Supplier Impersonation Fraud Detection using Bayesian Inference

Rémi Canillas
SiS-id - Liris
INSA Lyon
Lyon, France
remi.canillas@sisnet.fr

Omar Hasan
Liris
INSA Lyon
Lyon, France
omar.hasan@insa-lyon.fr

Laurent Sarrat
SiS-id
Lyon, France
laurent.sarrat@sisnet.fr

Lionel Brunie
Liris
INSA Lyon
Lyon, France
lionel.brunie@insa-lyon.fr

***Abstract*—In this paper, we introduce ProbaSIF, a supplier impersonation fraud detection system that relies on a Bayesian model to perform the classification of a new transaction as legitimate or fraudulent. ProbaSIF is divided in two parts: an intra-company analysis that aims to recreate the vision of a specific client about the legitimacy of the account used in a transaction with one of its supplier, and an inter-company analysis that uses all the accounts used to pay a supplier to model the supplier's payment behavior and take into account transactions issued by other clients. We use a dataset composed of more than 2 million transactions issued by real companies, provided by the SiS-id platform, to fit our Bayesian model, and evaluate the classification results of ProbaSIF using an other set of 108,000 transactions labeled by SiS-id expert system. Our study of a representative client shows that both of the approaches described in ProbaSIF show good precision (0.927 and 0.836) for the 255 transactions tested. Results also shows that ProbaSIF gives results consistent with the expert system provided by SiS-id. Finally, after evaluating ProbaSIF approaches on all the clients available in our dataset, we demonstrated that our classification system was accurate for a wide set of different clients.**

***Index Terms*—Fraud detection systems, bayesian models, fraud prevention, anomaly detection, unsupervised learning.**

In this paper, we propose ProbaSIF, an unsupervised supplier impersonation fraud detection system based on statistical analysis and Bayesian Inference. ProbaSIF uses a set of historical transactions issued on a Business-to-Business ecosystem involving several companies exchanging goods and services. The goal of ProbaSIF is twofold: Firstly, identify the behavior of a company emitting payments (called the client company) and of a company receiving payments (called supplier company). Secondly, produce an alert if an unusual transaction is emitted by a client company for a supplier company with respect to both of their behavior models, indicating a potential impersonation fraud. This alert can then be transmitted to both companies' fraud investigation team in order to validate or cancel the transaction.

Our system first uses probability theory to compute probability distributions representing the underlying payment behavior of a client and a supplier. This step can be conducted in an offline fashion, where the distribution of probability of an account being used to pay a supplier is calculated. Then, when a new transaction's legitimacy needs to be established, the account used in the transaction is compared with the client and supplier's probabilistic models. If the probability to see this account used to pay the supplier is lower than an user specified threshold, then the transaction's legitimacy is deemed low and the transaction considered fraudulent. The focus of this fraud detection system is to model a client and a supplier's behavior and then determine how a new transaction fits this model.

Using the assumption that a relatively simple data-driven system can be effective in a complex situation such as Fraud Detection [14], ProbaSIF relies on two separate analysis: one conducted using only the information available for the client conducting the transaction, in order to model the client's view when assessing the potential fraud, and the other one using the information gathered by the supplier receiving the transaction. The probability of occurrence of the account used in a new transaction is determined using both of these models, and compared with a risk threshold in order to assert the legitimacy of the transaction.

ProbaSIF is evaluated using a set of real transactions provided by the SiS-id platform, and its results shows that it achieves an accuracy of 92.7% for a low time to detection (570 ms) in our experimental setting. Its performance are also consistent with expert knowledge and even show more decisiveness, and moreover, ProbaSIF is performative to a wide set of different clients.

The contributions of this paper are the following:

1) A Bayesian model describing the probabilty of occurrences of an account in a set of real-life transactions between companies interacting in a B2B ecosystem (collected by SiS-id), in order to model the underlying behavior of these companies.
2) A performative classification system allowing to attribute a legitimacy label to a new transaction in a simple and interpretable way, in order to detect fraudulent transactions.
3) An experimental setup allowing us to perform a comparative analysis of ProbaSIF's results and SiS-id expert system's results, using a set of real-life transactions previously labeled by SiS-id's expert system.

2375-9356/20/$31.00 ©2020 IEEE
DOI 10.1109/BigComp48618.2020.00-53

I. Background

Lately, Supplier Impersonation Fraud (SIF) is on the rise, resulting in the loss of hundreds of thousands of Euros in 2018, and ranked 1st most frequent fraud affecting French companies in the latest survey about cyber-criminality conducted in 2019 by Euler Hermes and DFCG [11].Supplier impersonation fraud consists in a fraudster impersonating a member of a company providing goods and service to another, in order to trigger a payment on an account controlled by the fraudster. [2]. More and more companies are using digital tools to process, authorize, or even conduct transactions due to numerous advantages provided by digitalization such as the ability to conduct transaction all over the globe in a timely fashion. However, digital transactions make frauds against companies more effective, firstly due to the difficulty to formally identify and trust remote interlocutors that are sometimes geographically very distant from the company headquarters, and secondly due to the increased speed of wired transactions, allowing money to be moved from accounts to accounts in a very short amount of time, thus hindering the process of recovering it after a fraud. The SiS-id company proposes to build a platform aggregating the transactions issued by a set of companies involved in a client/supplier relationship in order to build an accurate model of the B2B ecosystem created by the flow of payments. This model can then be used to detect and prevent Supplier Impersonation Fraud by investigating anomalous behaviors in this ecosystem.

