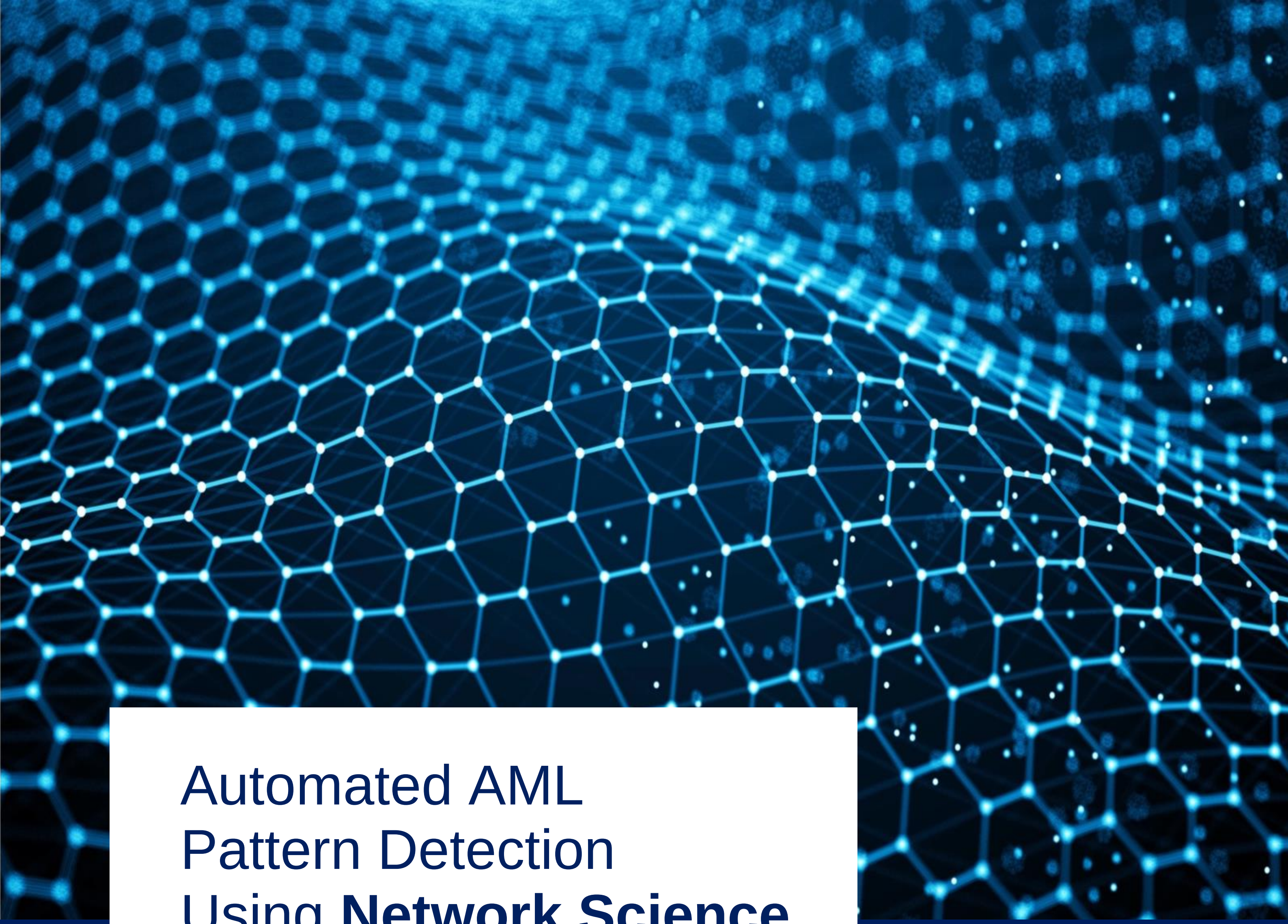


Automated AML Pattern Detection Using **Network Science**

Tookitaki Whitepaper

August 2021

Table of Contents

Executive Summary

Anti-Money Laundering (AML) is a significant burden for large and small banks, irrespective of their line of business. Money laundering techniques are evolving faster than we could think. In order to detect sophisticated organised crime, there is a compelling need to bring to light concealed connections among multiple but seemingly unrelated human money laundering networks, ties among actors of those schemes, and amounts of funds transferred among those entities. However, legacy systems with static rules and traditional approaches like configuration of new rules and thresholds are ineffective as they generate ultra-high false positives and fail to identify false negatives buried deep inside the mountain of legitimate transactions. Result: Suspicious Activity Reports (SARs) are filed late or worse no filing at all.

