AI against money laundering networks: the Colombian case

Olmer Garcia-Bedoya
Department of Engineering, Universidad de Bogota Jorge Tadeo Lozano, Bogota, Colombia

Oscar Granados
Department of Economics, Universidad de Bogota Jorge Tadeo Lozano, Bogota, Colombia, and

José Cardozo Burgos
Todosistemas STI, Bogota, Colombia

Abstract

Purpose – The purpose of this paper is to examine the artificial intelligence (AI) methodologies to fight against money laundering crimes in Colombia.

Design/methodology/approach – This paper examines Colombian money laundering situations with some methodologies of network science to apply AI tools.

Findings – This paper identifies the suspicious operations with AI methodologies, which are not common by number, quantity or characteristics within the economic or financial system and normal practices of companies or industries.

Research limitations/implications – Access to financial institutions' data was the most difficult element for research because affect the implementation of a set of different algorithms and network science methodologies.

Practical implications – This paper tries to reduce the social and economic implications from money laundering (ML) that result from illegal activities and different crimes against inhabitants, governments, public resources and financial systems.

Social implications – This paper proposes a software architecture methodology to fight against ML and financial crime networks in Colombia which are common in different countries. These methodologies complement legal structure and regulatory framework.

Originality/value – The contribution of this paper is how within the flow already regulated by financial institutions to manage the ML risk, AI can be used to minimize and identify this kind of risk. For this reason, the authors propose to use the graph analysis methodology for monitoring and identifying the behavior of different ML patterns with machine learning techniques and network science methodologies. These methodologies complement legal structure and regulatory framework.

Keywords Money laundering, Artificial intelligence, Software architecture, Financial crime networks

Paper type Research paper

Introduction

Criminal activities are increasingly diverse and have been carried out in different physical and digital spaces. These activities involve a growing group of people, organizations and

The authors would like to thank anti-money laundering expert members in different financial organizations. This project received financial support of Minciencias from call for research proposals 816–2018, Project Grant 696481765389.

Journal of Money Laundering Control

1368-5201
DOI 10.1108/JMLC-04-2020-0033

channels, becoming a social problem that affects different communities around the world. Money laundering (ML) is one of those activities, but it is generally associated exclusively with drug trafficking. However, ML relates to other illegal activities such as human trafficking, arms trafficking and smuggling. In the past twenty years, organizations such as the United Nations and some research centers have carried out several studies that agree that illegal activities, mainly drug trafficking and organized crime, represent approximately 3.5% of world GDP and 75% of that value used different mechanisms of ML. Besides, there are other crimes against public resources such as corruption and tax evasion that are not included in traditional crimes, but that use similar mechanisms to launder the money that results from illegal activities with different patterns and environments.

This problem attracts the attention of academics from a wide range of disciplines and different perspectives. First, it has been analyzed from the general schemes of the economy as Davoodi and Tanzi (1997), Walker (1999), Masciandaro *et al.* (2007), Schneider and Windischbauer (2008), Walker and Unger (2009) and Schneider (2010). Others like Imanpour *et al.* (2019) review the microeconomics and game theory to define the best public policies against ML based on the incentives that criminals have to develop such activities. Ardizzi *et al.* (2018) implements an econometric analysis of the use of cash in Italy to distinguish the illegal component in legal transactions with cash, whereas Loayza *et al.* (2019) review the impact of illegal activities and ML on economic growth for that matter from Colombia. McCarthy *et al.* (2015) propose a microeconomic model to analyze how the offender and the ML professional interact, that is, separate the ML activity of the criminal organization and integrate it with the professional activities developed by bankers, accountants and lawyers.

In a second group are the proposals from complex systems, systems theory and network science. To list a few, for example Luna-Pla and Nicolás-Carlock (2020) use an empirical approach to model corruption using some methodologies of complex systems and network science. Another analysis is Ribeiro *et al.* (2018), who use the dynamic structure of corruption networks as a proxy of illegal activities or Dreżewski *et al.* (2015), who applied some algorithms of social network analysis to detect the ML activities. Demetis (2018) uses the concept of structural coupling of systems theory to represent the dynamic relationship between the computer profile and the social profile in ML behavior.

In a third group, the social sciences approximation explains from sociology to law different aspects of ML. Some examples of a large list are as follows. Verhage (2017) tried to explain the anti-ML policy and its real impact on ML. The influence of criminal organizations in some social groups (Toner, 2009) or the digital transformation of financial services that created an ecosystem to financial crime as illegal financial activities, financial fraud or ML (Picard, 2009).