II. Related Work

Supplier Impersonation Fraud (or SIF) is a relatively uncovered area of Fraud Detection, as it is both a sensitive topic, as being a victim of fraud on both the client and supplier side can lead to a breach of trust and thus hinder the economical relationship built between a supplier and a client, and because it requires the analysis of confidential data in the form of the transactions issued by a client to company to its suppliers. This data is sensitive because, in the hand of a competitor, it can be used to launch devastating economical attacks. For these reasons, very few papers [7] in the literature proposes frameworks and solutions in order to mitigate SIFs.

However, one can take inspiration in other fraud detection frameworks as described in [1], [6], and [8]. Along the various existing techniques such as statistical analysis for fraud detection [4] and feature engineering for fraud detection [9], Bayesian-based solution stands out for their relative simplicity [13] and their ability to build unsupervised fraud detection systems that does not rely on a previously labeled dataset in order to compute their model [17]. These techniques have proved reliable even when used with consequent dataset [10], and have been democratized thanks to the worldwide adoption of the Nave Bayes Classifier [16].

Successful application of the Bayesian model can be found in various applications and topics such as modeling animal survival [5], detecting fraudulent ratings in a online shop [15], or detecting faulty battery in satellites [12]. However, to the best of our knowledge, our work is the first attempt to provide a Bayesian model to detect Supplier Impersonation Frauds.

TABLE I: Commonly used notations. Vectors are in **bold**.

Parameter	Interpretation
$t = \{c, a, s, d\}$	B2B transaction involving client c, supplier s and account a at time d.
C	No. of clients.
S	No. of suppliers.
$T^{c,s}, T^{s}$	No. of transactions involving client c and supplier s (resp. supplier s).
N	No. of accounts.
$N^{c,s}, N^{s}$	No. of accounts used by client c to pay supplier s (resp. No. accounts used to pay supplier s).
$\boldsymbol{a^{c,s}}$,$\boldsymbol{a^{s}}$	Vector $(a_i)_{i=1}^{N^{c,s}}$ (resp. $(a_i)_{i=1}^{N^{s}}$) of all accounts involving c and s (resp. s).
$\boldsymbol{t^{c,s}}$,$\boldsymbol{t^{s}}$	Vector $(t_i)_{i=1}^{T^{c,s}}$ (resp. $(t_i)_{i=1}^{T^{s}}$) of all transactions involving c and s (resp. s).
$\boldsymbol{p^{cli}_{c,s}}$	Probability of a random account to be used by c to pay s.
$\boldsymbol{p^{com}_{s}}$	Probability of a random account to be used to pay s.
$\boldsymbol{\alpha^{cli}_{c,s}}$, $\boldsymbol{\alpha^{com}_{s}}$	Dirichlet parameters for the accounts' distribution describing client c payment behavior with s (resp. supplier s payment behavior).
$\delta 1, \delta 2$	Risk thresholds for discretization.

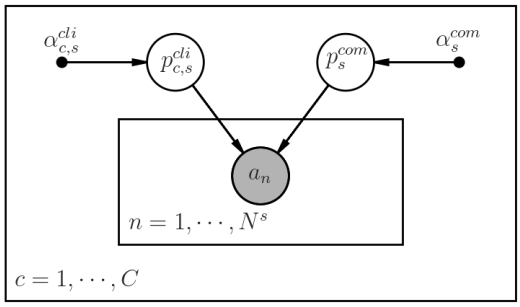


Fig. 1: Bayesian Model describing the probability of usage of an account.

III. Proposed Model

In this section we present the generative model used to describe a client and a supplier payment behavior in a B2B ecosystem. Table I summarizes the notations used in this paper. This model encompasses two different perspectives on the context in which a transaction is generated: as part of a client's behavior in the payment of its supplier, or as

part of a supplier's behavior in receiving payments. These two analysis are performed separately for each client $c \in C$. Figure 1 proposes a graphical representation of the described model.

A. Intra-Company Analysis

The aim of the intra-company analysis is to focus on the risk of fraudulence of a transaction t with respect to the behavior of the client that issued the transaction. This model aims to recreate the narrow vision of a singular client that uses only its own information to detect potentially fraudulent activity. This approach is motivated by the fact that the clients are particularly well-informed about how they pay their own suppliers, and thus aims to model this knowledge in a useful way for fraud detection.