Graph analytics also called network analysis, a modern technology that analyses relationship between entities, can be used to address this growing problem. Today, there are many graph analytics solutions being used in fields like social network analysis, fraud detection, supply chain and search engine optimisation. In the area of AML too there are graph analytics solutions that show a massive graph linking customer and related activity. However, they fail to automatically detect interesting connections, analyse and create relevant sub-networks (subgraphs) responsible for suspicious money movement from many legitimate or normal links. Although they aid as good visualisation tools, they fail to provide real benefits when it comes to financial crime detection.

This paper details Tookitaki's Network Science module which helps its Anti-Money Laundering (AMLS) solution to automatically and accurately detect complex money laundering patterns. In line with Tookitaki's overall efforts to use machine learning and big data analytics for AML purposes, the module was developed with ways to avoid subjectivity and ensure scalability. At Tookitaki, we envision that this advanced data analytics approach will enable financial institutions capture complex money laundering transactions and stop the bad actors with high accuracy and speed, improving returns and risk coverage.

Tookitaki's concept of ***Automated AML Pattern Detection Using Network Science*** was awarded the Monetary Authority of Singapore's Financial Sector Technology and Innovation (FSTI) Proof of Concept (POC) grant on 17th December 2020. The FSTI POC grant provides funding support for experimentation, development and dissemination of nascent innovative technologies in the financial services sector.

Financial Crime Detection Challenges

Money laundering transforms profits from illegal activities like corruption, forced labour, drug and arms trafficking into seemingly legitimate earnings by concealing the source of the acquired funds. Global estimates show approximately US$800 billion to US$2 trillion is laundered annually through the global banking system. That's roughly 2 to 5 percent of global GDP[1].

The risk of money laundering is increasing with introduction and large scale adoption of new payment methods like mobile wallets, digital currencies like Bitcoin, Ripple and such others. Today's money laundering landscape is complex and evolving. Detection of sophisticated organised crime involves unravelling concealed connections among multiple but seemingly unrelated human money laundering networks, ties among actors of those schemes, and amounts of funds transferred among those entities. Legacy rules systems and traditional approaches like configuration of new rules and thresholds generate ultra-high false positives and fail to identify false negatives buried deep inside the mountain of legitimate transactions. Result: Suspicious Activity Reports (SARs) are filed late or worse no filing at all.

In recognition of the growing problem, financial institutions (FIs) are exploring new technologies like graph analytics to examine connections between entities and better illuminate relationships. Today, the available graph analytics solutions show a massive graph linking customer and related activity but fail to automatically detect interesting connections, analyse and create relevant sub-networks (subgraphs) responsible for suspicious money movement from many legitimate or normal links in the massive graph. Such limitation calls for manual effort to query and extract the information needed and overlay with deep AML knowledge to accurately identify the relevant connections/networks (subgraphs) that showcase illicit fund transfers. Moreover, straightforward queries do not yield required results for complex AML patterns. Although the available graph analytics solutions are great visualisation tools and aids in investigation, they are time-consuming, loaded with subjectivity and limit scalability.