However, no proposal exists that tries to resolve this problem with a multidisciplinary approach. This paper presents the design of an architecture of a comprehensive system of administration and risk management to protect entities from potential acts of ML or financing of terrorism (LA/FT). The proposed system is a digital transformation tool based on artificial intelligence (AI) techniques, which intends to contribute to the integral construction of an anti-ML (AML) compliance management system, based on international standards such as ISO 31000, ISO 31010 and the recommendations of the Groupe d'action financière (FATF-GAFI.ORG). The contribution of this paper is how within the flow already regulated by financial institutions to manage the ML risk, AI can be used to minimize and identify this kind of risk. For this reason, we propose to use the graph analysis methodology for monitoring and identify the behavior of different ML patterns with machine learning techniques and network science methodologies.

After this introduction, the paper is divided as follows. In the next section, we present the first subsection of AML context in Colombia and the second one as a basic network science terminology that we used in the text. The following section is devoted to a brief description of the software solution architecture, especially the modules against ML. In the next subsection, we developed an AI structure as a tool to robustness the software architecture. In the following section, we present some simple networks visualization examples for the behavior of ML from an elementary set of rules associated with these activities. In the final section, we close the paper with a section devoted to highlight conclusion aspects of our approach and provide directions for future work.

Preliminaries

Both ML and terrorist financing represent a significant risk for countries, markets, and, in general, for society, because of the variety and impact of crimes. Therefore, at the local, regional and global levels, plans are being executed to combat this scourge through government policies, international cooperation, and coordination between countries. The countries' policies and laws incorporate regional or international agreements to combat crime, and additionally, each country creates increasingly robust structures through constitutional, regulatory and legal changes. For the reader convenience, we include here some elements of AML process in Colombia and a list of basic notions related to networks that are used in this work.

Colombia and anti-money laundering legal approach

In Colombia, crimes cataloged in the Criminal Code and incorporated into financial legislation through the "Instructions relating to the administration of the risk of ML and the financing of terrorism" contained in the Basic Legal Circular Part I, Title IV, Chapter IV of the Financial Superintendency. However, the offenses of ML are easily identifiable by the nature that any non-legal activity that tries to give the goods derived from such activities looks like legality or legalize, hide or cover up their origin automatically become ML. These crimes could be executed individually or collectively under the definition of a concert to commit crimes. The illegal activities (source crimes of ML) can be grouped in six types as established in the Table 1.

Money for this crimes must be reported not only by the officials of the institutions involved but also by those who have surveillance and mandatory reporting duties such as the superintendencies of ports and transport, notarial and registration, partnerships, private financial surveillance, health, tax address, the direction of gambling and direction of sports activities.

To regulate these reports, Colombia has created risk management system for money laundering and terrorism financing (ML and terrorist financing risk management system). This system seeks that all entities present suspicious transactions report on which operations must be listed by their number, quantity or characteristics is atypical in a specific segment. These anomalies should be determined around standard practices of a particular business or industry according to the uses and customs of the activity in question that has not been reasonably justified.

In this regard, the companies present difficulties in interpreting how to carry out the risk management system, from consolidating the risks to their monitoring and tracking. In the area of segmentation, the problem lies in the definition, application and, especially in sustaining a methodology to segment risk factors. In the case of the customer information, this is normally found in a wide variety of sources of information, resulting in a troublesome operative situation to generate the proper profiling of the customer, by not having the data

Group	Crime
Crime against people	Migrant traffic
	Human-trafficking
	Kidnapping and extortion
	Gender and age trafficking
Crime against the financial system	Securities fraud
	Misuse of client funds and resources
	Massive and regular money collection
	Unauthorized transactions with shareholders and associates
Crime against public administration	Misuse of information
	Transnational bribery
	Tax fraud and evasion
	Traffic of influence and abuse of authority
	Criminal association against public administration
	Assets omission or non-existent liabilities inclusion
	Peculation, fraud, omission, bribery and concussion
Customs fraud	Smuggling
	Customs fraud
	Smuggling facilitation
	Facilitation of oil and derivatives smuggling
Traffic	Drug trafficking
	Arms trafficking
Crime against governments	Rebellion
	Terrorism financing
	Management of terrorist activities resources

Table 1.
Source-crimes for money laundering

in a timely manner. Likewise, nowadays monitoring the customers and their transactions with traditional techniques implies a very high operating load because of scattered sources of information.