Formally, we represent the client's payment behavior by the probability of using an account $a \in \boldsymbol{a^{c,s}}$ to pay a supplier s. As the account a is a categorical variable, we can model the account choice as a Multinomial($\boldsymbol{a^{c,s}}$,$\boldsymbol{p^{cli}_{c,s}}$) distribution where $\boldsymbol{p^{cli}_{c,s}}$ is a vector of length $N^{c,s}$ of non-negative entries that sums to 1 representing the probability of an account a to be used by c to pay s. As we have no pre-determined beliefs of the client c's payment behavior, we consider that we have no prior knowledge about this distribution, and thus we draw $\boldsymbol{p^{cli}_{c,s}}$ from a Dirichlet($\boldsymbol{\alpha^{cli}_{c,s}}$) distribution. Using this distribution to model the parameters of $\boldsymbol{p^{cli}_{c,s}}$ allows us to consider the variations in behavior of c through time. A comprehensive definition of the Dirichlet distribution can be found in [3]. This generative model is summarized below:

$$\mathbf{p^{cli}_{c,s}} \sim \text{Dirichlet}(\alpha^{cli}_{c,s})$$
$$a \sim \text{Multinomial}(a^{c,s}, \mathbf{p^{cli}_{c,s}})$$

B. Inter-Company Analysis

The aim of the inter-company analysis is to focus on the risk of fraudulence of a transaction t with respect to the behavior of the supplier that received the transaction. This addition to the model is motivated by the fact that a single supplier s can used several accounts to be paid, thus showing a more complex behavior than the one witnessed by only one client. This approach is motivated by the fact that some suppliers might use different accounts when they are being paid, and thus an unusual account for a specific client can be legitimate for the supplier, as this account has been used to receive payment from another client.

Formally, we represent the supplier's payment behavior by the probability of using an account $a \in \boldsymbol{a^s}$ to pay the supplier s. We use the same motivation as before to consider a as drawn from a Multinomial($\boldsymbol{a^s}$,$\boldsymbol{p^{com}_s}$) distribution where $\boldsymbol{p^{com}_s}$ is a vector of length N^s of non-negative entries that sums to 1, representing the probability of an account a to be used to pay s. The foremost difference with the intra-company analysis model is that this distribution encompasses all the accounts used by all the clients involved in a transaction with s, and not only a single client. Similarly to the intra-company analysis, we have no pre-determined beliefs of supplier s's payment behavior, and thus we consider that we have no prior knowledge about the distribution $\boldsymbol{p^{com}_s}$, and so we draw $\boldsymbol{p^{com}_s}$ from a Dirichlet($\boldsymbol{\alpha^{com}_s}$) distribution. This generative model is summarized below:

$$\mathbf{p^{com}_s} \sim \text{Dirichlet}(\alpha^{com}_s)$$
$$a \sim \text{Multinomial}(a^s, \mathbf{p^{com}_s})$$

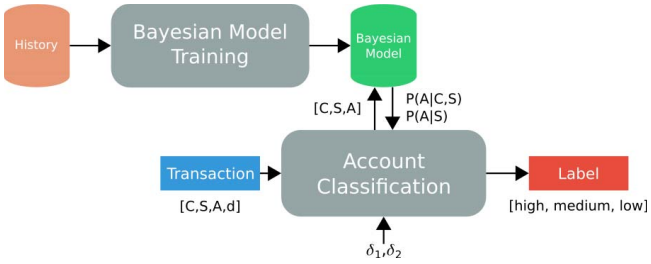


Fig. 2: Overview of the proposed fraud detection system.

IV. Proposed Algorithms

In order to assert the legitimacy of a transaction using the proposed Bayesian model, we propose to update the values $\boldsymbol{p^{cli}_{c,s}}$ and $\boldsymbol{p^{com}_s}$ according to the evidences found in a set of historical transactions, and then use $p^{cli}_{c,s}(a)$ and $p^{com}_s(a)$ the probability of occurrence of an account a used in a transaction t as an indicator of t's legitimacy. We then apply a threshold function to transform the probability value in a class label indicating the legitimacy of a transaction, according to user-defined risk thresholds $\delta 1$ and $\delta 2$.

Figure 2 shows an overview of the fraud detection process. First, the historical transactions are used as evidence in order to update the parameters of the bayesian model. Then, when the legitimity of a new transaction needs to be asserted, the probability of occurrence of the account are retrieved from the model and compared with user-defined thresholds, in order to assign a label to the tested transaction.

A. Probability Determination

In order to determine the value of each probability in $\boldsymbol{p^{cli}_{c,s}}$ and $\boldsymbol{p^{com}_s}$, we gather the historical transactions issued to s as a set of evidence that we use to update $\boldsymbol{\alpha^{cli}_{c,s}}$ and $\boldsymbol{\alpha^{com}_s}$ in a straightforward way described in 1, where $occ(a, \mathbf{t})$ is the number of occurrences of a in the set of transactions $\mathbf{t}$. $p^{cli}_{c,s}(a)$ and $p^{com}_s(a)$ are then determined as the point estimates of $\alpha^{cli}_{c,s}(a)$.

