



Graph Technology and AML

A New Perspective on
Financial Crime

◆ Monocle Research Team



Contents

The Current State of AML	1
The Potential of Graph Technology	2
Graph Algorithms and Anti-Money Laundering	4
Use Cases	6
Where to Start	7
How Monocle Can Assist	8

During Monocle's webinar on the topic of the reliability of regulatory returns and its success through data integration, a question was raised regarding how regulatory reporting within the banking industry should protect the global economy from financial shocks. Francis Gross, a senior advisor for the European Central Bank on the webinar panel, addressed the concept of a developing a "network of contracts", whereby regulators can oversee financial transactions and contracts across global financial markets and link their relationships and dependencies to one another. This would allow regulators the capability to comprehensively view financial impacts and contagion risks and ultimately take preventative action.

For anti-money laundering (AML) operations within financial services, graph technology has emerged as an innovative solution to store and analyse the **network of relationships** that exist in market conduct data – something that traditional relational databases and rules-based transaction monitoring (TM) systems are unable to properly cater for. In this paper, Monocle reviews the current state of AML operations within banking and how graph technology can strengthen banks' suspicious transaction monitoring capabilities.

The Current State of AML

Money laundering starts with the **placement** of illicit funds into the financial systems as criminals look to clean or launder these funds to make them indistinguishable from legitimate sources of money. These funds are broken down into inconspicuous amounts as to avoid the filing of a suspicious transaction/activity report by the banks they operate in. These funds are also transferred between various companies, financial institutions and third parties in order to obscure the identification of the parties involved – a process called **layering**. Finally, the funds are then **integrated** back into the criminal's legitimate financial account. It is these sophisticated and intricate networks of accounts that AML operations must first identify and then unravel.

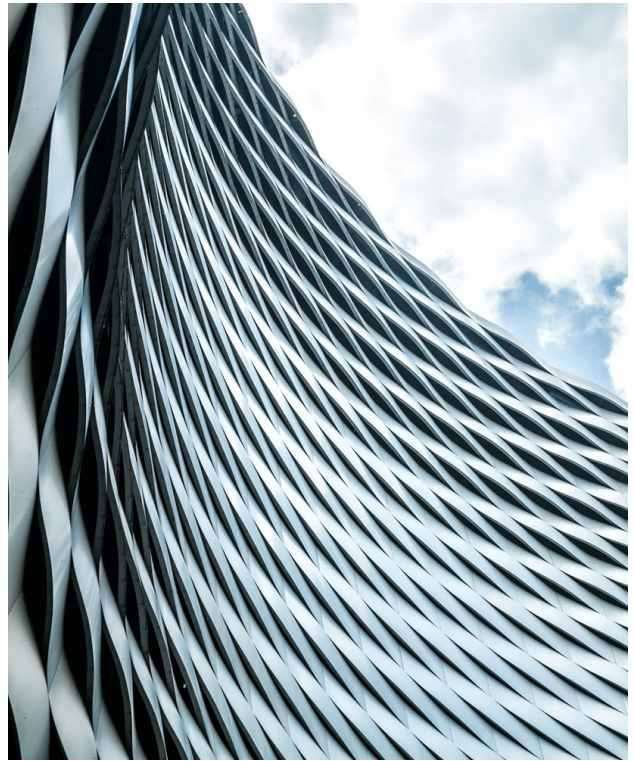
AML Fines in 2021

AML fines in Europe, the Middle East and Africa saw the largest increases in the world throughout 2021 with fines and sanctions increasing from \$1 billion in 2020 to \$3,5 billion in 2021.

For UK financial services in particular, the Financial Conduct Authority (FCA) is set to continue to make greater use of its regulatory powers including far more substantive fines and criminal prosecutions. The need to reinforce AML operations has become of paramount importance due to the acute risk of reputational and financial damage.

