CHAPTER 2

LITERATURE REVIEW

2.1 Challenges to Detect Suspicious Transactions in Money Laundering

Financial institutions are dealing with a number of challenges to identify money laundering activities (Prisznyák, 2022). It is challenging to identify suspicious transactions with regard to money laundering activities as there are no absolute universal regulations that define standards for what makes a transaction suspicious. This is because money launderer's tactics are always evolving to avoid detection by the system, hence the rules need to keep updating in order to capture the money laundering's latest modus operandi (Labanca et al., 2022). Furthermore, the complexity of detecting suspicious transactions arisen with large volumes of transactions data. In this digital era, financial services are available online for the convenience of their customers to perform transactions at anytime and anywhere. Due to vast volumes of transactions with evolving money laundering tactics, mitigating this financial crime has become more complex than ever before (Kute et al., 2021).

In addition, it is reported that most transaction monitoring models used by financial institutions have a false positive rate over 98 percent, where legitimate transactions are wrongly flagged as suspicious transactions (Buehler, 2019). These false alarms are making compliance officers wasting their time conducting unnecessary investigations while the real money laundering activities continue to happen behind their back (Oad et al., 2021). Moreover, financial institutions are confidential in nature. Due to privacy reasons and specific regulations, real transaction dataset are hard to obtain, hence most of the research are using synthetic dataset generated based on money laundering patterns defined by Financial Action Task Force (FATF) (Labanca et al., 2022). This creates a gap between the application in simulation environment and the current method available in real world to solve the problem (Kute et al., 2021).

2.2 Common Suspicious Transactions Indicators for Money Laundering

There are some general criteria that can be measured to identify suspicious financial transactions in money laundering activities. Financial transaction is the activity of transferring funds, investments or other assets that can be performed through many ways such as wire transfers, checks and credit cards. The existence of unusual financial transactions and customer behaviours are considered as anomalies with high risk for money laundering (Labanca et al., 2022). The anomalous transactions are defined by different aspects such as time, type of transactions, frequency of transactions, amount of money involved, and level of internationalization (Tundis et al., 2021).

Five categories of transactional anomalies are described below as per suspicious patterns' examples provided by Financial Action Task Force (FATF) (Labanca et al., 2022):

(a) *Small but highly frequent transactions generated within a short timeframe:* A lot of small transactions just below the threshold limit in a short period of time might be one of money launderer's tricky way to avoid suspicious detection.

(b) *Transactions with rounded normalized amounts bought or sold within an account*: Real trades in capital markets often involve non-rounded amounts. Therefore, it is unusual to have a perfectly rounded transaction amounts such as $100,000.

(c) *Security bought or sold at an unusual time*: Transactions involving trading of securities usually occurs during normal hours of stock markets. So, it might be suspicious if the trading takes place outside the specific timeframe of stock exchange.

(d) *Large asset withdrawal:* The account suddenly withdraws or transfers a large amount of money that highly deviates from the customer's normal transactions and does not match any valid business reason.

(e) *An unusually large amount of collateral transferred in and out of an account within a short period of time*: It is not common for an investment to move large amounts of collateral in and out of account so quickly as people did not simply trade collateral only.

2.3 Typical Typologies of Money Laundering Activities

Typologies of money laundering describes the flow of money in a diverse technique planned by criminals to cover up and disguise their illegitimate money (T. H. Phyu & S. Uttama, 2023). The AMLSim dataset produced by International Business Machines Corporations (IBM) presented eight Anti-Money Laundering typologies which are scatter-gather, gather-scatter, fan-in, fan-out, bipartite, mutual, forward, and cycle (B. Oztas et al., 2023). Five popular examples of typologies are briefly explained below and depicted in the Figure 1-5.

Fan-In and Fan-Out Typology also known as Star Pattern usually happens at the placement and integration phase (Cheng et al., 2023). Figure 1 is the Fan-In typology in which the money is aggregated from different account sources into one central target accounts. Meanwhile, Figure 2 is the Fan-Out typology in which the money from central source is transferred to multiple target accounts (T. H. Phyu & S. Uttama, 2023).