Today, the available graph analytics solutions uncover suspicious behaviours via network anomaly detection, leveraging classical graph techniques like cycle detection, PageRank, degree distribution, label propagation, community detection, etc., as well as the recent development of deep learning in graphs. But these techniques fail to automatically detect relevant subgraphs indicating the suspicious money trail from thousands of legit links and patterns. A subgraph can be defined as a graph whose nodes and edges are subsets of a bigger graph. In a money laundering scenario, the nodes represent entities and their attributes and edges represent the links or connections between them. Automatic detection of relevant subgraphs among voluminous, diverse data that are remote from suspicious transactions is difficult primarily because of (i) a single money laundering scenario is often buried in the graph containing a large number of subgraphs that are not relevant (ii) frequent patterns with many automorphisms and overlapping embeddings are observed in the graph and (iii) missing AML domain knowledge combined with right algorithms to pinpoint the accurate subgraph/s representing the money laundering scenario, which can provide the right contextual explanation to a bank's investigating team for faster decision-making.

[1] https://www.unodc.org/unodc/en/money-laundering/overview.html

Tookitaki Case Detection: A Two-Pronged Approach

The proposed solution is an automated new case detection module that leverages a combination of network science and a growing repository of AML patterns. The solution is a paradigm shift as it automates the process of converting the AML domain knowledge into multiple signals that can be easily ingested by network models to detect relevant suspicious patterns/subgraphs from a single large connected graph with millions of nodes and billions of edges, helping financial institutions to increase the coverage and accuracy of filing Suspicious Activity Reports (SARs) on time. Outcome: Reputational and financial risks are minimized significantly with limited cost.

Case Detection Components and Workflow

The new case detection module uses a two-pronged approach. Picture 1 in the next page depicts the workflow of the approach in a visual manner:

(i) First, relevant data across customer/entities, transactions, counterparty, interested parties (C2C, C2A) and external data sources (watchlists and data providers like Bureau Van Dijk) are ingested and pre-processed into our standardised data scheme

(ii) The standard data inputs are then fed into an intelligent risk indicator engine which uses advanced machine learning techniques to automatically generate thousands of risk indicators (in data science terms they are called features)

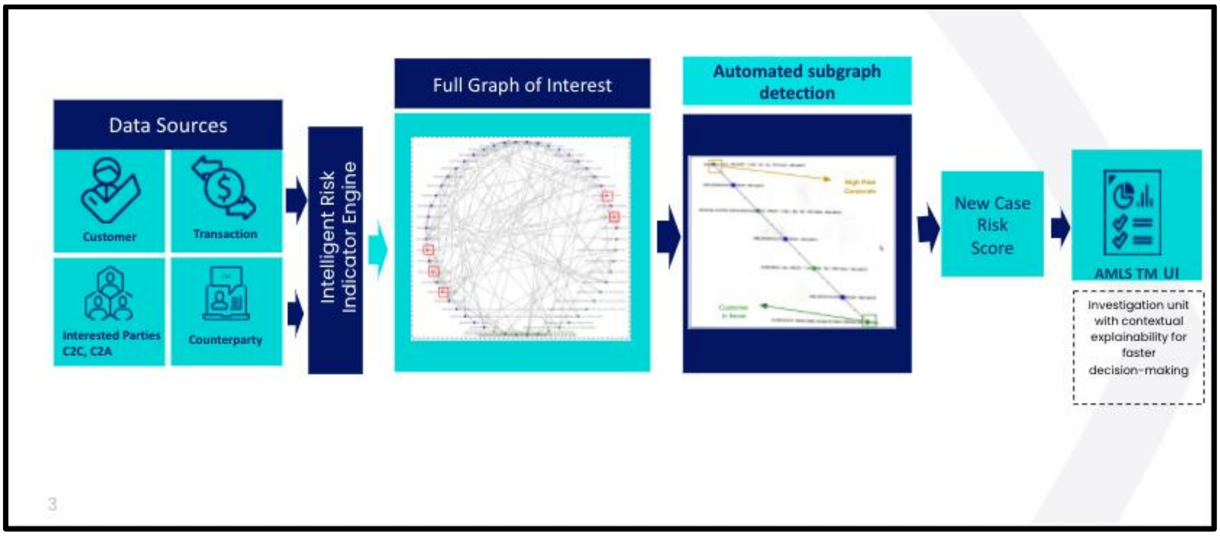
(iii) These risk indicators are fed into a network model to get the full graph with millions of nodes and billions of edges

(iv) Now, graph algorithms like (a) cycle detection (b) community detection including label propagation and Louvain algorithms (c) path finding linking targeted nodes to high-risk entities (d) spectral localization to find abnormal nodes and (e) representation learning like graph convolutional networks are applied in conjunction with predefined AML deep domain risk indicators to accurately detect the subgraph/s of interest.

