Article

A Multi-Agent System Based Approach to Fight Financial Fraud: An Application to Money Laundering

Claudio Alexandre [1,*] and João Balsa [1]

[1] Faculdade de Ciências da Universidade de Lisboa (BioISI-MAS), Campo Grande, 1749-016, Lisboa, Portugal

* Author to whom correspondence should be addressed.

Academic Editor: name
Version December 19, 2017 submitted to Appl. Sci.

Abstract: The anti-money laundering (AML) process has failed both in identifying suspicious cases in due time as in assisting the AML analysts in decision making. Starting from a new generic anti-fraud approach, this article presents the main aspects related to the development of a multi-agent system that goes beyond the capture of suspicious transactions, seeking to assist the human expert in the analysis of suspicious behaviour. First, a transactional behavioural profile of clients is obtained in a data mining process. A set of rules, obtained through data mining over a real database, in conjunction with specific rules based on legal aspects and in the expertise of the AML analysts make up the agents' knowledge base. The cases for which the system was unable to suggest a decision are flagged as requiring more detailed analysis. The system analysed 6 months of real transactions and indicated several suspicious profiles, a set of these suspects was investigated by the AML analysts who proved the suspicion of several cases, including some that had not been identified by the systems in execution.

Keywords: multi-agent system; decision support; anti-money laundering; anti-fraud

1. Introduction

In the last decades, money laundering has been increasingly recognized as a significant global problem and was given special attention by almost every government in the world. An evidence that money laundering is a global worry was the prioritization of its combat at the same level of the most relevant global issues [1]. The large amount of money involved in this crime and the social issues involved, justify the prioritisation in anti-money laundering (AML) [2]. The stages and a graphical scheme of a typical money laundering process was shown and explained in [3].

Most of the current rules and recommendations are targeted to transactions involving cash, which are the most common at the beginning of the money laundering process. This type of bank transaction, normally carried out in person, favours monitoring. Unlike subsequent virtual transactions, whose objective is to hinder the monitoring of this money's trajectory.

Most financial institutions already use semi-automated processes, determined by current regulations, for signalling suspicious transactions of money laundering, based on client information bank register, averages, standard deviations and pre-established fixed rules, usually with origin in empirical observations or the human experience of the AML analysts. However, the growing volume of transactions, coupled with the frequent publication of new national and international regulations, have led to inefficiency in this human analysis process.

With the aim of making the AML process more effective, we developed a multi-agent based approach to support decision making in this process. Introducing a version of the methodology of client behavioural profiles, intelligent agents analyse and signal suspicious transactions, and assist the AML analyst in making signalling decisions. The Belief-Desire-Intention (BDI) model, adopted in the development and implementation of the system, easily uses the behavioural profiles found, as well as the implementation of specific rules, both normative based and according to the risk involved. These characteristics, among other benefits, improve the efficiency of the process, facing the increase in the volume of transactions with the future gradual reduction of human intervention.

The experiment we conducted, over one year of real transactions, revealed a set of suspicious profiles considered adequate by the human specialists of the bank that finances this project, since, in addition to the cases already known, new suspects were signalised and validated.

In the following sections, the new anti-fraud strategy is presented (section two), followed by the analysis of related work (section three). The process of creating and building the agents' knowledge bases is presented in section four. The AML multi-agent system design and implementation is described in section five. The article concludes with an analysis of results (section six) and with some conclusions and future work comments (section seven).

2. Standard Anti-Fraud Strategy and a New Approach

The Know Your Customer (KYC) policy, defined by the Basel Committee[1] in [4], is like a best practices guide in the sense that it details the procedures to be followed to prevent fraud. Based on the cited document it is possible to map a generic flow to combat fraud or swindling in economic activity (Figure 1), in the cases in which this type of crime occurs in operations directly or indirectly computerized.

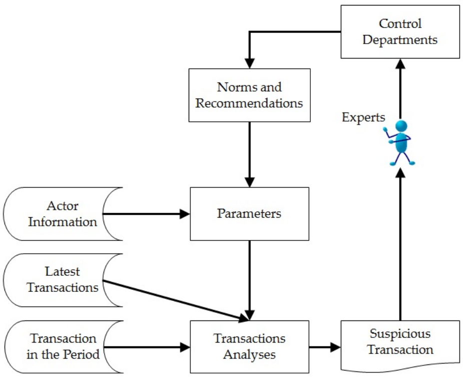


Figure 1. General Flow to Fight Fraud

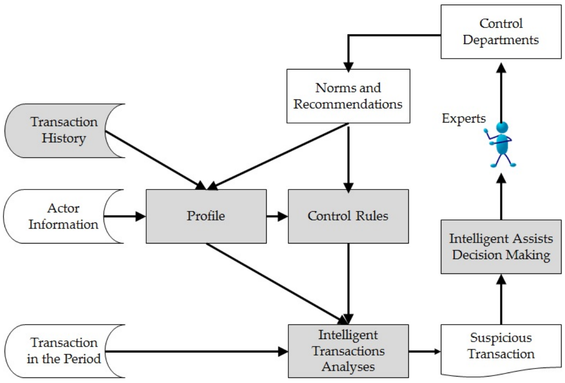

Control Departments
Norms and Recommendations
Transaction History
Experts
Actor Information
Profile
Control Rules
Intelligent Assists Decision Making
Transaction in the Period
Intelligent Transactions Analyses
Suspicious Transaction

Figure 2. New General Flow to Fight Fraud

The generic process of fraud prevention is based on regulations and recommendations issued by control offices and uses parameters that set limits for quantities and values involved in transactions. The analysis of the transactions is carried out in a short period of time and the majority is carried out by a human analyst.

Some authors classify money laundering as a predicate crime, i.e., a crime that always occurs due to some other underlying crime that would illicitly provide its author proceeds that later he, or others, will intend to camouflage [5]. However, it is possible to analyse the money laundering crime as an instance of a generic process of fraud or swindle and and we can see the AML activity as a special case of what happens in a general flow of fraud-fighting in the financial sector.

[1] The BCBS - Basel Committee on Banking Supervision, linked to the BIS - Bank for International Settlements, was established in 1974 to enhance financial stability by improving the quality of banking supervision worldwide, and to serve as a forum for regular cooperation between its member countries on banking supervisory matters to strengthen regulation, supervision and best practices in the financial market.