Just one of the problems that banks face when combatting money laundering, however, is that over 95% of system-generated alerts from transaction monitoring systems are reported to be false positives.¹ This staggering statistic reflects just how difficult it has become for banks to predict money laundering, as well as has become to predict as well as hinting at the inefficiencies in AML operations that exist at present.

Banks have continued to heavily invest in suspicious transaction monitoring solutions that target specific transaction behaviours primarily through rules-based triggers. These types of traditional transaction monitoring systems can process millions of transactions daily but are often hamstrung by a lack of flexibility in rule construction and can easily make use of rules that have become irrelevant or ineffective as criminals adopt new techniques. In order to refine TM and reduce false negatives, technology firms are promoting graphical databases and knowledge graphs that are designed to align to the nature of money laundering – networks and relationships.



The Potential of Graph Technology

Modern bank data infrastructure is built primarily using relational databases. While the reliance of storing data in rows and columns and linking data through primary and foreign keys has made data easier to manage, model and visualise, it has its limitations. Relational databases are often slow and rigid, requiring analysts to join multiple tables to produce consolidated views. AML investigations are hindered as data must be consolidated to determine payment trails and client networks across various datasets.

In contrast, graph databases and the use of knowledge graph data models are entity orientated rather than table orientated and therefore, allow for the modelling of numerous relationships between entities (accounts, clients, customer details), as we would consider them in real-life. Knowledge graphs (how data is modelled in a graph database) are built using “triples” or “triplets”, consisting of **nodes** (the subject and the object) and **edges** (the relationship between the subject

and object). These models can be analysed using mathematical graph theory and graph algorithms to uncover relational insights.

To allow for additional dimensionality (and better machine readability), graph databases make use of taxonomies and ontologies. Taxonomies include names for objects and fundamental relationships, while ontologies define the types of nodes and relationships, classify concepts into meaningful categories, ascribe attributes to nodes and edges and define possible relationships between nodes (much like schemas). Using an ontology as a framework and data together you create a knowledge graph.