Examples of predefined AML risk indicators are (a) for entities: incorporation in tax haven countries, business type of high risk etc (b) for transactions: overseas fund transfer amount and volume of the cycle, total cash withdrawal amount and cash flow through and such others (c) for counterparties: transactions between the parties whose business are not related, money flow to the accounts that belong to the same beneficial party etc.

(v) A risk scoring technique leveraging graph analytics and machine learning is applied generating risk scores for each node and subgraph. Based on this methodology, the solution automatically flags out the suspicious cases buried deep inside the mountain of legit transactions.

(vi) All the newly detected suspicious cases have scores and are ranked in order for a bank's AML team to investigate. For faster decision-making, the solution leverages Tookitaki's patent-pending XAI (Explainable AI) framework and provides global and local explainability. The global explainability helps to explain the overall model decision-making and the local explainability provides the context and breaks down the subgraphs into simple English explanations that can be readily consumed by AML investigators.



Investigation unit with contextual explainability for faster decision-making

Picture 1: Case detection using network science workflow

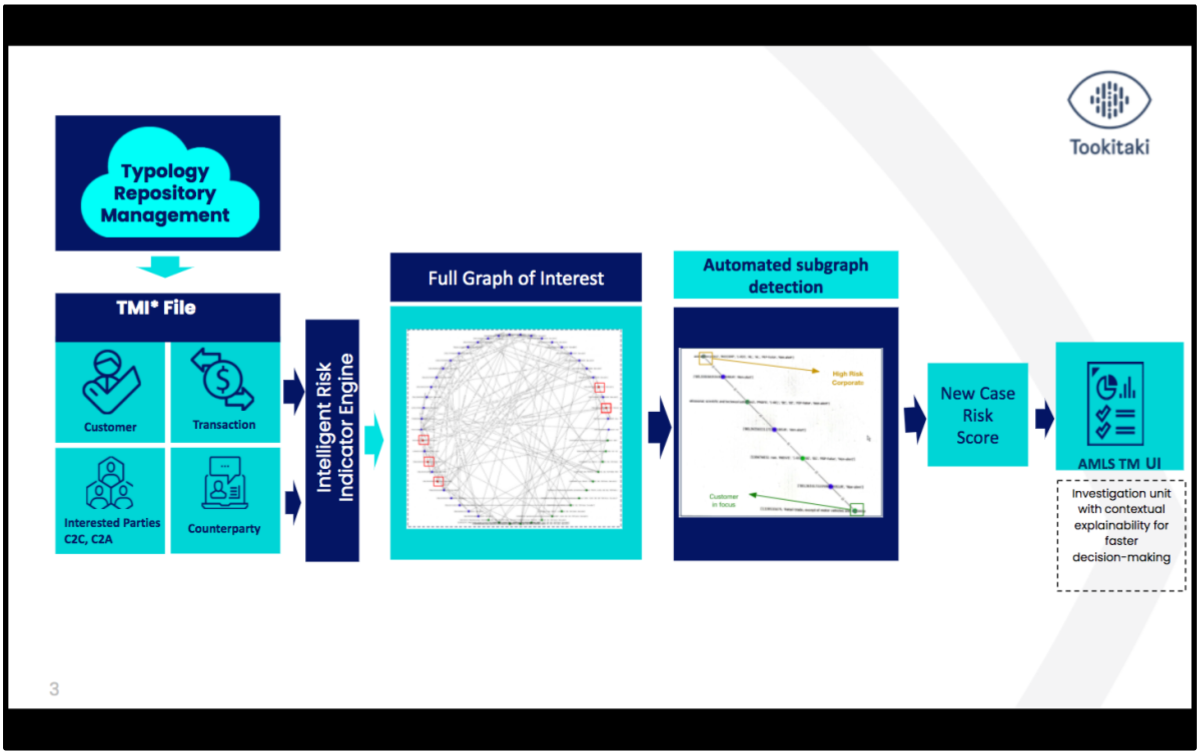
The solution further is enhanced by a second layer where FIs can get access to a growing repository of AML patterns and detect new cases for target typologies of their choice. This is made possible as we integrate the new case detection module with the recently-launched typology repository management (TRM) system. TRM is a growing library of AML patterns contributed by regulators, AML experts and financial institutions on a continuous basis. Additionally, it can improve the predefined AML risk indicators set that is used in combination with aforementioned graph query algorithms to auto-detect subgraph/s representing the suspicious money trail. Picture 2 in next page depicts the end-to-end workflow of the two-pronged approach in a visual manner:

(i) From the TRM dashboard, users can ingest typologies of their choice. The typologies are stored as AML patterns representing suspicious behaviour. A money laundering pattern in TRM comprises four key components – entity, counterparty, transactions and network relation. None of the components hold any personally identifiable information (PII) data and any rules and thresholds.

(ii) The four components are broken into properties (a) value, volume and velocity of money transactions (b) channels and products of money transfers (c) geographical property, industry type and watchlist info of counterparties involved in money transfers (d) money flow characteristics across multiple entities (e) relationships between different entities and (f) other static entity information like occupation, industry, type, location, adverse media etc.

(iii) These business properties are translated into machine-readable inputs and consumed by the intelligent risk indicator engine. The translation has been made possible through a middleware, which is a proprietary framework converting AML knowledge into machine instructions.

(iv) The intelligent risk indicator auto-generates all relevant risk indicators for ingested typologies and the process follows steps (iii) to (vi) of Picture 1. The new case detection module flags out the suspicious cases of the injected typologies.


Tookitaki
Typology Repository Management
TMI* File
Customer
Transaction
Interested Parties C2C, C2A
Counterparty
Intelligent Risk Indicator Engine
Full Graph of Interest
Automated subgraph detection
High Risk Corporate
Customer in focus
New Case Risk Score
AMLS TM UI
Investigation unit with contextual explainability for faster decision-making
3

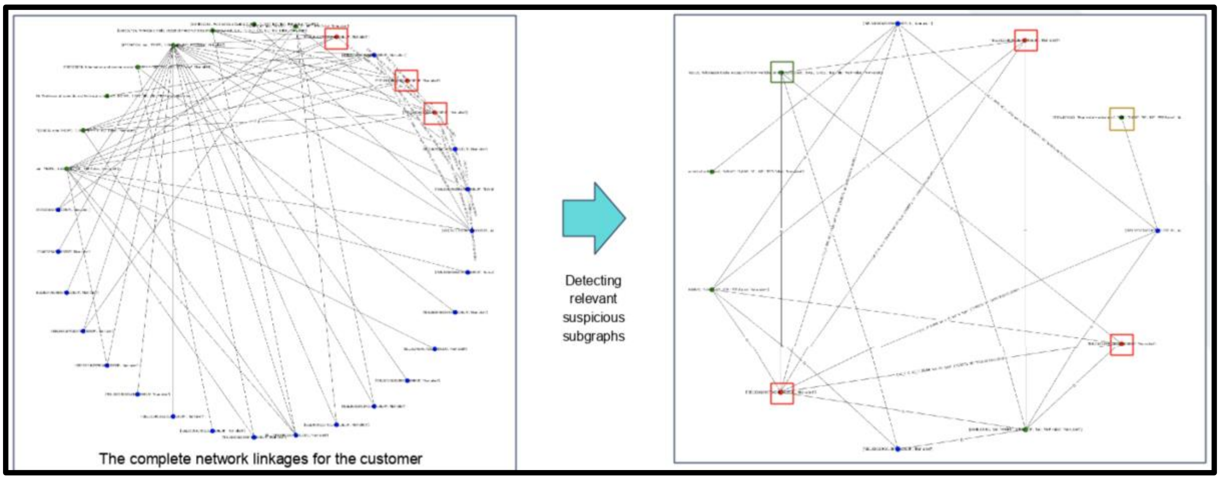
Picture 2: TRM and case detection – Two-pronged approach