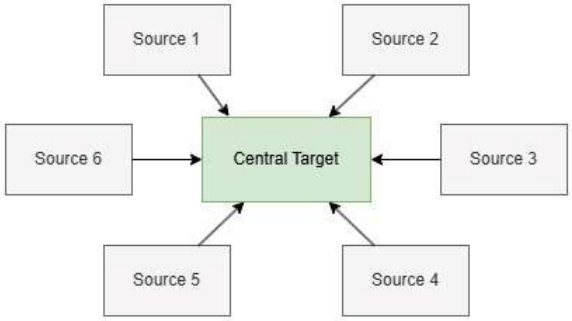

Figure 1: Fan-In Typology

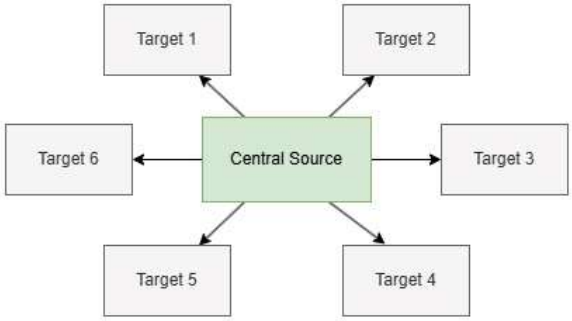

Figure 2: Fan-Out Typology

Mutual Typology as in Figure 3 describes the scenario in which two accounts transferring money to and back with each other. It usually happens at the layering phase. In example, at first, Account A transfer money to Account B, while Account B transfer money to Account C. Then, after a while, the money reverse back from Account C to Account B, and back to Account A (T. H. Phyu & S. Uttama, 2023).

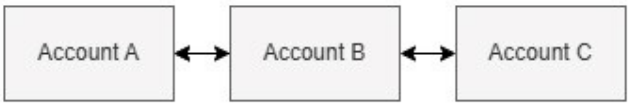

Figure 3: Mutual Typology

Figure 4 shows Forward Typology (T. H. Phyu & S. Uttama, 2023) or also known as Chained Pattern. It is most likely to be occurred at layering phase where the funds move in a sequential order through a chain of accounts that act as a bridge. This pattern could consist of multiple Star Patterns where the money transfer from source account to target accounts through various bridge nodes, forming numerous chains (Cheng et al., 2023).

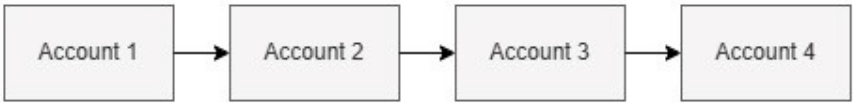

Figure 4: Forward Typology

Lastly, Cycle Typology or Cyclic Pattern is a combination of two or more Chained Pattern (Cheng et al., 2023). The money transfers through a sequence of transactions, comprising numbers of accounts, and ultimately return back to source account which is Account A in Figure 5 (T. H. Phyu & S. Uttama, 2023).

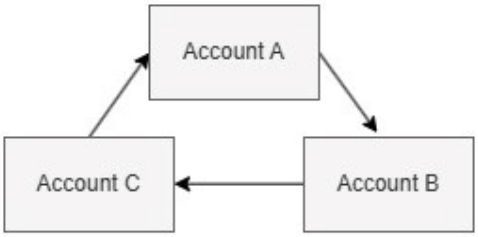

Figure 5: Cycle Typology

2.4 Example of Money Laundering Cases

There are three typical cases related to money laundering. Firstly, the money originated from criminal activities such as drug trafficking in which the criminals must clean the money first before they can use it legally. Secondly, the businessman or company makes false declarations on their legitimate income to avoid paying higher taxes. And the third case involves the recipients of the money that do not want to disclose the source of money due to security or ethical reasons such as receiving rewards from anonymous donor for exposing corruption (Oad et al., 2021).

The infamous case of money laundering in Malaysia is the 1 Malaysia Development Berhad (1MDB) scandal where it involved embezzlement and bribery of funds amounted to billions US dollars in which the money are mostly laundered outside of Malaysia (Jones, 2020).