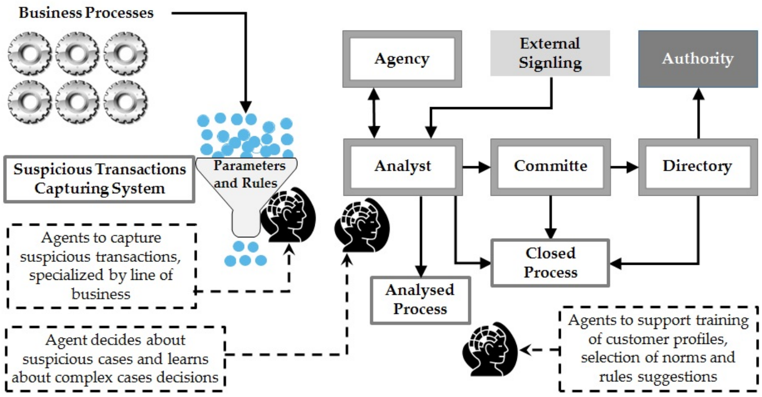

Business Processes
Suspicious Transactions Capturing System
Parameters and Rules
Agents to capture suspicious transactions, specialized by line of business
Agent decides about suspicious cases and learns about complex cases decisions
Agency
External Signling
Authority
Analyst
Committe
Directory
Analysed Process
Closed Process
Agents to support training of customer profiles, selection of norms and rules suggestions

Figure 3. New Anti-Money Laundering Flow

On the other hand, experience has shown that the AML process running as an instance of the generic flow to fight fraud is an inefficient process, because it fails in identifying and reporting suspicious transactions. This inefficiency causes a systemic error because the lack of real cases interferes in the development of new norms and recommendations with good quality.

A new version of the general flow to fight fraud is showed in Figure 2, whose objective is to mitigate the identified risk. The creation of profiles of the actors participating in the activity are based on 1) their complete history of performed operations; 2) the replacement of the current fixed parameters by production rules, based on these profiles and the norms and existing recommendations; and 3) the use of intelligence at some points of the process. This way, the quality level in the capture of suspicious operations and, especially, of decision making by the specialist, is clearly improved.

Creating a new instance of the generic flow proposed in Figure 2 for the ML crime, it is possible to draw a stream like the one shown in the Figure 3, that corresponds to what is commonly used by financial institutions, and presented in [6].

3. Related Work

The adaptability of the "modus operandi" of fraudsters and the lack of systematized information linking suspect transactions to evidence of crime are obstacles to a more rapid advance in automating the process of preventing and combating money laundering activities. However, since the first widely publicized system in the AML area, the FinCEN Artificial Intelligence System (FAIS) [7], developed and used by the Financial Crimes Enforcement Network (FinCEN), many artificial intelligence techniques have been used in the search for a good system.

Data mining, machine learning and clustering techniques have been widely used in the attempt to identify suspected money laundering cases, as in Zhang [8], where a discretization process was applied to a data set to find a more adequate set of clusters. Kingdon [9] proposed that an artificial intelligence approach should model individual clients and look for *unusual* rather than *suspicious* behaviour. There are also statistical approaches, like the ones described in Liu and Zhang [10] and Tang and Yin [11].

In Le-Khac and Kechadi [12], the authors present a case study corresponding to the application of a knowledge base solution that combines data mining techniques, clustering, neural networks and genetic algorithms to detect money laundering patterns. Chang and Chang [13] proposed the use of decision trees based on C4.5 to induce rules and use them to validate the identified cluster. Larik & Haider [14] focus their work on the debit and credit information made by clients of a financial institution to identify suspicious transactions.

In his survey about clustering, Sabau [15] asserts that clustering has proven itself a recurrently applied solution for detecting fraud and concludes that k-means based clustering algorithms with Euclidean distance as dissimilarity metric are the most commons used ones.

Regarding the use of agent based approaches in AML, there are few authors that have considered them. In Gao [16], an agent architecture is defined to include a set of specialized agents, such as data collecting agents, monitoring agents, a behaviour diagnosis agent, and a reporting agent.

Xuan and Pengzhu [17] present an agent-based approach that, besides the inclusion of reporting and user agents, include Negotiation and Diagnosing agents that are ultimately responsible for the most critical decisions, taken on the basis of information provided by two other groups of agents: data collecting and supervising. Rajput [18] use ontologies and rules in the creation of a specialized system to detect money laundering suspicious transactions.

The works mentioned above focus only on the signalling of suspicious transactions stage; use a small amount of data for testing; and do not assist the human expert in decision making. The system we present in this paper seeks to overcome the above mentioned limitations.

4. Building the Agents' Knowledge Bases

A database with real data reflecting the transactional behaviour of clients has a significant amount of attributes necessary for the control and management of the business involved. Of course, not all of these attributes are relevant to a search for suspicious transactions. The selection of the relevant attributes and the possible generation of new attributes must reflect the transactional database and allow the identification of suspicious behaviour [19].

Two years of current account transactions from a Brazilian bank were used in the first stage of the process. The accounts of the 5.2 million clients of this bank receive, on average, 85 million transactions per year. The analysis of the database showed that less than 10% of the clients are made up of corporate entities (commercial companies, industries, governments, etc.), however, they are responsible for more than 90% of the values transacted.

Tests carried out showed that the client type attribute was not enough to offer a good characterization of the groups formed in this unsupervised data base. Thus, to better characterize the clients, the database was divided: one with individual clients and another with corporate type clients. The procedure described below was performed independently for each database.

4.1. Client Transactional Behaviour Strategy

The rules that constitute the Agents' knowledge bases were obtained through a process detailed in Algorithm 1. From each base, transactions whose characteristics are irrelevant in this context (for example tariffs, commissions, interest, taxes, etc.) were excluded, resulting in 35 million relevant transactions (Algorithm 1 - Step 1).

With the purpose of establishing transactional behaviour profiles, in a certain period, we created a set of attributes, which aggregate quantities and segment characteristics for each actor of the process. The period measured is directly related to the nature of the business involved, presenting the maximum possible duration, for example quarterly, semi-annual, annual, etc.

