering, spectral clustering, evolutionary algorithms, or modularity optimization[86].

To find optimal communities, a specific function is optimized: the Modularity Formula. Given a network represented as a weighted graph and partitioned into communities or modules, this formula depends on the specific structure of the graph representation, and expresses the mathematical definition of modularity in terms of weights:

$$Q = \frac{1}{2w}\sum_i \sum_j \left(w_{ij} - \frac{w_i w_j}{2w}\right)\delta(C_i, C_j)$$

Where Ci is the community to which node i is assigned, w_ij represents the value of the weight in the link between the nodes i and j (0 if no link exists), $w_i = \sum_j w_{ij}$, and $w = \sum_i w_i$. Finally, the function δ corresponds to the kronecker delta function: $\delta(i,j)$ takes the value 1 if the nodes i and j are in the same module and 0 otherwise.

Stage 3: Use of features

Once the target communities have been identified within the network, specific metrics or features can be defined to evaluate the depth and importance of the relationships, or the risk of the connections between entities. These features can be used in rules or machine learning algorithms to enhance the predictive capabilities of the models by reducing false positives and identifying better suspicious patterns. The rule-based approach incorporating "enriched" features may be useful to produce qualitative alerts as they incorporate new information apart from the traditional transactional base related to the customer (see figure 8). However, machine learning techniques can unveil stronger relationships which allow to separate true positive alerts and false positive alerts.

[86]Optimal algorithm for pattern detection have been developed by several authors. See L. Alsedà, A. Awasthi, Jörg Lässig (2012).

Figure 5. Depiction of community detection algorithm (steps (a) to (d)) using the modularity optimization.

The upper community contains a suspicious behavior (target) corresponding in this use case to the Customer A network of Stage 1 (shown in figure 7 in the next page).

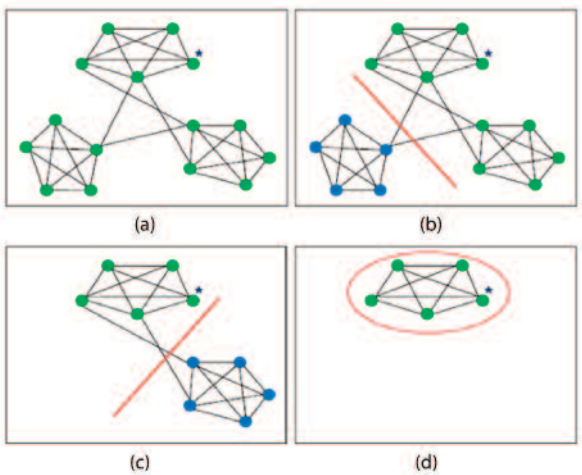
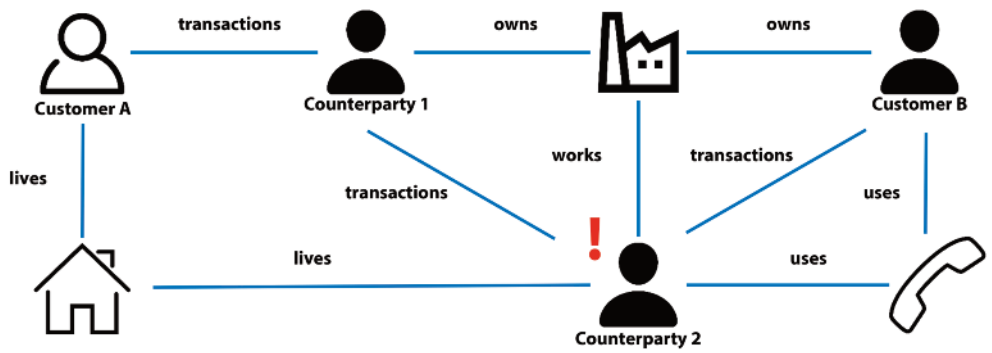


Figure 6. Simplified logic for a scenario based on features obtained through network analysis. Other features such as transactional amounts, may be incorporated according to the pattern or combination of patterns to be detected. This example is for illustration purposes only.

Logic and Parameterization of detection rules with network features:

- If entity type = **Individual**
- If suspicious community identified: suspicious connections identified = **Yes**
- If number of paths to suspicious node > = **num_paths_to_suspnode**
- If distance to suspicious node is <= **max_dist_to_suspnode**
- If primary connections to suspicious node = **primary_connection_type**

Figure 7. Depiction of network relationships related to Customer A including a suspicious node (Counterparty 2 is blacklisted) and possible synthetic entities (nodes related to the Company).



In the example shown opening this user case, figure 7, whose network information is presented in figure 8, Customer A and Customer B pertain to the same suspicious cluster with connections to the suspicious entity (Counterparty 2), but Customer B has the strongest relationship, both personally and professionally with Counterparty 2.

Based on this scenario, if thresholds were calibrated to be num_path_to_suspnode = 1, max_dist_to_suspnode = 5 and primary connection to suspicious node = "all" (either transactional, personal or any type), then both Customer A and B will be flagged as suspicious entities (or their related transactions, etc.). However, considering a more traditional approach, without using the network analysis, only Customer B would be flagged; Customer A does not have transactional connections with the Counterparty 2.

Complex features can be evaluated and different types of machine learning algorithms can be trained resulting in higher risk assigned to Customer B and associated transactions. Incorporating new features into the models also allows to increase the accuracy and detect more potentially risky behaviors (reducing false negative alerts), while discriminating better the risk among those behaviors identified (reducing false positive alerts).

Figure 8. Information on customers for suspicious connections identification.

Entity	Min distance to suspicious node	Primary connection to suspicious node	Personal data connection	Number of paths to suspicious node	Cluster identified	Suspicious connections identified
Customer A	2	Transactional	Yes	2	1	Yes
Customer B	1	Transactional	Yes	4	1	Yes