Example of a suspicious case

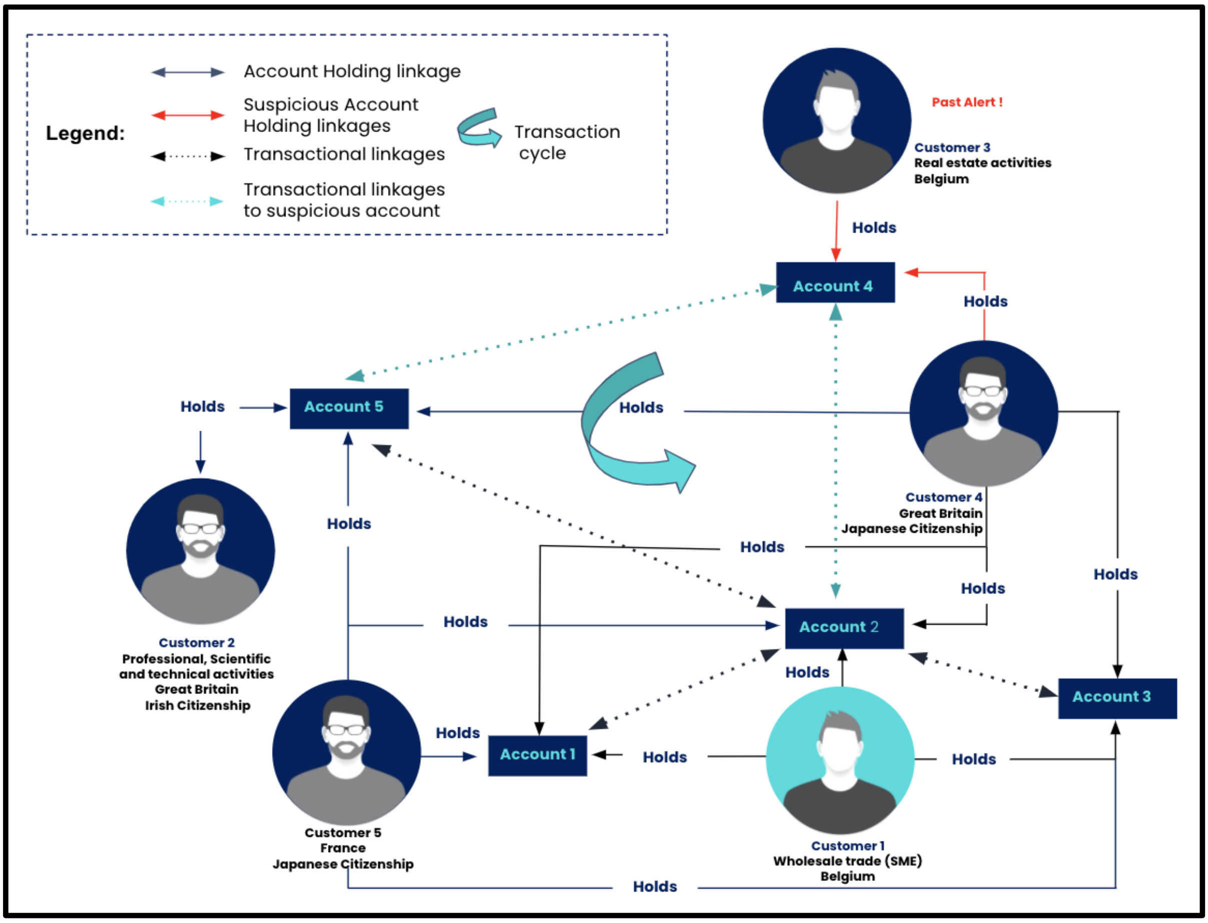
To validate the above hypothesis we developed a working prototype and tested it on a pilot exercise we carried out with a large European bank. We had lots of data limitations in the pilot - no counterparty information, watchlist etc. Picture 3 in the next page is a specific example of a suspicious case, which we discovered during the pilot and it showcases the full graph and subgraph. Picture 4 is a visual representation of the subgraph for illustration purpose only.