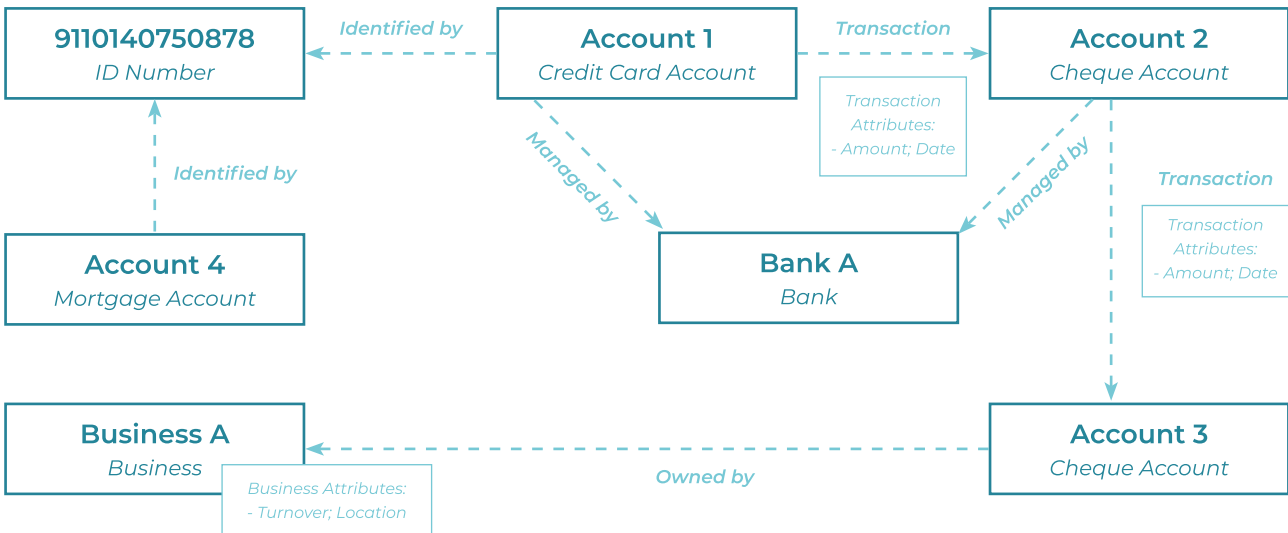


1. Reuters. (2018). *Anti-money laundering controls failing to detect terrorists, cartels and sanctioned states*. Available at: <https://www.reuters.com/article/bc-finreg-laundering-detecting-idUSKCN1GP2NV>

Triple/Triplet



Knowledge Graph



Knowledge graphs are effective in producing visualisations that are easy to understand and explain. Above is an example of a knowledge graph using various data from across a financial institution.

This results in a polyhierarchical model, where entities are categorised multiple times to produce a graph-view. This is one of the key advantages of graph technology in that the data is represented in a way that makes sense to the human brain and is considered a close representation of how we would “whiteboard” these data landscapes during planning sessions.

Furthermore, knowledge graphs do not rely on joins that can become tedious to code as well as constrain memory and CPU resources to execute. From an AML perspective, this is a problem as investigations are often slowed by the process of consolidating disparate datasets when examining suspicious transaction events. Graph technology allows for considerably faster ad-hoc querying to identify relationships that would often be unattainable through SQL and relationship databases.



Developments in Graph Technology

While graph theory, mathematical modeling of relations between objects, has existed for several centuries, graph databases and knowledge graphs (graph specific data models) have only started to gain popularity over the past 20 years.

With its origins linked to the idea of the Semantic Web (the internet that is machine readable through ontologies) as popularized by the inventor of the world wide web, Tim Berners-Lee, graph technology is playing a crucial role in data management for companies like Amazon, Microsoft, and Google. Similarly, the mathematics of graph theory supports various algorithms designed for data analysis of graph databases.

Gartner predicts that graph technology will reach majority adoption in three to six years and with a high “mass”, is set to have a substantial impact as a technology driving business, process and data changes in the market.²

This makes graph technology an ideal prospect to consider now for the future of financial institutions’ AML operations.

Graph Algorithms and Anti-Money Laundering

Money laundering by its nature is a complex network of transactions or relationships between accounts which makes it an ideal use case for graph databases and knowledge graphs. The appeal of this technology is the capability to connect the dots across various data sets to uncover previously hidden networks, using graph algorithms. These algorithms can query the data to produce insights and metrics beyond the capability of relational database queries. The most common these algorithms are:



Pathfinding

Arguably the most common graph algorithm, pathfinding is used to explore routes between nodes to identify optimal routes and the shortest paths between a starting point and a destination. This is useful in determining degrees of separation between nodes and can be used to identify

payment trails between nodes of interest such as political exposed persons. Google makes use of shortest path algorithms as part of their Google Maps offering to calculate optimised routes and provide driving directions. Pathfinding can also measure “**betweenness**” which measure the number of shortest paths a node is on which is related to centrality.



Centrality

These algorithms are designed to measure the impact or importance of certain nodes in a network and have been used extensively for social network analysis. There are a variety of centrality metrics that can be used to determine key actors amongst AML networks, particularly when layering money across various accounts. Centrality metrics

2. Gartner. (2021). *5 Impactful Technologies From the Gartner Emerging Technologies and Trends Impact Radar for 2022*. Available at: https://www.gartner.com/en/articles/5-impactful-technologies-from-the-gartner-emerging-technologies-and-trends-impact-radar-for-2022?utm_medium=social&utm_source=linkedin&utm_campaign=SM_GB_YOY_GTR_SOC_SFL_SM-SWG&utm_content=&sf253358796=1

include **degree**, the number of connections a node has, PageRank (eigenvector centrality), determining a node's importance based on the importance of its relationships to other popular nodes, and **betweenness** as mentioned above. Twitter makes use of PageRank in determining personalised recommendations to its users including accounts they may be interested in following.