For any company and organization, it is extremely important to implement a technology solution of self-management of ML and terrorist financing risks. This paper presents below an idea of the architecture and modules for this solution, taking into account Colombian regulations and the financial action task force recommendations (Muñoz Forero and Bautista Bernal, 2018).

Money laundering networks

The interactions that generated in the ML source crimes can be exemplified as the number of agents N that represents the number of network components, i.e. their size and are identified as $i = 1,2,\ldots,N$. Additionally, the number of links L represents the total number of interactions between agents. These interactions become a property, which defined as the degree k_i of an agent, which represents the number of links with other agents, i.e. the number of associated connections. Thus, the degree of an agent in a non-directed network defined by the total number of links L and can be expressed as the sum of the agent's degrees:

$$L = \frac{1}{2}\sum_{i=1}^{N} k_i$$

With the agents and their links, it is possible to build a network G with N agents and Links consisting of a group of intersection points $V(G) = v_1,v_2,\ldots,v_n$ and a group of intersections $E(G) = e_1,e_2,\ldots,e_m$. This network is represented by an adjacency matrix, A_{ij} which indicates the links between the nodes i and j. Where $A_{ij} = 1$, if the link is effective otherwise it will be $A_{ij} = 0$. Then, the agents are linked and their links list through a proximity matrix given by:

$$A_{ij} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

The importance of each agent depends on the proximity degree of its neighbors because the importance of the node i determined by x_i and the centrality of the agent's vectors i is proportional to the sum of the centrality of the eigenvectors of all the agents that connect to i. Thus:

$$x_i = \frac{1}{\lambda}\sum_{j\in N_i} x_j = \frac{1}{\lambda}\sum_{j=1}^{n} A_{ij}x_j,$$

where N_i is the agent group connected to the agent i, n is the total number of agents, A_{ij} is the adjacency matrix and λ is a constant. What is the probability that a randomly selected agent in the network has a degree k? For this, the degree distribution p_k provides the probability, then:

$$\sum_{k=1}^{\infty} p_k = 1,$$

where p_k in a network of N agents is given by $\frac{N_k}{N}$. The degree distribution is important to understand some properties and phenomena of networks, especially, their robustness and propagation (Barabási and Pósfai, 2016). In this case, highly connected agents (hubs) can be identified through their degree and establish how important the other agents are from the degree distribution.

However, in the ML, there are not only simple connections between two agents but also develop interaction networks to try to hide the trajectory of the crime.

In this way, the distance d_{ij} between nodes i, j is the smallest number of links to go from i to j. Therefore, the neighbors of the agent i are all agents j that are connected to this agent through a simple link $d_{ij} = 1$. ML can have multiple trajectories of the same length d_{ij} between a couple of agents, but the fundamental element is the direction that the interactions between them have. Because in a non-directed network $d_{ij} = d_{ji}$, i.e. the distance between nodes i y j is the same distance; while in a directed network $d_{ij} \neq d_{ji}$, i.e. the existence of a path from the node i to node j does not guarantee the existence of a trajectory from j to i. This can be written as:

$$d_{ij} = \min\left\{\sum_{k,l\in P_{ij}} a_{kl}\right\},$$

where P_{ij} is the connection path between the agent i and the agent j. From the distance between the agents i, j the average path length established as ζ, that if this distance is

smaller, there may be a greater possibility of interaction between agents involved in ML, although, to reduce suspicious activities the distance may be greater. Finally, the average path length ζ is given by:

$$\zeta = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i \geq j}^{n} d_{ij}$$

These elements that define the interactions between agents allow a basic identification about ML from a network science perspective. However, we include some patterns of static networks., which integrates with the against ML framework to determine the software architecture.

- *Path.* The money launderer agent develops many transactions to clean the illegal money and to send to another money launderer agent. [Figure 1(a)]
- *Cycle.* The money launderer agent does not care how many transactions are needed to clean the illegal money but at some point, this money returns to the money launderer agent [Figure 1(b)].