In this algorithm, the parameter $\alpha^{com}_s(a)$ that represents the probability of occurrences $p^{com}_s(a)$ of each account $a \in \mathbf{a^s}$ to be used to pay supplier s is computed as the number of occurrences $occ(a, \mathbf{t^s})$ of account a in the set historical transactions $\mathbf{t^s}$, divided by the total number of transactions where s is involved T^s. Additionally, if a is also part of the set of transactions $\mathbf{t^{c,s}}$ that c emits to pay s, then the parameter

Algorithm 1: Classification of a transaction using our Bayesian model and risks thresholds.

Data:

- $t^{c,s}$: set of historical transactions involving c and s
- t^{s}: set of historical transactions involving s
- $T^{c,s}$: total number of historical transactions involving c and s
- T^{s}: total number of historical transactions involving s
- $a^{c,s}$: set of account used by c and s
- a^{s}: set of account used by s

Result:

- $\boldsymbol{\alpha^{cli}_{c,s}}$
- $\boldsymbol{\alpha^{com}_{s}}$

foreach a *in* a^s **do**
  if a *in* $a^{c,s}$ **then**
    $\alpha^{cli}_{c,s}(a) = \frac{occ(a,t^{c,s})}{T^{c,s}}$;
  $\alpha^{com}_{s}(a) = \frac{occ(a,t^{s})}{T^{s}}$;

$\alpha^{cli}_{c,s}(a)$ that represents the probability of occurrences $p^{cli}_{c,s}(a)$ of each account $a \in \mathbf{a^{c,s}}$ to be used by c to pay supplier s is also computed. By definition, we have $\alpha^{com}_{s}(a) \geq \alpha^{cli}_{c,s}(a)$, as $\alpha^{com}_{s}(a)$ can be seen as the marginalization of $\alpha^{cli}_{c,s}(a)$ over all the clients in C.

A limitation of the system arises here: if T^s and/or $T^{c,s}$ are low, then the Law of Large Number [18] cease to apply and the values computed by the algorithm start to take extreme values instead of representing the actual payment behavior of the client or supplier. However this issue might be beneficial in our setting, as extreme values will automatically trigger an either extremely favorable answer for known accouts, or defavorable answer for unknown accounts, which is a desired outcome in practice.

B. Classification

Once $\boldsymbol{p^{cli}_{c,s}}$ and $\boldsymbol{p^{com}_{s}}$ are computed, we use Algorithm 2 to append a legitimacy label to a proposed transaction. In this algorithm, the values $p^{cli}_{c,s}(\tilde{a})$ and $p^{cli}_{c,s}(\tilde{a})$ representing the probability of occurrence of the account $\tilde{a}$ used in the investigated transaction $\tilde{t}$ are compared with two user-defined risk thresholds $\delta 1$ and $\delta 2$ representing the minimum and maximum probability value from which a transaction is considered as fraudulent or legitimate. Any probability value found in between $\delta 1$ and $\delta 2$ is considered as moderately suspicious.

V. Experimental Results

In order to evaluate the two approaches implemented in ProbaSIF, we use a dataset of more than 2 millions historical transactions provided by SiS-id to fit our model. We then evaluate the classification results of ProbaSIF with a set of 108.000 transactions already labeled by SiS-id's expert system. However, instead of analyzing the whole dataset globally, we partition the evaluation process in C different analysis, one for each client, in order to answer the following questions:

Algorithm 2: Fitting the parameters for the Bayesian model using historical transactions.

Data:

- $\tilde{t} = \{\tilde{c}, \tilde{a}, \tilde{s}, \tilde{d}\}$: Transaction to label.
- $\boldsymbol{p^{cli}_{c,s}}$: Probability distribution of accounts (client analysis)
- $\boldsymbol{p^{com}_{s}}$: Probability distribution of accounts (supplier analysis.)
- $\delta 1, \delta 2$: Risk thresholds.

Result:

- $c^{cli}(\tilde{t})$: Class label for $\tilde{t}$ (client analysis)
- $c^{com}(\tilde{t})$: Class label for $\tilde{t}$ (supplier analysis)

Get $p^{cli}_{c,s}(\tilde{a})$;
if $p^{cli}_{c,s}(\tilde{a}) > \delta 2$ **then**
  $c^{cli}(\tilde{t}) =$ "high"
else
  if $p^{cli}_{c,s}(\tilde{a}) > \delta 1$ **then**
    $c^{cli}(\tilde{t}) =$ "medium"
  else
    $c^{cli}(\tilde{t}) =$ "low"
Get $p^{com}_{s}(\tilde{a})$;
if $p^{com}_{s}(\tilde{a}) > \delta 2$ **then**
  $c^{com}(\tilde{t}) =$ "high"
else
  if $p^{com}_{s}(\tilde{a}) > \delta 1$ **then**
    $c^{com}(\tilde{t}) =$ "medium"
  else
    $c^{com}(\tilde{t}) =$ "low"

- Q1 - Precision on real data: does ProbaSIF two approaches precisely detect fraudulent transactions ?
- Q2 - Consistency with expert knowledge : does the proposed approaches are consistent with the expert-based system ?
- Q3 - Adaptability : is ProbaSIF performative for every client in our dataset ?