One year of transactions was used to generate the client transactional behaviour profile, that is, database information such as: the account age, the number of transactions generated, the number of used services, the amount sent to other banks and to the accounts of the own bank, and the number of movements divided into six ranges of values. A twelfth attribute was created and named *debt percentage*, representing, in a weighted way in the period analysed, the time that the money remained in the customer account [20] [3] (Algorithm 1 - Step 2).

This database table with active customer profiles in the analysed year has 2.4 million lines, each line representing a unique element formed by 3-tuple: client, agency and account. Over this table, the non-supervised inductive learning procedure was executed using clustering, seeking groups of clients with similar and mutually exclusive characteristics. The algorithm K-means was used in classification and the algorithms PART and J48 in generation of production rules, executed 11 times (number of attributes minus 1) (Called algc, algr1 and algr2, respectively, in Algorithm 1 - Step 3).

Two sets of rules, one of each algorithm used, with the least number of incorrectly sorted instances were selected (Algorithm 1 - Step 4) and, together with the clusters that gave rise to these rules, represent the result of the process (Algorithm 1 - Step 5).

Algorithm 1 Learning Process

```
Input : TotalTransactionsBase-TTB
procedure LearningProcess( )
    RelevantTransactions-RT ← TTB analysis                              ▷ Step 1
    ClientProfiles-CP ← RT attribute selection / generation             ▷ Step 2
    NumberofClusters-k ← Number of CP Attributes - 1                    ▷ Step 3
    while $k > 1$ do
        SetofClusters-Cl ← Classification algorithm (algc (CP, k))
        VectorError1-E1 (k) ← Calculates Classification Error (algr1 (Cl))
        VectorError2-E2 (k) ← Calculates Classification Error (algr2 (Cl))
        k ← k - 1
    end while
    IndexofMinorError-idx ← Finds Minor (E1, E2)                        ▷ Step 4
    FinalClassification-FC ← algc (CP, idx)
    RuleSet1-R1 ← algr1 (FC)
    RuleSet2-R2 ← algr2 (FC)
    returns FC, R1, R2                                                  ▷ Step 5
end procedure
```

4.2. Cautious Approach to Risk

The analysis of the generated clusters, for the two customer segments, allowed the identification of characteristics such as: high turnover of high values, with full transfer to other financial institutions (high risk); or movement of values close to the legal limit to communication to regulatory agencies (moderate risk). This analysis resulted in the classification shown in Table 1.

With this classification, it is possible to define a better strategy, offering differentiated treatment to the groups of clients, according to their level of risk. Despite the excellent level of accuracy obtained in the evaluation of the generated rules, around 99% for both customer segments, an error of one percent represents more than 26 thousand transactions and cannot be ignored.

The confusion matrix generated by the rules algorithms has identified the rules that, because they are applicable to two or more groups of clients, represent the one percent error mentioned. That is, rules classify customers as belonging to more than one profile. Probably there is a configuration of the parameters of the algorithms used that allows to minimize this error, however, the number of rules increases significantly, making the cost / benefit little attractive. The decision was to reclassify the profiles that do not represent risk or have low risk (profiles one, two and three), as shown in Table 1. Thus, the transaction that belongs to one of these three groups will be reclassified, only for analysis process, if there is a rule that satisfies the condition. Thus, the problem was corrected conservatively using the same amount of rules.

For example, in the database used for data mining, 33 rules classify customers as individual client belonging to risk profiles two and three, corresponding to 1.85% of the total. However, these rules also classify 0.06% of customers originally belonging to the standard profile. The reclassification consists of, during the process of searching for a suspicious transaction, consider these standard clients as belonging to the risk groups, without modifying the original classification.

The classification of profiles also allows the creation of specific rules, whether based on current regulations or inspired by transactional behaviour. As already mentioned, this work used one year of information to generate the profiles, obtaining monthly totals and allowing to select maximum values

Table 1. Classification of Generated Profiles

Profile	Individual Client	Other type Customer	Reclassifies Profile
1. Low Utilization	Cluster4	Cluster4	Yes
2. Standard Client	Cluster2	Cluster1	Yes
3. Low Risk	Cluster1	-	Yes
4. Moderate Risk	Cluster5	Cluster3	No
5. High Risk	Cluster3	Cluster2	No

in the year for each relevant attribute. It was established that the search for suspicious transactions will always start one month before the date requested for analysis. In this way the behaviour for one month of transactions will always be used for comparison with the profiles.

5. AML Multi-Agent System

The goal of this system is to support the AML process in a financial institution and the global system architecture is presented in Figure 4.

The system will keep up a profile for each customer, based on the transaction history, which will be used along with the rules created from official regulations to combat money laundering, to the capture and signal suspicious transactions processed by the various business systems. The system will decide on some marked cases and learn from the aid provided by the AML analyst during the decision making process of the most complex cases. It will also monitor norms and recommendations posted by organs of control, flagging those involving money laundering. New rules to capture suspicious transactions, including new regulations and possible changes in profiles, will be suggested.

The database with profile history reflecting the learning period and a set of rules constitute the primary knowledge base of the agents. The set of rules is formed by:

1. DM Rules – rules generated in the data mining process, which will be used to review the original classification of clusters
2. Legal Rules – rules based on legal regulations and general guidelines that guide the fight against the crime of money laundering
3. Profile Rules – rules that reflect the specific knowledge of the bank about AML and the risk management according to the groups of profiles generated

The profiles to be analysed complete the set of information that will be manipulated by the system in the process of capture and analysis of suspicious transactions, the main focus of this work.

To achieve the proposed objectives, it is necessary to define a set of entities capable of making decisions, using existing knowledge and learning from the decisions taken by the human analysts. These entities need to have action autonomy and be able to communicate with one another. Besides, the system should be scalable and flexible. These features point to a set of deliberative agents that are able to review and extend their knowledge, working in a environment fully observable, deterministic, static and discrete [21].

To model this system we chose a methodology based on Wooldridge [22], and in the study about the major Agent-Oriented Software Engineering (AOSE) carried out by Bawa [23]. Bawa performed a comparison between five methodologies. The *Prometheus* methodology was selected, tested and verified to comply with all the criteria that we consider relevant to this work (ease of use; few graphics representation; accessible documentation, clear and preferably with examples; support tool for easy installation, preferably for windows environment).