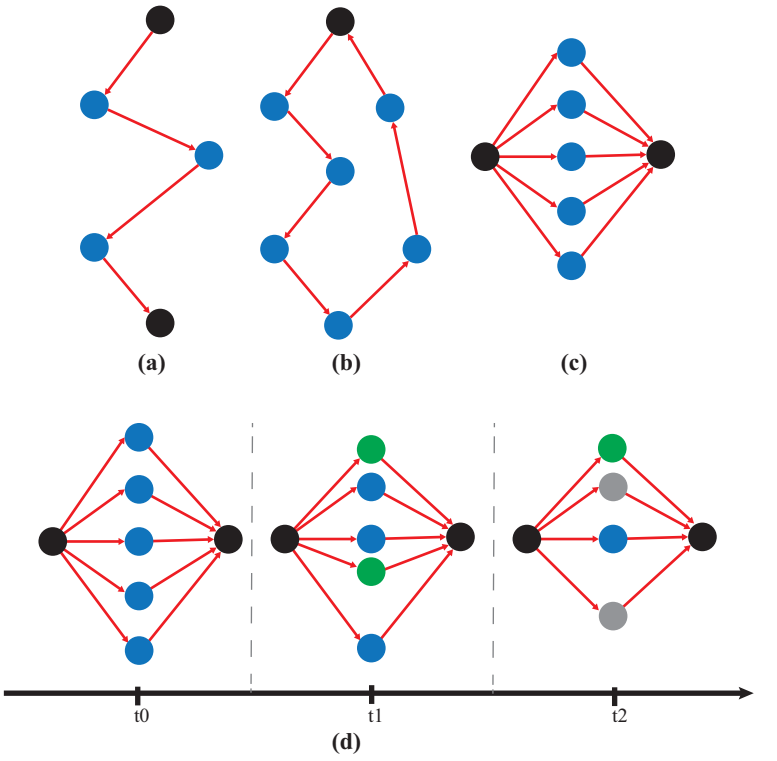


Figure 1.
Some money laundering topologies

- *Smurf*. The money launderer agent uses the breaking up of money into smaller amounts. The money is then deposited into several bank accounts either by different persons or by a single person over a period [Figure 1(c)].

In the digital dimension, money launderers have created a dynamic scheme that affects very quickly the regulations and government policies because they use different accounts or financial instruments that close or liquidate very soon [Figure 1(d)]. For this reason, we need to understand the dynamic structure of ML, and AI can be a complement to network science to identify in real-time the dynamical topology of ML networks. In the next section, we integrate these elements with the AML framework to determine the software architecture.

Software solution architecture

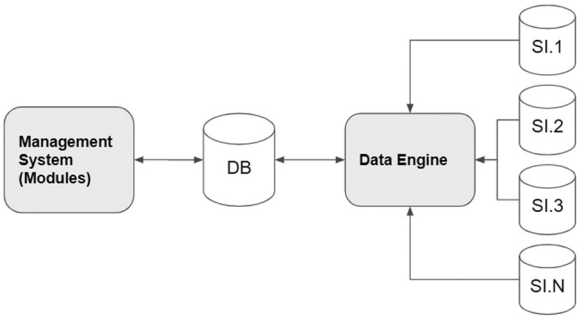
In this section, we present a software solution architecture. We begin with a brief description of the modules that exhibit the principal characteristics of the software architecture and we finish with AI and visualization tools to robust the architecture (Figure 2).

Modules architecture

A solution with a component-based software structure is ideal approach to connect more components and be scalable. This allows to manage modules, heavy processes and other functionalities in separate services. The main features of this configuration are reusability of components, easy replacement, scalability and modularity. Therefore, this modular design supports an event-driven architecture or a service-oriented architecture, including a micro-services architecture (MSA).

The first component is the management system, the different modules the companies need to control and manage their ML/FT risk system that are placed in here. From a pattern design point of view, these modules can be separated not only according to the unique responsibility principle but also according to physical components as services. The information handled through the modules is located in a database structure to access it at any time.

The second component is the Data Engine, in charge of receiving the information and processing it by validation and business rules, and specially through AI techniques, to detect possible events of ML/TF in the customers' operations. The data can be supplied to this engine from different sources of information, which is the



Notes: DB: Database; SI: External source of information

Figure 2. Component architecture

reason why it is important for this component to be able to receive it no matter the type of the source and its configuration. Besides, this information and the results of the processing by AI and rules are located in a common database structure for both, Management System and Data Engine, to use it and to have the entire data available to manage the ML/FT risk system.