Here is the detailed explanation of the suspicious activity:

(i) The new case alert is generated on Customer 1, which is a SME corporate account and is into wholesale trade business. It is designated as a low risk customer by the bank.
(ii) The subgraph shows multiple entities across Europe funnelling money to a historical alerted entity in real estate business through the accounts jointly held by these multinational entities.
(iii) The subgraph also depicts cyclic transaction behaviour between these joint accounts
(iv) Further, the flow-through indicator of outgoing fund transfer over incoming transactions for the last 90 days is very high for Customer 1. At the same time, the customer in focus showed a self-profile deviation on its average monthly transactional amounts over the last three months as against the remaining nine months (a) ~4x spike on overall transactional activity and (b) 10x spike on overseas fund transfers.


The complete network linkages for the customer
Detecting relevant suspicious subgraphs

Picture 3: Example of a suspicious case



Picture 4: Visual representation of the subgraph

Benefits to the Industry

Since the working prototype generated encouraging results, we are looking forward to developing the full-scale solution and launch in the market. We believe the industry will benefit immensely from the solution as it is a novel approach to detect complex money laundering scenarios involving cryptocurrency and other new payment methods. It will also help in the detection of ultimate beneficial ownerships (UBOs) among others. The following are the perceived benefits of using our solution:

Automated detection of money laundering scenarios

Financial institutions typically deal with thousands of transactions every day. It is often difficult for them to pin point money laundering scenarios buried deep inside a large pile of transactions. With network analytics, our solution enables automated detection of subgraphs representing money laundering scenarios among thousands of normal transactions. It is done with zero manual effort, without subjectivity and no information loss. The end benefit is no more missed SARs and late SAR filings.

Faster alert disposal

For compliance analysts, it is often time-consuming to triage and investigate alerts as they need to manually do analysis and verify information. Our solution provides contextual explainability of the subgraph in easy-to-comprehend business language so that analysts need to spend much time in searching for information or doing maths. It enables faster alert disposition with 50-60% reduction in effort.

Intelligence on emerging money laundering patterns

The new cased detection module has been developed in such a manner that financial institutions can ingest machine readable typologies from Tookitaki Typology Library in a seamless manner. This feature equips financial institutions with the intelligence on the latest money laundering strategies. Our solution allows quick testing and adoption of new and emerging typologies with a guided user interface that requires minimal manual inputs.

The solution benefits is summarised in the following table:

	Approach	Benefits
1	Automated detection of subgraphs representing the money laundering scenario buried deep inside hundreds & thousands of normal transactions	Zero manual effort. No subjectivity and information loss. SARs are not missed and filed on time.
2	Contextual explainability of the subgraph	Faster alerts disposition with 50-60% effort reduction
3	Seamless ingestion of AML patterns into the new case detection module without any manual intervention	Allows quick testing and adoption of new complex typologies with limited manual inputs and a guided user experience.

Concluding Thoughts

As money laundering has become increasingly complex, especially during the pandemic times, financial institutions are finding it hard to detect and report suspicious activities on a timely manner. They need to go beyond regular analytics that relies on statistics, computer programming and operations research to uncover insights into financial crimes. Graph analytics with its techniques such as clustering, partitioning and shortest path algorithms can better analyse relationships between entities and prove highly helpful in the battle against money laundering. However, existing graph analytics have not gone beyond improving investigations and they have proven to be ineffective in finding the bad actors with actionable insights. It is time for financial institutions to move beyond the current state and seek synergies to adopt modern technology in an efficient manner if they want to achieve the goal of sustainable compliance. For this it is essential they embark on a journey with next-gen technology partners who present strong technology capabilities to prioritise solutions with demonstrated efficiency and effectiveness improvements with reduced load on human resources and enhanced risk coverage.

Tookitaki

North America	Singapore	India
+1-704-444-0435	+65-6250-2620	+91-804-951-8262

www.tookitaki.ai **Request demo:** sales@tookitaki.com