Padgham [24] presents all aspects of this methodology, accompanied by a set of examples. The support tool is called PDT (Prometheus Design Tool), and works as an Eclipse plug-in. All these agents

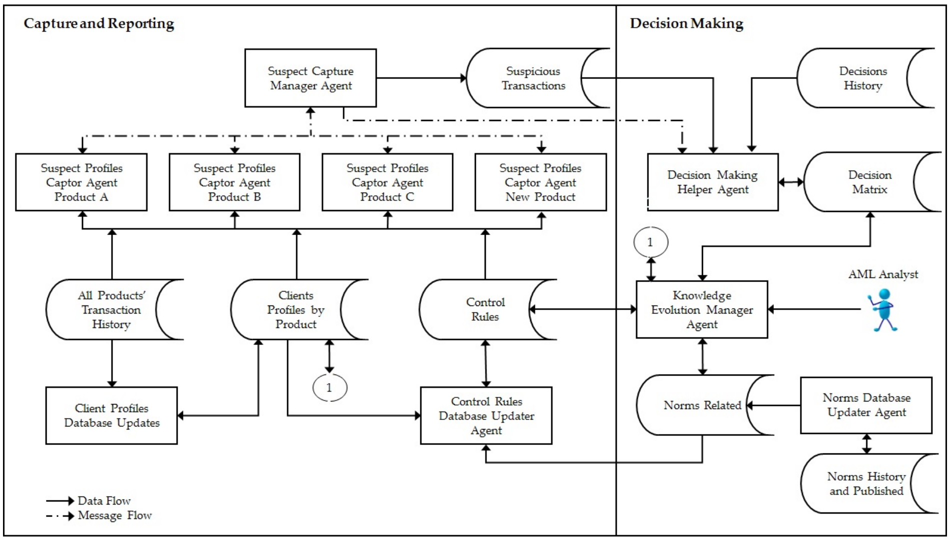

Capture and Reporting
Decision Making
Suspect Capture Manager Agent
Suspicious Transactions
Decisions History
Suspect Profiles Captor Agent Product A
Suspect Profiles Captor Agent Product B
Suspect Profiles Captor Agent Product C
Suspect Profiles Captor Agent New Product
Decision Making Helper Agent
Decision Matrix
1
AML Analyst
All Products' Transaction History
Clients Profiles by Product
Control Rules
Knowledge Evolution Manager Agent
1
Client Profiles Database Updates
Control Rules Database Updater Agent
Norms Related
Norms Database Updater Agent
Norms History and Published
Data Flow
Message Flow

Figure 4. Global System Architecture

and environmental characteristics are supported by the selected methodology. Although the detailed design produced by PDT can be straightforwardly converted to a specific and proprietary language, it is generic and can be used in a range of agent programming platforms.

Considering characteristics such as: quantity and relevance of available historical data that will be used for decision making; the aforementioned need, during the process, for review and expansion of the knowledge acquired; the possibility of developing new sub-objectives/objectives; we understand that Belief-Desire-Intention (BDI) is the appropriate model to adopt. Moreover,*Prometheus* is strongly targeted to BDI. Figure 5 shows the main system elements that make up the adopted BDI model and how they link to our problem.

Figure 6 shows a partial system overview (only the main elements) that results from the application of the methodology. The graphical symbols used in the system overview diagram represent *actors* like the AML Analyst; *data* which can be internal data indicating knowledge bases as the KB Control Rules or external as the DB Transaction History; *percepts* as Requests Suspicious Transactions; and *agents*. Many other diagrams, with additional details, were generated, however, they are omitted in this paper[2]. The system overview is enough in this context and the following sections describe the agents.

5.1. Agents Description

There is a Suspect Profiles Captor (SPC) agent for each product (current accounts, investment funds, currency exchanges, etc.). This approach has two advantages. Firstly, this allows us to model each agent's knowledge according to the specificities each product has. Secondly, it makes scalability easier, in the sense that the creation of a new product can be incorporated in the system just by adding a new agent specialized in it. In the analysis of transactions, agents use the current control rules generated by the previously described process, based on customer profiles, the norms on AML and in the internal rules of the financial institution.

Whenever a SPC agent identifies a suspicious transaction, it informs the Suspect Capture Manager agent (SCM) that in its turn is responsible for forwarding it to some other SPC agents. So, SPC agents have two working modes: *transaction oriented*, in which the agents try to capture suspicious transactions with no assumptions regarding clients; *client oriented*, in which agents try to capture