These two components work together to implement the modules which support the compliance of an organization for combat the ML. The first modules group is as follows:

- *Risk management and control of AML*. This module, based on ISO 31010, includes the stages of identification, measurement, evaluation, control and monitoring of risk. Risk classified as reputational, legal, operational or contagious and have as risk factor customers, transactional channels and products. Risks are estimated through the probability of occurrence and impact assessment, as shown in Figure 3, which are adapted through statistical tools and with the controls proposed to identify new risks, validate the efficiency of controls or the need to design new controls.
- *Determination of customer risk profile*. This module is based on the transactionality made by customers and is related to the risk factors of products, channels, and jurisdictions. The risk factors, initially defined by the entity, are automatically adapted by the statistics of the transactional monitoring module.
- *Documentation of suspicious transaction reports*. When different internal stakeholders define warning signals, the system creates a suspicious operation report with traceability characteristics. This report lets to be monitored and processed by all responsible. The network of people with whom interactions occur is automatically generated in the report.
- *Politically exposed persons (PEP) clients and associated network management*. This module identifies and manages clients that have the status of being PEPs according to the regulations of the control entities.
- *Campaign and political party management*. A specific segment of users who must comply with additional regulations based on the records of maximum amounts established by the corresponding authority.
- *Attention to legal requirements* This module manages the information requirements by competent authorities, which allows these requirements to serve the risk management in two ways.

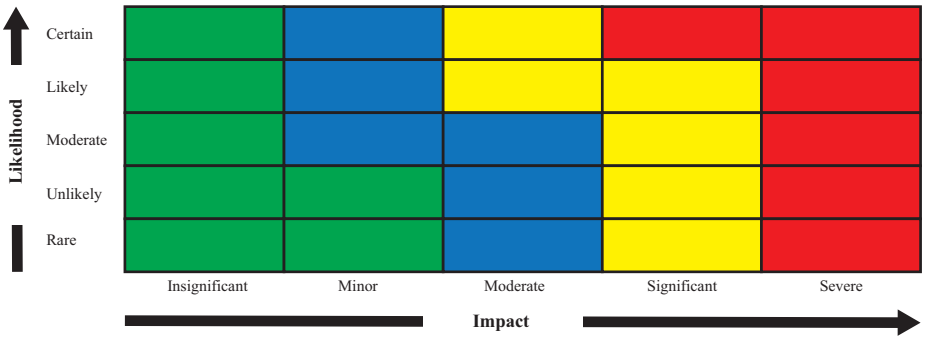


Figure 3.
Risk matrix

- *Management of international agreements* It allows managing clients that have special requirements with other organizations. Among the international agreements currently supported are FACTA (The Foreign Account Tax Compliance Act) from the Americas and Common Reporting Standard of the Organisation for Economic Co-operation and Development.

Each module generated both individual and statistical reports of the activities carried out on it, which allows traceability for the risk management and associated processes. Additionally, the software generates some statistics that support the decision-making based on data (data-driven) for which there is a tool for generating reports and data visualization. The next section explains the second module group characterized by machine learning use in the back to support the AML strategy.

Artificial intelligence

There are three modules that information has characteristics such as volume, variability, and speed that can be cataloged as big data modules. We designed them with an architecture where their core functionalities learn from the data to make AML's task more robust and effective.

Transactional monitoring of customers and users

Financial entities obliged to report the operations carried out by their clients that they determine as unusual (report of suspicious operation), for which this module must allow identifying the clients who perform them and report them to the regulatory entity. This module is composed of the following phases.

(1) *Operations consolidation*: In this phase, the entity provides information from the different sources where operations carried out by clients. This information must be provided within a structure defined by the entity according to its client operations.

(2) *Definition and rules parameterization*. Rules are conditions that define the entity to identify what could be considered unusual, according to the risk profile of the entity, its niche market (customers), and products offered. This functionality created alert conditions. As an example, Algorithm 1 presents a transaction rule created by the entity.

- Generate alarm.
- The system has the functionality to generate its own rules through Machine learning techniques on the clients' operations data of the to establish the transactional behavior. The procedure of how these models generated is performed according to the cross industry standard process for data mining methodology (Wirth and Hipp, 2000), which contemplates the general stages for the construction and development of any model. The first stage is business understanding, which requires that the person that configure the software has the expertise over AML. The second and third stage is data understanding, where a descriptive analysis is carried out to analyze and detail the behavior and obtaining the critical information, cleaning, and transformation methods that are required to be modeled. These stages are a service of the company that develop the software depending on the company data.
- After that, there come two stages that are modeling and evaluation, which programmed in the software to take the cleaned data and test different models and select the best. In the case of transaction monitoring, models tested are

logistic regression, decision tree, random forest, Xgboost, and AutoEncoder Deep Neural Network. All these models evaluated over validation data to obtain the better model which gives the best indicator of ML detection. It is essential to say that in this module, the accuracy is not a good indicator because there exist many data which is not fraud, so many evaluation indicators are proposed to the organization.