For the sake of brevity, we propose to first use a single representative client as a test case to study Q1 and Q2. We then repeat the same experimental process to evaluate the performance of ProbaSIF for each of the 83 clients founds in the set of transactions evaluated by SiS-id expert system in order to answer Q3.

A. History Dataset

In order to fit our models, we use a set of B2B transactions provided by the SiS-id platform, aggregating the transactions carried between July 2016 and July 2019 between 5,921 companies. We dubbed this dataset "History". Table II sums up the features available from this dataset. Depending on the transactions' sources, more data can be available, such as the amount of the transaction, or details about the good or services included in the transaction, but these pieces of information are not available for every records, thus we left them out of the modeling process. This dataset contains more than 2 millions

TABLE II: Features describing a transaction between two companies.

Feature	Type	Description
Client	Nominal (ID)	Identification number of the client issuing the transaction.
Supplier	Nominal (ID)	Identification number of the supplier receiving the transaction.
Account	Nominal (ID)	Identification number of the bank account to which the money is transferred.
Date	Continous (Timestamp)	Timestamp indicating the date when the transaction took place.

transactions.
We split this dataset in order to fit the model of each client: first each of the transactions where the client is found are selected, then all the transactions involving the suppliers found in this subset of transactions are also added to the dataset. This dataset is then used to fit our model. For the test case of our representative client, 16,168 transactions involving the representative client are first selected, then 423,464 transactions are added from the suppliers, thus aggregating a dataset of 439,632 transactions.

B. SiS-Id Expert System

The SiS-id expert system is a rule-based system that relies on the expert knowledge of SiS-id's fraud investigation team. The expert investigators identified a set of fraudulent and legitimate patterns corresponding to potential fraud or legitimate transactions. While the exact inner workings of the expert system are confidential, we consider its results as the expert opinion of the investigation team about the legitimacy of a specific transaction. When tasked to label a transaction, the expert system analyzes the known patterns of transactions and then outputs a label: *high* means that the transaction fits a legitimate pattern found in the rule base, *low* means that the transaction fits a fraudulent pattern found in the rule base, and *medium* indicate that the transaction does not fit any of the patterns found in the rule base and thus the expert system is unable to give a meaningful estimation of the transaction's legitimacy.

C. Audit Dataset

A second set of transactions is provided by SiS-id. It consists in the list of transactions that were analyzed using the expert system of the company in the past 2 years (July 2017 - July 2019). The dataset, called the "Audit" dataset, is composed of 108,102 suspicious transactions submitted by 171 unique client companies. These transactions are attributed a legitimacy label by SiS-id's fraud detection plateform that we use as ground truth for our analysis. In order to perform the evaluation of our system, we use the 86 clients found in both the History dataset and the Audit dataset. For the client selected in our test case, 251 transactions were labeled by the expert systems, with the following label distribution: 97 were labeled with the *high* legitimacy label, 99 with the *medium* legitimacy label, and 55 with the *low* legitimacy label.

D. Precision

First and foremost, the most important metric in order to assert the performance of the two approaches of ProbaSIF is the capacity to detect frauds. In order to evaluate this capacity in the context of the representative client, we use the 55 transactions labeled with the *low* label as our ground truth, and evaluate the result of the classification algorithms based on the fitting previously performed using the historical transactions. Table III shows the confusion matrix of each of the model's classification results. We focus on the rightmost columns in order to evaluate the accuracy of ProbaSIF two approaches.
Table IIIa shows the confusion matrix for the intra-company analysis: 51 of the 55 transactions are correctly labeled with the *low* label by this approach, thus the precision of the intra-company analysis is $Pre^{cli} = \frac{51}{55} = 0.927$.
Table IIIb shows the confusion matrix for the inter-company analysis: 46 of the 55 transactions are correctly labeled with the *low* label by this approach, thus the precision of the inter-company analysis is $Pre^{com} = \frac{46}{55} = 0.836$.

Figure 3 shows a Venn diagram representing the index of the transaction attributed a *low* legitimacy label by each of the classificaton systems. This figure shows us that the 46 transactions with a *low* legitimacy label found by the inter-company analysis were also found by the inter-company analysis. It is interesting to notice that a large number of labeled transactions (136 of 162 and 141 transactions for the intra-company and inter-company respectively) are similarly labeled by the two ProbaSIF approaches. This fact might be explained by the fact that they share a similar perspective for their fraud detection.