[2] The complete model can be accessed at: https://goo.gl/zHZFJG

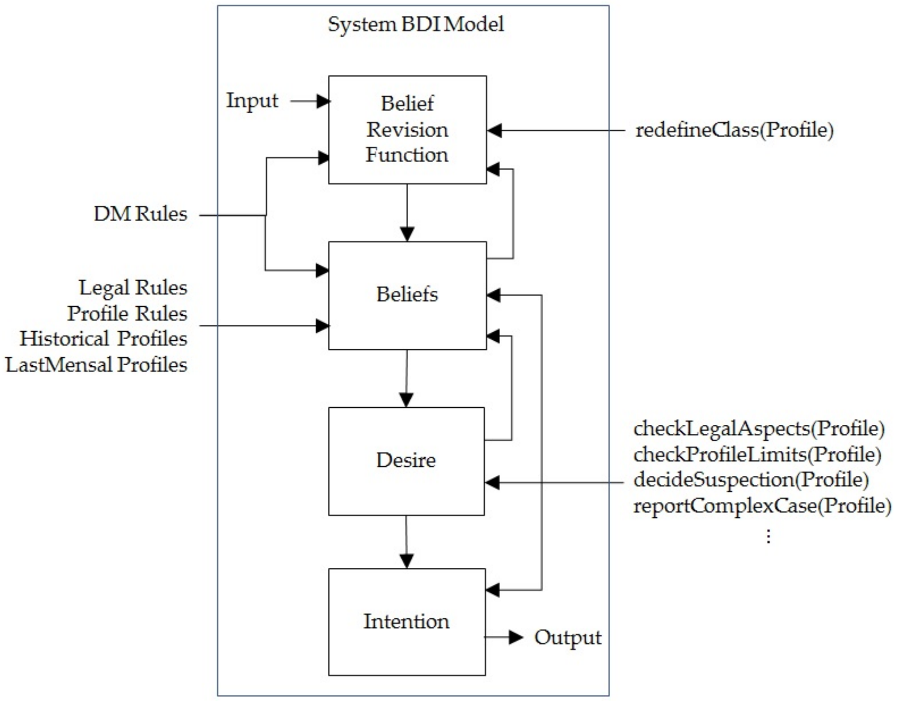


Figure 5. System BDI Model

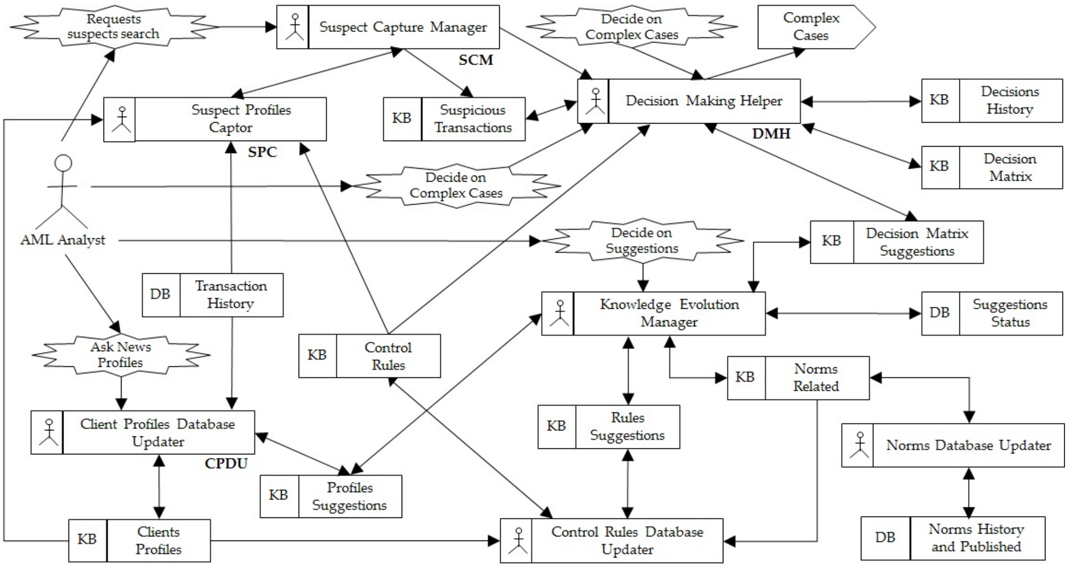


Figure 6. Partial System Overview

suspicious transactions for clients that were identified by another SPC agent and communicated by SCM.

The SPC agent decision making occurs when at least one rule in Legal Rules or Profile Rules is triggered and decision is based on the risk level of the profile, characterized mainly by the cluster to which it belongs.

A Suspect Capture Manager agent can receive an external analysis request and command for execution by specialist SPC or command it autonomously. When receiving information from an SPC that a suspicious transaction has been identified, it forwards it to some other SPC agents search in client mode. After receiving all reply messages, SCM announces to the Decision Making Helper agent (DMH) the existence of suspicious transactions. The default time for automatic capture execution and knowledge about existing SPC agents compose the basic information for agent decision making.

The Decision Making Helper agent performs the analysis of the previously signaled transactions and conducts a learning process. The DMH agent has autonomy to decide amongst three possibilities regarding a signalization: accept it, discard it, or send it for further (human) analysis. So, this agent assumes the role of the AML Analyst in the analysis of suspicious transactions. Decisions taken by this agent and those reported by the AML Analyst, are stored in the historical decisions database and are used in the learning process to evolve the decision matrix. This agent is also responsible for the changes suggested in the decision matrix and the update of this knowledge base, after the suggestions are validated by the AML Analyst.

The decision making of this agent uses the defined decision matrix and what was learnt from the history of decisions about complex cases.

The internal knowledge of the system is formed by the profiles base, the norms and recommendations base, the control rules base and the decision matrix. For each one of these knowledge bases there is an agent responsible for its evolution.

The Client Profiles Database Updater (CPDU) agent acts in the analysis of transaction history to generate client profiles and subsequent comparison with the base of existing profiles. This process can be triggered by a user request or autonomously by the agent. New arising profiles are suggested to the AML Analyst. The CPDU agent updates the profiles data-base with the profiles that are validated.

The Norms Database Updater (NDU) agent purpose is to search for norms and regulations newly published by official agencies, and that are related to the AML process. This agent acts independently seeking new published norms, saving them and then selecting those related with AML, suggesting them to the AML Analyst for study and possible validation. The validated norms are then incorporated into the historic database of norms.

The Control Rules Database Updater (CRDU) agent has as main objective to ensure the evolution of the control rules database, that is a key element in the system architecture. The first stage of this development process is the generation of new rules, confronted with the current rules and based on new profiles and new norms existing. CRDU then suggests these new rules to AML Analyst and the validated rules are incorporated into the control rules database.

The Knowledge Evolution Manager (KEM) agent executes tasks such as: controling the suggestions made by agents for evolution of the knowledge bases; maintaining interface with AML Analyst in the validation procedure, deletion or extension of these suggestions; and maintain communication with other agents to command and control all process. The interface maintained with the authenticated AML Analyst allows it to analyze and validate the suggestions.

The evolution of the knowledge bases and the learning process existing in system are intended, primarily, to mitigate the risk of the occurrence of false positive and / or false negative, common in systems based only on a set of rules and behaviour norms[16].