2.5 Rule-Based Method in Money Laundering Detection

Rule-based method also known as an expert system (Labanca et al., 2022) detects illicit transactions by using heuristic algorithms (Ouyang et al., 2024) that relies on the expert knowledge in which the suspicious transactions are flagged based on predefined rules and thresholds (X. Luo et al., 2022). The rules set the boundaries for transactional actions that may be a part of money laundering process. If the rules match, the investigator will receive alerts, and case will be created for the customer (Oad et al., 2021). For example, the system will raise an alert if the transactions made is above $10,000 and the money is coming from or transferred to an overseas account (Kute et al., 2021). This approach ease the compliance officers to interpret the system's output and use the information straightforwardly (Labanca et al., 2022).

However, this approach has some disadvantages as it is unable to detect new emerging money laundering patterns. Therefore, the rules need to be regularly updated so that they manage to capture evolving changes of this financial crime (Kute et al., 2021). Furthermore, since this approach does not have the capacity to cover unknown anomalous transactions, hence it leads to false negatives ignoring the occurrence of suspicious transactions (Labanca et al., 2022). Moreover, using static thresholds generates a high number of false positives resulting to more alerts and subsequently, more efforts for manual investigations which require more times for the compliance officers to scan through the cases (Oad et al., 2021).

2.6 Machine Learning Algorithm in Money Laundering Detection

Machine learning is a section in Artificial Intelligence (AI) which employs the algorithm techniques that facilitates predictions based on large volumes of data (Prisznyák, 2022). It has the capabilities to assess the hidden correlations and extract the insights by learning the patterns existed in the dataset (Labanca et al., 2022). Moreover, it can overcome the disadvantages of rule-based method by reducing the time to review the alerts manually and reducing false positives (Labanca et al., 2022). Therefore, numerous research are done in the last decades to identify the best machine learning methods for money laundering detection (X. Luo et al., 2022). However, there is no real ideal algorithm to detect money laundering as the performance of machine learning algorithms varies depending on the underlying theoretical logic, (Prisznyák, 2022) adaptability, suitability and generalization capabilities to the dataset experimented (Cheng et al., 2023).

Supervised machine learning algorithm learns the patterns of normal and suspicious transactions using training dataset that has been labelled by subject matter experts or based on past confirmed money laundering's cases (Ouyang et al., 2024). This algorithm often resulting in high detection rate compared to unsupervised machine learning algorithm. However, it is not as effective to detect new suspicious patterns that is not exist in the dataset (Labanca et al., 2022). Example of supervised machine learning are classification algorithms such as Decision Tree, Random Forest, Support Vector Machines and, regression algorithms such as linear regressions (Prisznyák, 2022).

Meanwhile, unsupervised machine learning algorithm is not labelled in advanced, hence it uses the patterns and correlations recognised from the dataset to measures the deviations of the transactions from the norm to label the anomalous transactions (Prisznyák, 2022). While this algorithm has capability to discover unknown anomalies and new patterns that is hidden in the dataset, the subject matter expert still needs to validate whether the predictions are correct. This is because, unsupervised machine learning tend to generate high number of false positives when actually some anomalous transactions are acceptable as normal transactions (Labanca et al., 2022). Some examples of unsupervised machine learning are anomaly detection such as nearest neighbour methods and, clustering model such k-means in the process to find behavioural patterns that differs significantly from licit cases (Ouyang et al., 2024).

2.7 Applications of Machine Learning Algorithms used in Money Laundering Detection

Reference	Machine Learning Algorithm	Dataset	Performance Metrics	Result
(X. Luo et al., 2022)	Dynamic Transaction Pattern Aggregation Neural Network (DTPAN) - utilize two feature extractors: 1) DBFE- Dynamic Behaviour Feature Extractor (to learn the dynamic features of transaction behaviours) 2) DSFE- Dynamic Structure Feature Extractor (to learn the evolution of transfer relationship between accounts)	Real-world dataset provided by law enforcement agency	Precision, Recall, F1, Accuracy	DTPAN enhances the performance of Machine Learning by exploring the dynamic information of transactions
(Cheng et al., 2023)	Group-Aware Graph Neural Network (GAGNN)	Real-world dataset from one of the largest bank card alliances worldwide (UnionPay)	AUC (Area under ROC Curve) and R@P_N (Recall rate when precision rate equals N)	GAGNN can be applied widely to detect organized behavior