Finally, after discussion with SiS-id investigation team and a careful analysis of the expert system, it appears that the 4 transactions missed by the intra-company analysis where given a *low* legitimacy label by the expert system due to the fact that the account identification number found in the transaction was misspelled. As spell-checking is not a part of our model, it is understandable that these four transactions have been missed.

E. Consistency

The second most important question that we ask is: how consistent the classifications made by ProbaSIF's two algorithms are with the expert knowledge ?
In order to answer this question, we first analyze the Venn diagram shown in Figure 4a, that shows the transactions assigned with the *medium* label by each of the classification systems. While it is clear that ProbaSIF two models have labeled less transactions with the *medium* label (1 for the intra-company analysis and 5 for the inter-company analysis), the most surprising result is that none of these transactions as been given the *medium* label by the expert system.
This result might be explained by the fact that the *medium* label doesn't mean the same thing for the expert system

TABLE III: Confusion Matrix between SiS Rule Engine's results and ProbaSIF two approaches' results for $\delta 1 = 0.50$ and $\delta 2 = 0.90$.

(a) Confusion Matrix: intra-company analysis

Expert System → intra-Client analysis ↓	High	Medium	Low
High	**41**	39	4
Medium	1	**0**	0
Low	51	60	**51**
Total	97	99	55

(b) Confusion Matrix: inter-company analysis

Expert System → inter-Client analysis ↓	High	Medium	Low
High	**60**	39	6
Medium	2	**0**	3
Low	35	60	**46**
Total	97	99	55

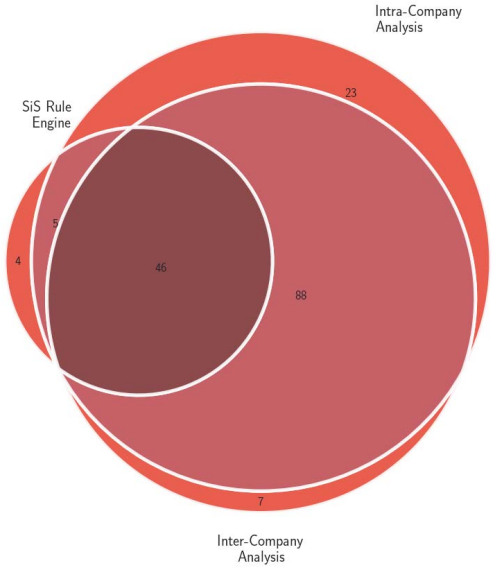


Fig. 3: Overlap of results from the Rule-based System, intra-company analysis and inter-company analysis for $\delta 1 = 0.50$ and $\delta 2 = 0.90$ (fraudulent transactions).

and ProbaSIF: while it indicates that an account's occurrence probability lies in the range $[\delta 1, \delta 2]$ for ProbaSIF, it means for the expert system that the transaction doesn't fit any identifying patterns, thus representing a lack of knowledge rather than an actual assertion of legitimacy.

We can see using Table III that for both of ProbaSIF's classification systems, 39 of the 99 *medium* label transactions have been labeled with a *high* label by ProbaSIF approaches, while 60 of the 99 have been labeled with a *low* legitimacy label. The fact that these results are similar might be explained by the overall similarity in the decision of the two ProbaSIF's algorthims.

Then, we examine the Venn diagram representing the transactions assigned a *high* legitimacy score by each of the classification systems. We see here that the classificaton systems did not agree on almost half of the transactions (43 of the 97 labeled as *high* by the expert system). Moreover, Table III shows that a significant number of transactions labeled as *high* by the expert system were assigned a *low* label (51 for the intra-company analysis and 35 for the inter-company analysis), thus indicating a lot of false positive, indicated by a False Positive Rate (FPR) relatively high for both of ProbaSIF's algorithm: $FPR^{cli} = \frac{51}{97} = 0.526$ and $FPR^{com} = \frac{35}{97} = 0.361$. However, this high FPR might be explained by the fact that our risk threshold $\delta 2$ was arbitrarily set to a high (0.90) value, and thus that the probability needed to be labeled with the *high* label by both the inter-company analysis and the intra-company analysis was too high to be realistically achievable. An other issue arising with the use of user-defined threshold is the fact that they directly impact the number of transactions with the "medium" label. The analysis of the impact of the risk thresholds on the consistency of the model is currently under investigation.