5.2. Agents Interaction

The framework chosen to implement the system, which will be discussed below, uses natively the internal structure of the decision making mechanism of the Procedural Reasoning System (PRS)

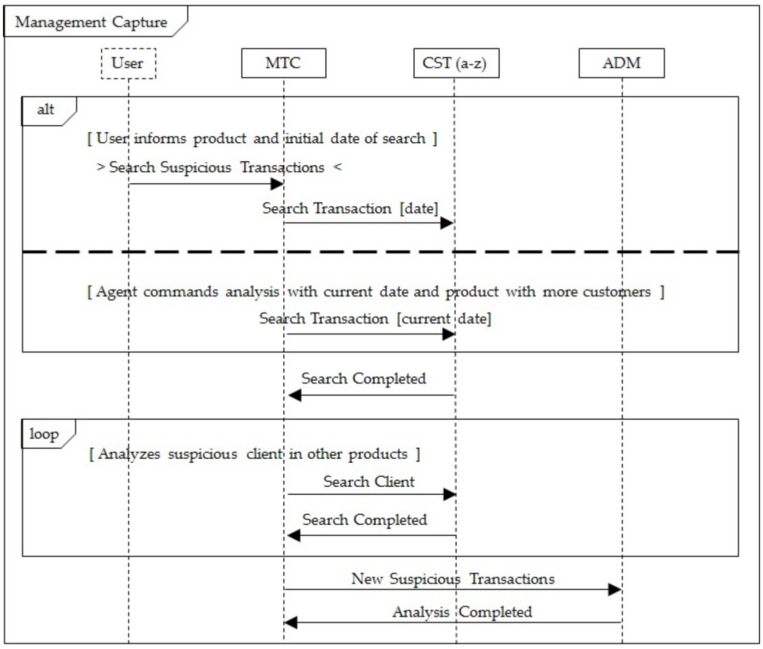

Management Capture
User
MTC
CST (a-z)
ADM
alt
[User informs product and initial date of search]
> Search Suspicious Transactions <
Search Transaction [date]
[Agent commands analysis with current date and product with more customers]
Search Transaction [current date]
Search Completed
loop
[Analyzes suspicious client in other products]
Search Client
Search Completed
New Suspicious Transactions
Analysis Completed

Figure 7. Example of the Interaction Process Between Agents

model and the Knowledge Query and Manipulation Language (KQML) to establish a simple but efficient communication process. "For example, the KQML performative *tell* is used with the intention of changing the receiver's beliefs, whereas *achieve* is used with the intention of changing the receiver's goals", that is, with the performative label it is possible to identify the intent of the message sender [25].

Cooperation among SPC agents happens through their direct interaction with the SCM agent, that coordinates the tasks and receives the results. Only the SCM agent knows all the SPC agents existing; besides, it can identify the products that the suspect client uses in the institution. Figure 7 shows, briefly, an example of the interaction process between agents.

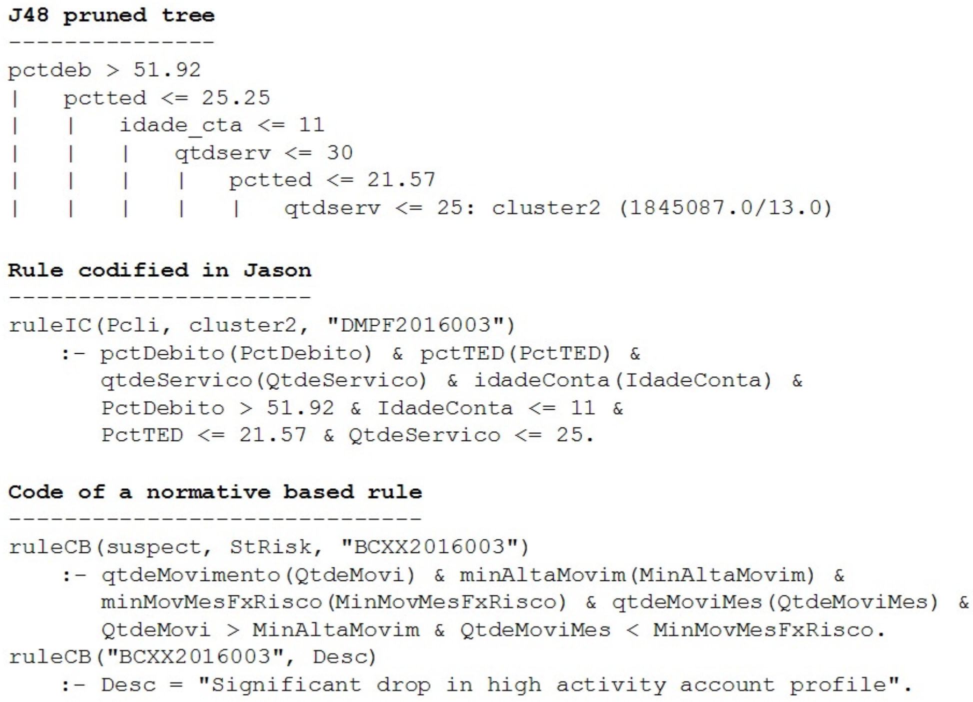
Interaction amongst other agents (DMH, MER, NDU, CRDU, CPDU) has the role to trigger in each of these agents the goal to perform the task under its responsibility. In other words, all agents have their own specific expertise, and they have independent and not conflicting goals. On the other hand, there is plenty of cooperation for the achievement of a common goal.

5.3. System Implementation

Departing from [26] and based on [27] and [28], the JaCaMo framework [29] was chosen. The tutorial presented in AAMAS2015 [30] reinforced this understanding. The framework is flexible in interaction with other programming languages, mainly java; native access to major database platforms; good interface and documentation; license free; and BDI native.

JaCaMo is based on three independent platforms: a) *Jason* for programming the agents level, inspired by the BDI architecture; b) *CArtAgO* for programming the environment level, that is composed of one or more workspaces, used to define the topology of the environment; c) *Moise* for programming the organizations level, defining the existing social groups and sub-groups, social tasks and behaviour rules that will allow the achievement of social goals [29] [25].

The rules obtained in the data mining process were written as Jason rules and used in a similar process as used in Prolog. The knowledge bases, and external databases, are accessed in MySQL using artefacts written in Java to improve the performance in access, due to the volume of data. All knowledge and external databases are used by the system as beliefs, representing the states of the environment and the knowledge acquired.



```
J48 pruned tree
----------------
pctdeb > 51.92
|   pctted <= 25.25
|   |   idade_cta <= 11
|   |   |   qtdserv <= 30
|   |   |   |   pctted <= 21.57
|   |   |   |   |   qtdserv <= 25: cluster2 (1845087.0/13.0)

Rule codified in Jason
----------------------
ruleIC(Pcli, cluster2, "DMPF2016003")
    :- pctDebito(PctDebito) & pctTED(PctTED) &
       qtdeServico(QtdeServico) & idadeConta(IdadeConta) &
       PctDebito > 51.92 & IdadeConta <= 11 &
       PctTED <= 21.57 & QtdeServico <= 25.

Code of a normative based rule
------------------------------
ruleCB(suspect, StRisk, "BCXX2016003")
    :- qtdeMovimento(QtdeMovi) & minAltaMovim(MinAltaMovim) &
       minMovMesFxRisco(MinMovMesFxRisco) & qtdeMoviMes(QtdeMoviMes) &
       QtdeMovi > MinAltaMovim & QtdeMoviMes < MinMovMesFxRisco.
ruleCB("BCXX2016003", Desc)
    :- Desc = "Significant drop in high activity account profile".
```

Figure 8. Examples of rules, as codified in the KB.