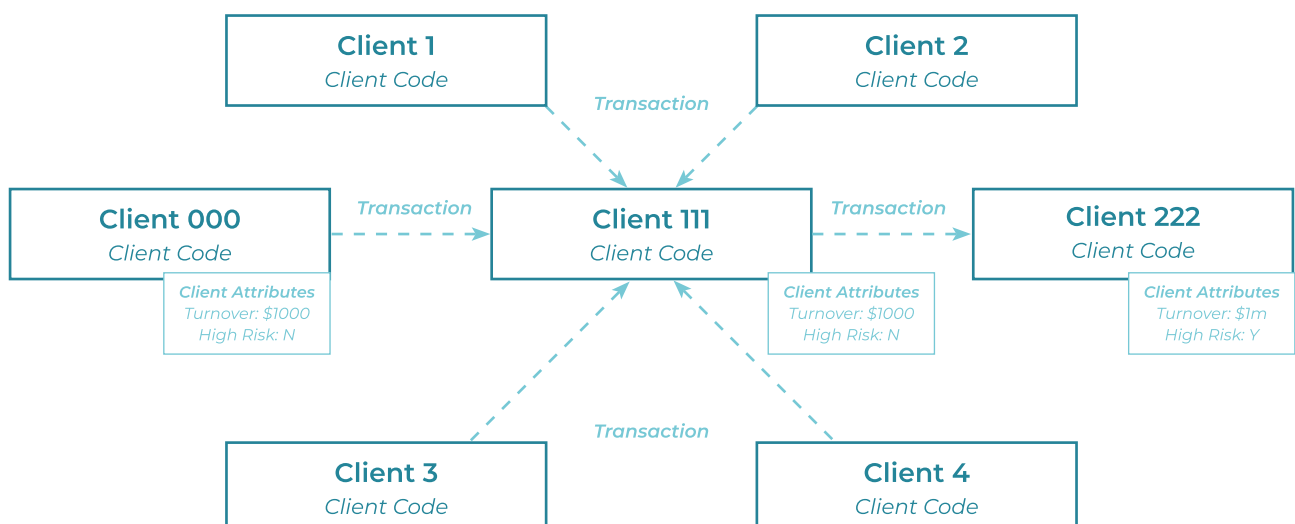


Community Detection

Also known as clustering, these algorithms can identify communities of nodes that share similar edges. This can be used to identify syndicates of money laundering intermediary accounts due to their connection to other suspicious parties. There are various algorithms that can be used to determine clustering from **triangle count** and **clustering coefficients**, that measure nodes that form triangles through their edges, to more complex algorithms that incorporate relationship weights and degree metrics.

Significantly, graph technology can support other AML monitoring approaches by developing metrics or variables across nodes that can be fed into your traditional rules-based transaction monitoring systems or machine learning models. These metrics help to build more complex rules, as well as providing valuable insights to train ML models to identify hidden AML nuances and trends not yet identified by current typologies.

AML Specific Knowledge Graph



In the above example, a pathfinding algorithm would detect that client 000 is indirectly related to client 222 through their relationship with client 111 through their transaction history. The algorithm can be set to determine the shortest path of transaction from one node to nodes that have been flagged as high-risk (attribute of the node) and help to identify flows of funds towards suspicious parties. Client 222 has been identified as high-risk which may lead investigators to examine the legitimacy of the flow of these funds in this payment network.

Similarly, Client 111 is not considered a suspicious person, but a centrality algorithm would detect that client 111 is receiving multiple payments from various client nodes, making it an important intermediary to client 222 who is considered a high-risk person. Investigators can easily understand the importance of client 111 in a visualised knowledge graph due to high degree of relationships.