F. Adaptability

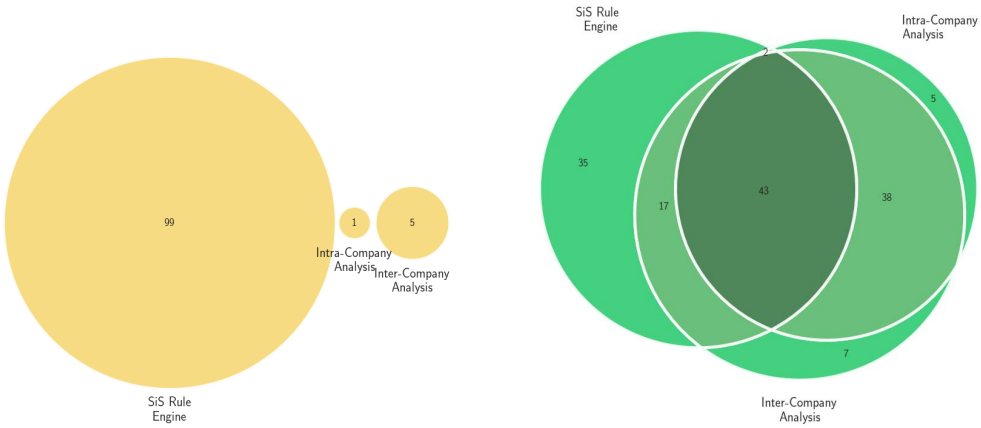
Then, we investigate how well the ProbaSIF model and classification algorithms perform in different context. Figure 5 shows the results in terms of accuracy for ProbaSIF's two approaches when used with the other clients founds in the B2B ecosystem (the clients' company identification numbers have been removed to protect their anonymity). In order to evaluate the performance of the inter-company and intra-company analysis on these different clients, an arbitrary precision threshold of 0.70 has been defined, and a performance under this threshold means that ProbaSIF's approaches are not suited for this client. Our results shows that only 12 of the 73 clients do not meet the objectives, and thus ProbaSIF is usable for 81% of the considered clients.

Additionally, the efficiency in terms of classification time is also shown in Figure 5 : it usually takes around 570 ms on a laptop with 7,4 GiB of RAM to classify a transaction using ProbaSIF. None of the experiments made for each of the clients goes above the threshold of 700 ms.

VI. Conclusion

In this chapter, we introduced ProbaSIF, a SIF detection system based on Bayesian models that uses historical transactions of a client company to compute the probability of occurrence of a specific account in a transaction with a supplier company. A label is assigned to a new transaction based on the probability for the account involved in the transaction to be used to pay the supplier.

We described the two main algorithms composing ProbaSIF, each linked to a specific probability distribution fit by the transactions found in a dataset provided by SiS-id. Firstly,



(a) Overlap of *medium* labels

(b) Overlap of *high* labels

Fig. 4: Overlap of results from the Rule-based System, intra-company analysis and inter-company analysis for $\delta 1 = 0.50$ and $\delta 2 = 0.90$ (undetermined transactions).

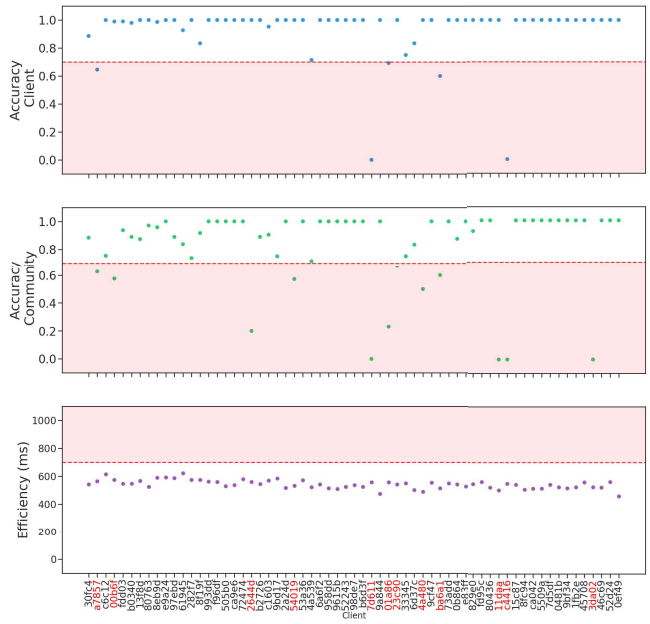


Fig. 5: Global results of ProbaSIF systems for $\delta 1 = 0.50$ and $\delta 2 = 0.90$.

we introduce the intra-company analysis that focuses on the probability of an account to be used to pay a supplier, knowing that a specific client performed the transaction. This probability distribution reflects the narrow vision of a client, limited to the knowledge of its own transactions when dealing with a supplier in the B2B ecosystem. Secondly, we describe the inter-company analysis that does not take into account the client issuing the transaction and focuses on the accounts used to pay the supplier by the all the clients of the ecosystem. This approach aims to model a more general view the supplier's transaction behavior in order to detect a possible discrepancy. We then described a classification algorithm that use both of these probability distribution to assign a label to a new transaction.