(Labanca et al., 2022)	Amaretto - Active Learning Framework -combines supervised (Random Forest) and unsupervised learning techniques (Isolation Forest)	Synthetic dataset provided by industrial partner (trading in international capital market)	AUROC (Area under the receiver operating characteristics), TPR (True Positive Rate), FPR (False Positive Rate), FNR (False Negative Rate), AUC (Area under ROC Curve), Accuracy, Precision, FScore, MCC (Matthews Correlation Coefficient), Norm. Cost	Amaretto improves up to 50% detection and reduces the overall computing cost by 20%
(Yang et al., 2023)	Combining two methods: 1) combines heuristic rules, Long Short Term Memory (LSTM) and Graph Convolutional Neural Network (GCN) 2) ensemble learning for anomaly detection, to identify anomaly transaction data that may be missed by heuristic rules	Elliptic dataset 2019, a bitcoin transaction dataset	Precision, Recall Rate, F1, AUC (Area under ROC Curve)	Accurate identification of anomaly transactions and low false-negative rate in identifying abnormal data

(J. Luo et al., 2024)	Edge-Node Fusion algorithm for Transaction-Level prediction (ENFT) -based on principal neighbourhood aggregation -includes multi-task edge prediction method (MEP) and conditional edge prediction method (CEP)	Synthetic dataset generated by AMLSim Simulator	Accuracy, precision, recall, and F1 Score	ENFT model with two-round training method enhance prediction on illicit transaction edges.
(Tundis et al., 2021)	Comparing five supervised machine learning: 1) Decision Trees (DT) 2) Support Vector Machines (SVM) 3) Random Forest (RF) 4) Linear Regressions (LRs) 5) Naïve Bayes (NB)	Synthetic open-source financial transaction dataset that resembles the normal transactions and malicious behaviour related to money laundering	Accuracy, precision, recall, F1 Score, TPR (True Positive Rate), TNR (True Negative Rate), FPR (False Positive Rate), FNR (False Negative Rate)	Random Forest has the best performance with an accuracy, recall and F1 Score greater than 94% and lower False Positive Rate (FPR)
(Reite et al., 2024)	XGBoost	Data from bank in Norway on Small and Medium-sized Enterprises	Mean AUC, AUC std, Youden's J, Min Euclidean distance, Youden's TPR/FPR, Min distance TPR/FPR	Client risk classification model with additional accounting data and credit score information

		(SMEs) customers to examine how various client risk classification models can predict suspicious transactions		can predict suspicious transaction accurately and reducing number of false positives.
(Pambudi et al., 2019)	Support Vector Machine (SVMs) with Random Under Sampling (RUS) techniques to handle imbalance dataset and reduce model training time	Synthetic Financial Dataset for Fraud Detection	Precision, recall, F1 Score, TPR (True Positive Rate), TNR (True Negative Rate), FPR (False Positive Rate), FNR (False Negative Rate)	The model can detect fraud more accurate with an increase in precision by 40.82% and F1-Score of 22.79% compared to previous study
(Zhang & Trubey, 2019)	1) Bayes logistic regression 2) Decision Tree 3) Random Forest 4) Support Vector Machine 5) Artificial Neural Network	Actual transaction data from US financial institution	AUC (Area under the ROC Curve)	ANN has the best performance compared to other four algorithms.

References

B. Oztas, D. Cetinkaya, F. Adedoyin, M. Budka, H. Dogan, & G. Aksu. (2023). Enhancing Anti-Money Laundering: Development of a Synthetic Transaction Monitoring Dataset. *2023 IEEE International Conference on E-Business Engineering (ICEBE)*, 47–54. https://doi.org/10.1109/ICEBE59045.2023.00028

Buehler, K. (2019). Transforming approaches to aml and financial crime. *McKinsey*, *Query date: 2024-12-14 02:04:51*.