Use Cases

Association to Suspicious Entities

Using pattern-matching scores, direct relationships are identified between parties and then scored based on their proximity to suspicious entities. These scores are based on the quantity, quality and distance of the relationship and can be set up to identify relationships across transactions and the flow of money to politically exposed persons (or the flow of unusual transactions across high-risk geographies or unrelated industries).

Similarly, centrality and community detection can be useful in identifying money laundering techniques such as structuring (also known as smurfing). Structuring is a money laundering technique whereby illicit funds are broken into multiple less conspicuous amounts and then transferred through various parties the financial system until they reach their intended target (a central party using a community of intermediaries). These payment chains can be detected through various patterns, where one node received a high number of incoming deposits, for example, and sends a few large transactions to high-risk parties (or is a high risk high-risk party themselves). Graph technology is ideal at performing queries to discover these networks across banks' transactional data.

Shared Attributes:

The application of graph algorithms can be expanded further to also find similarities in various attributes including personal details such as phone numbers, physical addresses and IP addresses. Analysts can determine relationships once a shared attribute is

identified between separate parties' accounts which can highlight cases of fraudulent identity by criminals who have setup multiple accounts to be used for layering and other money laundering activities.

Single View of Customer

Graph databases and knowledge graphs have applicability far beyond just AML, from supply chain management to the algorithms behind Google Maps. Entity resolution through graph theory lends itself well to financial institutions looking to clean their customer data to produce a single view of each customer.

With many business units operating in data siloes, the duplication of customer accounts across various systems leads to disparate records. When looking to establish a single version of truth, graph theory enables the identification of shared attributes that can confirm if various records are related to one consumer.

Where to Start

Introducing any new technology can be a daunting task however, graph databases and knowledge graphs are supported by a reputable list of vendors including Oracle, Neo4J and SAS. SAS in particular offers OPTGRAPH procedures for graph algorithms and network analysis.

1. Database and Data Model Setup:

Setting up a graph database will require the development of your taxonomy and ontology to map out and model your data. This requires determining what will constitute your nodes (classes) and your edges (relationships). From an AML perspective, it makes sense to establish your customers and clients as your nodes and transactions as edges. In this way, you can begin to build payment networks between your various client nodes and add additional attributes (i.e. turnover, name, accounts, address) as required to add further detail to the knowledge graph.