6. Results

The real data used in this work refer to two years, with 30.5 million and 35.2 million relevant transactions, respectively. The client transactional behavioural profiles were generated from the data of the first year, the reference base. The search for suspicious transactions was executed over six months of relevant transactions in the second year, resulting in 17.1 million transactions. Over these transactions, 3.2 million transactional client behavioural profiles were generated for those months.

The systems in use in most banking institutions are strongly based on client information bank register, that's why they can identify as suspicious the incompatibility between the sum of transaction values and the customer income or billing information. The system proposed in this article does not use the same client information bank register and therefore will not signal such situations as suspicious. This system is based on the client's transactional behaviour and should be adopted as a complementary tool. The initial strategy used to select suspect profiles for verification by human analysts and examples of reports issued by system was shown in [31].

The characteristics described above justify why the result obtained by the system here presented did not identify the same suspects identified by systems actually running. However, it is important to highlight that this system has signalled cases that were confirmed by human analysts and that were not identified in the past by the systems in execution.

The process of searching for suspicious transactions, implemented to date, can be divided into 3 phases: reclassifying the profiles or adjusting the confusion matrix; the capture of suspicious transactions; and the analysis of the captured transactions. The system uses as beliefs a set of 132 rules (104 of classification, 28 normative and profiles based), as well as the classification of customers in the indicated profiles. These beliefs are re-evaluated only once for each capture process. For example, in the first month evaluated, 418 profiles were reclassified, leaving the "standard profile" to "medium" and "high risk" profiles, only in that analysis. Table 2 shows a summary of the results obtained, where profiles are indicated as suspicious in the six months analysed. Figure 8 shows some examples of the rules used in the process.

Table 2. Summary of results for 6 months

Reference	Profiles			Suspects by Profiles Risc Classification				
Month	Analyzed	Reclassified	Suspicious	Low Utilization	Standard Client	Low Risk	Medium Risk	High Risk
Jan/2015	444.819	418	182	8	32	1	47	94
Feb/2015	408.302	406	148	9	22	0	42	75
Mar/2015	447.542	395	254	21	19	1	66	147
Apr/2015	427.415	392	216	13	20	0	58	125
May/2015	447.683	391	246	13	19	1	67	146
June/2015	474.484	388	297	15	23	0	76	183
Average Totals	**441.708**	**398**	**224**	**13**	**23**	**1**	**59**	**128**
		0,09%	**0,05%**					

Table 3. AML Analist analysis result

	Suspects Gruped by Risc Classification				Previously Signalized by Other System	
Result of Human Analysis	**Low Utilization and Standard Client**	**Low Risk and Medium Risk**	**High Risk**	**Total**	**Yes**	**No**
Not Suspect	13	6	7	26	19	7
Suspect but needs to deepen analysis	2	3	1	6	5	1
Suspicion Confirmed	4	0	2	6	2	4
Total	**19**	**9**	**10**	**38**	**26**	**12**

The average percentage of 0.05% of suspect profiles with respect to the period and the amount of profiles analysed is a feasible amount to be analysed by human analysts, considering a normal AML process. However, considering that the analysis of these results is an additional task to the analysts' daily routine, we have reduced the number of profiles that will be investigated. The strategy adopted was to investigate 38 profiles of which 16 were repeatedly flagged as suspects in the six months analysed plus 22 other profiles flagged in the five months.

Table 3 shows the AML Analyst investigation result. It is important to highlight the following aspects about the 38 analysed cases: a) 26 cases were not considered suspects, but the analysts did not consider them false positive, because there were indications of non-standard procedures; b) 12 cases were confirmed as suspects, six of which will required more in-depth investigations by bank branch involved, and six were clearly identified as suspects; c) all cases classified by the system as being "high risk" had not been previously reported; d) of the six fully confirmed suspects 4 had never been previously reported.

Considering the purpose of the system to assist the AML analyst and its learning ability, it is possible to observe that the analysis phase need refinement, considering the 26 cases indicated as suspicious but not confirmed, even though these were not considered false positive.

7. Conclusions

The systems currently in use by financial institutions are strongly based on client information bank register. They have failed in the steps of capture suspicious transactions and do not assist human specialists.

With our work we explored new approaches to combat both fraud and money laundering, and have described a multiagent system that is successful in this task, using specialization and cooperation between intelligent agents, in order to optimize and improve the quality of the process of signalling suspicious profiles in the anti-money laundering process.

To build the agents' knowledge bases we used some machine learning techniques to identify risk groups and to create the client transactional behaviour profiles. These profiles were used as a marker for future behaviour.

The results obtained show the feasibility of systematic use and establish a new front to combat this crime. The quality of the results has been attested by the verification realized by the anti-money laundering analyst in the signalled suspicious transactions, in which it is worth highlighting that *all cases classified by the system as being "high risk" suspects had not been previously reported*; of the 6 fully confirmed suspects 4 had never been previously reported by any other system running.

As next future step we will review and improve the analysis phase, considering the 26 cases indicated as suspicious but not confirmed, even though these were not considered false positive. The strategy will be to complete the learning based on the decisions of the AML analyst, thus allowing better signalling.