- The last part is implementation, where we can say that the software use R in the background to data processing and model creation. In this part, we also identified that these models have to be updated periodically to obtain a real digital transformation in the AML process and close the methodology loop.

(3) *Generation, assignment and management of alerts*: If a client meets one or more conditions, the software generates automatic alarms according to the periodicity with which the client information or transactions is updated. The software also allows generating a historical report of alert signal management, which evidences all the steps that were carried out on a particular case.

Customer segmentation

Financial institutions obliged to segment the risk factors guaranteeing homogeneity within segments and heterogeneity between them, according to the methodology previously established by the entity. At a minimum, the following factors must be segmented accordingly:

- *Clients*: economic activity, volume or frequency of their transactions and amount of income, expenses and equity.
- *Products*: type of product, characteristics and final users.
- *Distribution channels*: type of channels and characteristics.
- *Jurisdictions*: location, characteristics and type of transactions.

The software performs this process to allow us to visualize the typical characteristics of the transactions that developed and compare them with those made by clients to detect unusual operations. Here, other sets of machine learning algorithms implemented through the variables required by regulations as well as others defined by the financial entity. These models obtained through the same methodology explained in the section above. The software also allows the creation of manual segments, which serve to feedback and evolve AI models.

The purpose of this segmentation is to understand how the entity's client universe distributed, it identifies which segments are the most risky and which clients of each segment have atypical transaction behavior concerning the other clients that are part of the group (warning signs), and be able to establish unusual operations and the respective report to the financial information and analysis unit.

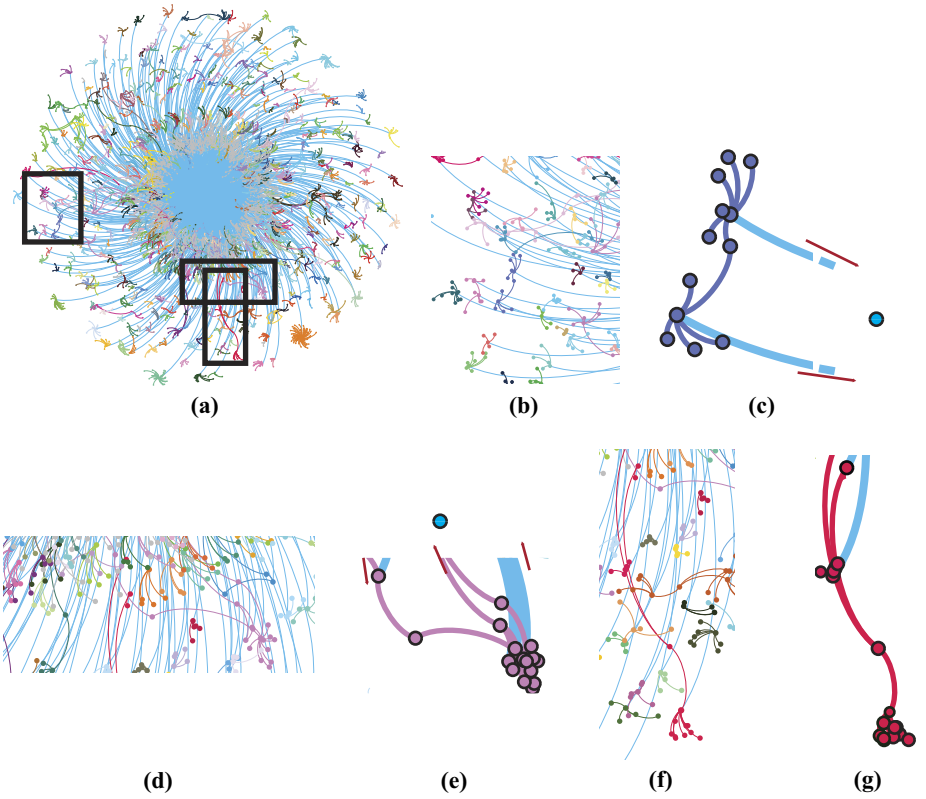
Graph visualization

To understand how complex relations between customers are, the implementation of graph theory supported by network science is important. The information is located in a graph database to show the results through the graph visualization interface. This implementation is part of the source of information on the Data Engine. Some networks are presented below.