Cheng, D., Ye, Y., Xiang, S., Ma, Z., Zhang, Y., & Jiang, C. (2023). Anti-Money Laundering by Group-Aware Deep Graph Learning. *IEEE Transactions on Knowledge and Data Engineering*, *35*(Query date: 2024-12-12 23:09:33), 12444–12457. https://doi.org/10.1109/TKDE.2023.3272396

Jones, D. S. (2020). 1MDB corruption scandal in Malaysia: A study of failings in control and accountability. *Public Administration and Policy*, *23*(1), 59–72. https://doi.org/10.1108/PAP-11-2019-0032

Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep Learning and Explainable Artificial Intelligence Techniques Applied for Detecting Money Laundering–A Critical Review. *IEEE Access*, *9*(Query date: 2024-12-12 23:09:33), 82300–82317. https://doi.org/10.1109/ACCESS.2021.3086230

Labanca, D., Primerano, L., Markland-Montgomery, M., Polino, M., Carminati, M., & Zanero, S. (2022). Amaretto: An Active Learning Framework for Money Laundering Detection. *IEEE Access*, *10*(Query date: 2024-12-12 23:09:33), 41720–41739. https://doi.org/10.1109/ACCESS.2022.3167699

Luo, J., Pan, W., & IEEE. (2024). Edge and Node Fusion for Transaction-level Prediction on Money Laundering Detection. *Beijing University of Posts & Telecommunications*, 138–143. https://doi.org/10.1109/ICAIBD62003.2024.10604451

Luo, X., Han, X., Zuo, W., Xu, Z., Wang, Z., Wu, X., & IEEE. (2022). A Dynamic Transaction Pattern Aggregation Neural Network for Money Laundering Detection. *Qilu University of Technology*, 818–826. https://doi.org/10.1109/TrustCom56396.2022.00114

Oad, A., Razaque, A., Tolemyssov, A., Alotaibi, M., Alotaibi, B., & Zhao, C. (2021). Blockchain-Enabled Transaction Scanning Method for Money Laundering Detection. *ELECTRONICS*, *10*(15). https://doi.org/10.3390/electronics10151766

Ouyang, S., Bai, Q., Feng, H., & Hu, B. (2024). Bitcoin Money Laundering Detection via Subgraph Contrastive Learning. *ENTROPY*, *26*(3). https://doi.org/10.3390/e26030211

Pambudi, B., Hidayah, I., & Fauziati, S. (2019). Improving Money Laundering Detection Using Optimized Support Vector Machine. *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, *Query date: 2024-12-12 23:09:33*, 273–278. https://doi.org/10.1109/ISRITI48646.2019.9034655

Prisznyák, A. (2022). Bankrobotics: Artificial Intelligence and Machine Learning Powered Banking Risk Management Prevention of Money Laundering and Terrorist Financing. *PUBLIC FINANCE QUARTERLY-HUNGARY*, *67*(2), 288–303. https://doi.org/10.35551/PFQ_2022_2_8

Reite, E., Karlsen, J., & Westgaard, E. (2024). Improving client risk classification with machine learning to increase anti-money laundering detection efficiency. *JOURNAL OF MONEY LAUNDERING CONTROL*. https://doi.org/10.1108/JMLC-03-2024-0040

T. H. Phyu & S. Uttama. (2023). Improving Classification Performance of Money Laundering Transactions Using Typological Features. *2023 7th International Conference on Information Technology (InCIT)*, 520–525. https://doi.org/10.1109/InCIT60207.2023.10413155

Tundis, A., Nemalikanti, S., Mülhäuser, M., & ASSOC COMP MACHINERY. (2021). Fighting organized crime by automatically detecting money laundering-related financial

transactions. *Technical University of Darmstadt*. ARES 2021: 16TH INTERNATIONAL CONFERENCE ON AVAILABILITY, RELIABILITY AND SECURITY. https://doi.org/10.1145/3465481.3469196

Yang, G., Liu, X., & Li, B. (2023). Anti-money laundering supervision by intelligent algorithm. *COMPUTERS & SECURITY*, *132*. https://doi.org/10.1016/j.cose.2023.103344

Zhang, Y., & Trubey, P. (2019). Machine Learning and Sampling Scheme: An Empirical Study of Money Laundering Detection. *COMPUTATIONAL ECONOMICS*, *54*(3), 1043–1063. https://doi.org/10.1007/s10614-018-9864-z