References

1. Madinger, J. *Money laundering: a guide for criminal investigators*, 3th ed.; CRC: Boca Raton, FL, 2012.
2. Cser, A. The Truth About Machine Learming in Fraud Prevention [WEBINAR]. Kount, 2017. Retrieved from http://www.kount.com/kount-events/webinar-the-truth-about-machine-learning-in-fraud-prevention.
3. Alexandre, C.; Balsa, J., New Advances in Information Systems and Technologies; Springer International Publishing: Cham, 2016; Vol. 1, chapter Integrating Client Profiling in an Anti-money Laundering Multi-agent Based System, pp. 931–941.
4. Committee in Basle. Customer Due Diligence for Banks. Technical report, Working Group on Cross-border Banking - Basel Committee on Banking Supervision - Bank for International Settlements, 2001.
5. Canas, V. *O Crime de Branqueamento: Regime de Prevenção e de Repressão*; 2004. 335p.
6. Alexandre, C.; Balsa, J. A Multiagent Based Approach to Money Laundering Detection and Prevention. Proceedings of the International Conference on Agents and Artificial Intelligence; Loiseau, S.; Filipe, J.; Duval, B.; van den Herik, H.J., Eds.; SciTePress: Lisbon, 2015a; Vol. 1, pp. 230–235.
7. Senator, T.E.; Goldberg, H.G.; Wooton, J.; Cottini, M.A.; Khan, A.F.U.; Klinger, C.D.; Llamas, W.M.; Marrone, M.P.; Wong, R.W.H. The Financial Crimes Enforcement Network AI System (FAIS) Identifying Potential Money Laundering from Reports of Large Cash Transactions. *AI Magazine* **1995**, *16*, 21–39.
8. Zhang, Z.M.; Salerno, J.J.; Yu, P.S. Applying Data Mining in Investigating Money Laundering Crimes. Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; ACM: New York, NY, USA, 2003; KDD '03, pp. 747–752.
9. Kingdon, J. AI Fights Money Laundering. *IEEE Intelligent Systems* **2004**, *19*, 87–89.
10. Liu, X.; Zhang, P. A Scan Statistics Based Suspicious Transactions Detection Model for Anti-money Laundering (AML) in Financial Institutions. Multimedia Communications (Mediacom), 2010 International Conference on, 2010, pp. 210–213.
11. Tang, J.; Yin, J. Developing an intelligent data discriminating system of anti-money laundering based on SVM. Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on, 2005, Vol. 6, pp. 3453–3457.
12. Le-Khac, N.A.; Kechadi, M.T. Application of Data Mining for Anti-money Laundering Detection: A Case Study. Proceedings of the 2010 IEEE International Conference on Data Mining Workshops; IEEE Computer Society: Washington, DC, USA, 2010; pp. 577–584.
13. Chang, W.; Chang, J. Using clustering techniques to analyze fraudulent behavior changes in online auctions. Networking and Information Technology (ICNIT), 2010 International Conference on, 2010, pp. 34–38.
14. Larik, A.S.; Haider, S. Clustering based anomalous transaction reporting. *Procedia Computer Science* **2011**, *3*, 606 – 610.
15. Sabau, A.S. Survey of Clustering based Financial Fraud Detecton Research. *Informatica Economica* **2012**, *16*.
16. Gao, S.; Xu, D.; Wang, H.; Wang, Y. Intelligent Anti-Money Laundering System. Service Operations and Logistics, and Informatics, 2006. SOLI '06. IEEE International Conference on, 2006, pp. 851–856.
17. Xuan, L.; Pengzhu, Z. An Agent Based Anti-Money Laundering System Architecture for Financial Supervision. Wireless Communications, Networking and Mobile Computing, 2007. International Conference on, 2007, pp. 5472–5475.

18. Rajput, Q.; Khan, N.S.; Larik, A.S.; Haider, S. Ontology Based Expert-System for Suspicious Transactions Detection. *Computer and Information Science* **2014**, *7*, 103–114.
19. Paula, E.L.; Ladeira, M.; Carvalho, R.N.; Marzagão, T. Deep Learning Anomaly Detection as Support Fraud Investigation in Brazilian Exports and Anti-Money Laundering. 15th IEEE International Conference on Machine Learning and Applications, ICMLA 2016, Anaheim, CA, USA, December 18-20, 2016, 2016, pp. 954–960.
20. Alexandre, C.; Balsa, J. Client Profiling for an Anti-Money Laundering System. *ArXiv e-prints* **2015b**, [arXiv:cs.LG/1510.00878]. http://adsabs.harvard.edu/abs/2015arXiv151000878A.
21. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall Press: Upper Saddle River, NJ, USA, 2009.
22. Wooldridge, M. *An Introduction to Multiagent Systems*, 2nd ed.; Wiley Publishing: Chichester, UK, 2009.
23. Bawa, A.; Attri, V.K. A Study of Tools Used in Implement Agent Oriented Software Engineering. *International Journal of innovative Research in Computer and Communication Engineering* **2015**, *3*. http://www.ijircce.com/upload/2015/may/40_A_Study.pdf, Acessado em 15/09/2015.
24. Padgham, L.; Winikoff, M. *Developing Intelligent Agent Systems: A Practical Guide*; John Wiley & Sons Ltd: Chichester, West Sussex PO19 8SQ, England, 2004; chapter 8, pp. 155–171.
25. Bordini, R.H.; Hübner, J.F.; Wooldridge, M. *Programming Multi-Agent Systems in AgentSpeak Using Jason (Wiley Series in Agent Technology)*; John Wiley & Sons, 2007.
26. Fisher, M.; Bordini, R.; Hirsch, B.; Torroni, P. Computational Logics and Agents - A Roadmap of Current Technologies and Future Trends. *Computational Intelligence* **2007**, *23*, 69–91.
27. Boissier, O.; Bordini, R.; Hübner, J.; Ricci, A.; Santi, A. Multi-agent Oriented Programming with JaCaMo. *Sci. Comput. Program.* **2013**, *78*, 747–761.
28. Shehory, O.; Sturm, A. *Agent-Oriented Software Engineering: Reflections on Architectures, Methodologies, Languages, and Frameworks*; SpringerLink : Bücher, Springer Berlin Heidelberg, 2014.
29. Bordini, R.H.; Dix, J. Programming Multiagent Systems. In *Multiagent Systems*, 2nd ed.; Weiss, G., Ed.; The MIT-Press: London, England, 2013; chapter 13, pp. 587–639.
30. Baldoni, M.; Baroglio, C.; Capuzzimati, F.; Micalizio, R. Programming with Commitments and Goals in JaCaMo+. Proceedings of the 2015 International Conference AAMAS; , 2015; pp. 1705–1706.
31. Alexandre, C.; Balsa, J. Um Sistema Multiagente no Combate ao Branqueamento de Capitais. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação* **2017**, pp. 1 – 17.