Visualization of money laundering networks

With the approach of finding ML groups and patterns using network science tools introduced above, in this section, we present the most frequent topologies in ML networks that confirmed the simple network framework introduced in a section above. The usefulness of these visualizations lies in catching other suspicious networks because these interactions are similar to patterns of ML. To demonstrate the validity of our approach, in the following we perform some numerical simulations with real data and we develop the ML visualization in static networks. As a first application, we include the static network cases in a dataset with approximately 6,000 nodes.

Figure 4 shows the emergence of ML suspicious groups in the real data (Panama Papers dataset) based on the records of the International Consortium of Investigative Journalists, but we have held all information regarding names of companies and people in anonymity. This data set contains information about 0.6 million agents (intermediaries, offshore entities, and officers) and almost 0.8 million links among



Notes: (a) One of the principal components of interaction between financial and non-financial agents in a real data set; (b) first network zoom; (c) cycle structure from the first zoom; (d) second network zoom; (e) smurf structure from second zoom; (f) third network zoom; (g) path structure from the third zoom

Figure 4.
Zoom of suspicious interactions

these agents. We compiled and curated the data, but especially, we used one of the three principal components of this network [Figure 4(a)]. This component has 5,751 agents and 6,039 edges (interactions). Additionally, as network metrics we considered the average degree of this component, as a part of the whole Panama papers network (399,145 agents and 446,387 edges), is 1.05 and the average path length between agents is equal to 1, i.e. the average distance between pairs of nodes in the network. In this case, we find that these metrics confirmed some of the traditional ML characteristics.

On the other hand, the modularity structure of this component presents approximately 863 communities but the principal element of this analysis is the ML structures that we find as paths, cycles, and smurf schemes. We presented some examples of these topologies that help us to identify suspicious activities. The first one is the cycle [Figure 4(c)] that connected two groups with the same node. Second, the smurf scheme that stars with interactions between principal-agent and few agents. After that, these few agents with a great number of agents. In a final stage, those agents connected with a few agents and these agents with principal-agent [Figure 4(e)]. In some situations, agents in the smurf scheme could connect directly with principal-agent.

The last topology structure is the path [Figure 4(g)]. In this case, the network shows that the path between a suspicious agent and others did not close a cycle, but only disguise the transitions from one place to another.

Finally, the temporal decomposition of ML networks is part of our AI proposal because the money launderer agents change their operations day by day. Thus, is very important the current data and the integration with data in different temporalities to check the topology evolution and changes in some patterns.

Discussion and conclusion

In this paper, we presented technical and empirical evidence of the AML framework with new methodologies as AI and network science. The software proposal looks to fill a gap between ML dynamics and current tools of AML because some of these tools as a result of a static analysis after several days or months of have been made the financial transactions.

Additionally, this paper presents some findings in three issues that can serve as guidelines for further research on AI and network science to detect anomalies in a growing financial transaction ecosystem. First, the software architecture presented in this paper is fundamental to reduce some patterns of ML and terrorism finance because it integrates in only one architecture traditional elements of the AML proceedings with AI and network science. Second, the AML activities need a multidisciplinary structure to solve the global problem of this kind of financial crime. Third, we need to develop a framework to connect different temporalities of financial networks, i.e. by the minute, by hour or by day.

On the other hand, we found some limitations. The network visualization module employed in the software architecture described in this paper suffers from several limitations. First, the module needs to implement other network science methodologies to strengthen the AML proceedings and visualization tools. Second, we need to train the AI algorithms with new datasets and a great group of simulations to consolidate the software and confront the ML dynamics and evolution.

Finally, ML complexity needs a multidisciplinary framework to reduce or mitigate it because the financial crimes related to ML increasing day by day. The AML schemes need a new paradigm: a multidisciplinary framework that provides a dynamic structure from which one might understand a great number of phenomena and patterns from financial

crime and, for this reason, the AI, and network science combination is one of the new options to resolve complex problems and illegal activities. The solutions have and need to implement a digital transformation to close the growing gap and complement the regulatory frameworks.

References

Ardizzi, G., De Franceschis, P. and Giammatteo, M. (2018), "Cash payment anomalies and money laundering: an econometric analysis of Italian municipalities", *International Review of Law and Economics*, Vol. 56, pp. 105-121, doi: 10.1016/j.irle.2018.08.001.