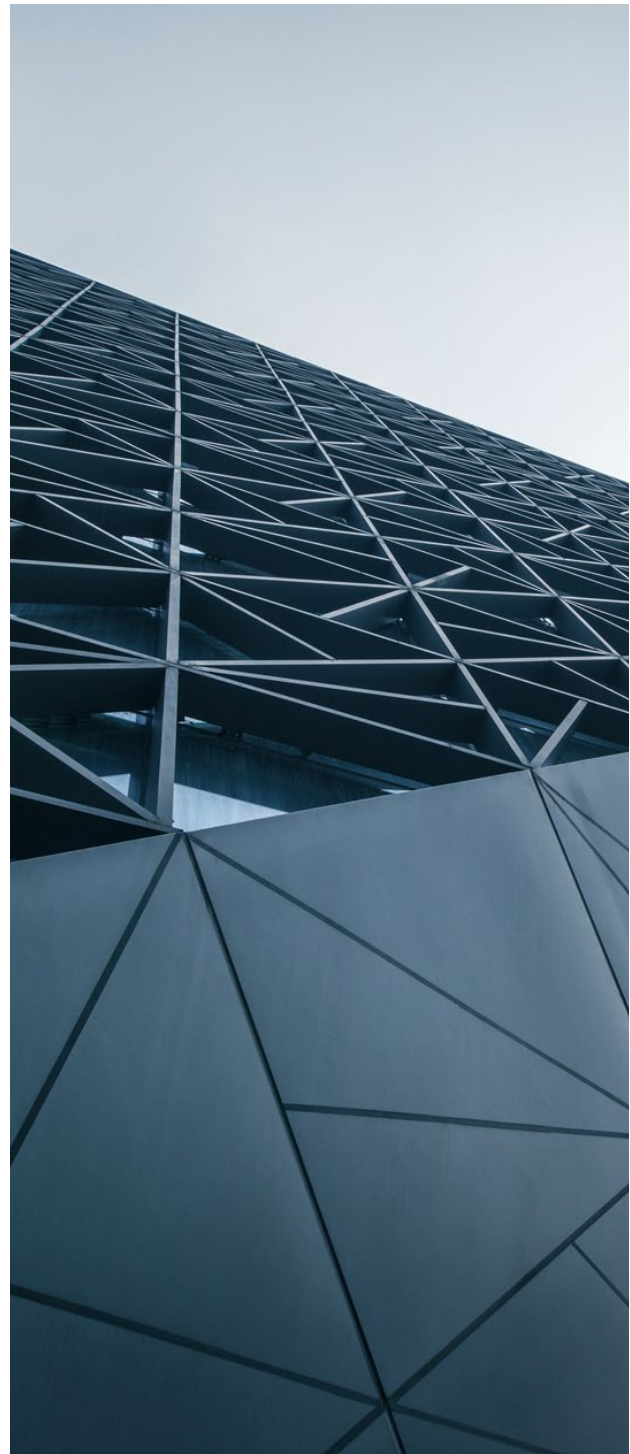
2. Data Preparation

Once this is established you will need to load your data sets into your graphical database. From an AML approach, you will require customer data to create your nodes. Following this, to build out your knowledge graph, you will need to load your transaction data that includes data of both the sender and recipient to establish the relationships across the network.

Once your graph has been created, you can begin to write your various algorithms as mentioned to above to identify AML activity directly or produce metrics that support your established transaction monitoring systems.

3. Graph Algorithms

Once your graph has been created, you can begin to write your various algorithms as mentioned to above to identify AML activity directly or produce metrics that support your established transaction monitoring systems.



How Monocle can Assist

The use for graph technology in financial services is becoming increasingly appealing as the technology matures and its use cases become more apparent. At Monocle we have extensive experience in AML and financial crime operations and data to assist our clients in designing and implementing their graph databases, knowledge graphs and AML graph algorithms.

Monocle's expertise in data management, data transformation and the automation of many of these processes and controls ensures that we can support our clients to successfully integrate graph technology into their current operations, as well as producing robust data quality that meets business requirements and regulator expectations. Furthermore, we have specialised skills related to financial crime dashboarding and reporting across various platforms and vendors.



About Monocle

Monocle is the largest independent management consulting firm in South Africa specialising in banking and insurance. Since our establishment in 2002, we have worked with industry-leading banks and insurance companies around the world, including institutions in the United Kingdom, Europe, Scandinavia, Asia, South Africa and throughout sub-Saharan Africa.

We design and execute bespoke change projects, from start to finish, bridging the divide between business stakeholders' needs and the complex systems, processes and data that sit under the hood. We offer several unique capabilities to our clients, which have been forged over time through the combination of a highly specialised skillset and extensive experience working with the systems, processes and people that are at the heart of the financial services industry.

Johannesburg

13th Floor, Greenpark Corner,
3 Lower Road, Morningside,
Sandton, South Africa

Phone: +27 (0) 11 263 5800
Fax: +27 (0) 11 263 5811
Website: www.monoclesolutions.com

Cape Town

301 New Cumberland,
163 Beach Road,
Mouille Point,
Cape Town, South Africa

Phone: +27 (0) 82 952 1415
Website: www.monoclesolutions.com

United Kingdom

1 Royal Exchange,
London,
EC3V 3DG, England

Phone: +44 (0) 2071 902 990
Website: www.monoclesolutions.com

Amsterdam

Weteringschans 165 C,
1017XD Amsterdam,
Netherlands



MONOCLE

Monocle Solutions © 2022