Barabási, A.-L. and Pósfai, M. (2016), *Network Science*, Cambridge University Press, Cambridge, available at: http://barabasi.com/networksciencebook/

Davoodi, H. and Tanzi, V. (1997), "Corruption, public investment, and growth", *IMF Working Papers*, Vol. 97 No. 139, doi: 10.5089/9781451929515.001.

Demetis, D.S. (2018), "Fighting money laundering with technology: a case study of Bank X in the UK", *Decision Support Systems*, Vol. 105, pp. 96-107, doi: 10.1016/j.dss.2017.11.005.

Dreżewski, R., Sepielak, J. and Filipkowski, W. (2015), "The application of social network analysis algorithms in a system supporting money laundering detection", *Information Sciences*, Vol. 295, pp. 18-32, doi: 10.1016/j.ins.2014.10.015.

Imanpour, M., Rosenkranz, S., Westbrock, B., Unger, B. and Ferwerda, J. (2019), "A microeconomic foundation for optimal money laundering policies", *International Review of Law and Economics*, Vol. 60, p. 105856, doi: 10.1016/j.irle.2019.105856.

Loayza, N., Villa, E. and Misas, M. (2019), "Illicit activity and money laundering from an economic growth perspective: a model and an application to Colombia", *Journal of Economic Behavior and Organization*, Vol. 159, pp. 442-487, doi: 10.1016/j.jebo.2017.10.002.

Luna-Pla, I. and Nicolás-Carlock, J.R. (2020), "Corruption and complexity: a scientific framework for the analysis of corruption networks", *Applied Network Science*, Vol. 5 No. 1, p. 13, doi: 10.1007/s41109-020-00258-2.

McCarthy, K.J., van Santen, P. and Fiedler, I. (2015), "Modeling the money launderer: microtheoretical arguments on anti-money laundering policy", *International Review of Law and Economics*, Vol. 43, pp. 148-155, doi: 10.1016/j.irle.2014.04.006.

Masciandaro, D., TakÃ¡Ts, E.Ã. and Unger, B. (2007), *Black Finance*, Edward Elgar Publishing, Cheltenham.

Muñoz Forero, L.S. and Bautista Bernal, V.K. (2018), Conocimiento del Lavado de Activos en Colombia: Enfoque a las Recomendaciones del Grupo de Acción Financiera Internacional (GAFI) y normatividad nacional.

Picard, M. (2009), "Financial services in trouble: the electronic dimension", *Journal of Financial Crime*, Vol. 16 No. 2, pp. 180-192, doi: 10.1108/13590790910951858.

Ribeiro, H.V., Alves, L.G.A., Martins, A.F., Lenzi, E.K. and Perc, M. (2018), "The dynamical structure of political corruption networks", *Journal of Complex Networks*, Vol. 6 No. 6, pp. 989-1003, doi: 10.1093/comnet/cny002.

Schneider, F. (2010), "Turnover of organized crime and money laundering: some preliminary empirical findings", *Public Choice*, Vol. 144 Nos 3/4, pp. 473-486, available at: www.jstor.org/stable/40835624

Schneider, F. and Windischbauer, U. (2008), "Money laundering: some facts", *European Journal of Law and Economics*, Vol. 26 No. 3, pp. 387-404, doi: 10.1007/s10657-008-9070-x.

Toner, G.A. (2009), "New ways of thinking about old crimes: prosecuting corruption and organized criminal groups engaged in labor-management racketeering", *Journal of Financial Crime*, Vol. 16 No. 1, pp. 41-59, doi: 10.1108/13590790910924957.

Verhage, A. (2017), "Great expectations but little evidence: policing money laundering", *International Journal of Sociology and Social Policy*, Vol. 37 Nos 7/8, pp. 477-490, doi: 10.1108/IJSSP-06-2016-0076.

Walker, J. (1999), "How big is global money laundering?", *Journal of Money Laundering Control*, Vol. 3 No. 1, pp. 25-37, doi: 10.1108/eb027208.

Walker, J. and Unger, B. (2009), "Measuring global money laundering: the Walker gravity model", *Review of Law and Economics*, Vol. 5, doi: 10.2202/1555-5879.1418.

Wirth, R., (2000), and J. and Hipp, "CRISP-Dm: towards a standard process model for data mining", in *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Springer-Verlag, London, pp. 29-39.

Corresponding author
Olmer Garcia-Bedoya can be contacted at: olmer.garciab@utadeo.edu.co