We presented the result of ProbaSIF first on a single client to investigate its performance locally, and we then generalized it to the other clients of our B2B ecosystem in order to investigate its global performances. Results shows that locally, most of the low legitimacy labels assigned by both of ProbaSIF's approach are shared by the rule engine, meaning that its results are consistent with expert knowledge. Furthermore, ProbaSIF shows very good accuracy in detecting fraudulent transactions (0.927 and 0.836 for the intra-company and inter-company approaches respectively), and that its time to detection is close to real-time (570 ms on average).

The global evaluation of ProbaSIF shows that the approach leads to very good adaptability of the fraud detection system, meaning that it can be used with a minimum tuning on a large set of clients with heterogeneous behavior. However, as it relies on the set of historical transactions in order to compute the underlying probability distributions used for classification, it is clear that when a low number of such historical transactions is available, the Law of Large Number will not apply [18]. However, we discussed the fact that this outcome might in fact be beneficial for our system, as less historical transactions means a more clear-cut decision from the classification systems, which might be a desired outcome in order to ward off fraudulent transactions in an unknown environment.

References

[1] Aisha Abdallah, Mohd Aizaini Maarof, and Anazida Zainal. "Fraud detection system: A survey". In: *Jour-*

nal of Network and Computer Applications 68 (2016), pp. 90–113. ISSN: 10958592. DOI: 10.1016/j.jnca.2016.04.007.

[2] AIG. *Impersonation Fraud Claims Scenarios*. https://www.aig.com/content/dam/aig/america-canada/us/documents/business/management-liability/impersonation-fraud-claims-scenarios-brochure.pdf. [Online; accessed December 17, 2019]. 2019.

[3] David M Blei, Michael I Jordan, et al. "Variational inference for Dirichlet process mixtures". In: *Bayesian analysis* 1.1 (2006), pp. 121–143.

[4] Richard J. Bolton et al. "Statistical Fraud Detection: A Review". In: *Statistical Science* 17.3 (2002), pp. 235–255. ISSN: 08834237. DOI: 10.1214/ss/1042727940.

[5] Stephen P Brooks, Edward A Catchpole, Byron JT Morgan, et al. "Bayesian animal survival estimation". In: *Statistical Science* 15.4 (2000), pp. 357–376.

[6] Michael H. Cahill et al. "Detecting Fraud in the Real World". In: *Computing Reviews* 45.7 (2013), pp. 911–929. ISSN: 0010-4884. DOI: 10.1007/978-1-4615-0005-6_26.

[7] Rémi Canillas et al. "Exploratory Study of Privacy Preserving Fraud Detection". In: 2018, pp. 25–31. ISBN: 9781450360166. DOI: 10.1145/3284028.3284032.

[8] Varun Chandola, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey". In: *ACM computing surveys (CSUR)* 41.3 (2009), p. 15.

[9] Alejandro Correa Bahnsen et al. "Feature engineering strategies for credit card fraud detection". In: *Expert Systems with Applications* 51.6 (2016), pp. 134–142. ISSN: 09574174. DOI: 10.1016/j.eswa.2015.12.030. URL: http://dx.doi.org/10.1016/j.eswa.2015.12.030.

[10] D. G.T. Denison et al. "Bayesian partition modelling". In: *Computational Statistics and Data Analysis* 38.4 (2002), pp. 475–485. ISSN: 01679473. DOI: 10.1016/S0167-9473(01)00073-1.

[11] Euler-Hermes DFCG. *Barometre Euler Hermes-DFCG 2019*. https://www.eulerhermes.fr/actualites/etude-fraude-2019.html. [Online; accessed December 17, 2019]. 2019.

[12] Mohamed Ahmed Galal, Wessam M Hussein, Mahmoud MA Sayed, et al. "Satellite battery fault detection using Naïve Bayesian classifier". In: *2019 IEEE Aerospace Conference*. IEEE. 2019, pp. 1–11.

[13] Andrew Gelman and Cosma Rohilla Shalizi. "Philosophy and the practice of Bayesian statistics". In: *British Journal of Mathematical and Statistical Psychology* 66.1 (2013), pp. 8–38.

[14] Alon Halevy, Peter Norvig, and Fernando Pereira. "The unreasonable effectiveness of data". In: (2009).

[15] Bryan Hooi et al. "BIRDNEST: Bayesian inference for ratings-fraud detection". In: *16th SIAM International Conference on Data Mining 2016, SDM 2016* (2016), pp. 495–503. arXiv: arXiv:1511.06030v2.

[16] Irina Rish et al. "An empirical study of the naive Bayes classifier". In: *IJCAI 2001 workshop on empirical methods in artificial intelligence*. Vol. 3. 22. 2001, pp. 41–46.

[17] Ethan Roberts, Bruce A. Bassett, and Michelle Lochner. "Bayesian Anomaly Detection and Classification". In: (2019). arXiv: 1902.08627. URL: http://arxiv.org/abs/1902.08627.

[18] Howard Wainer. "The most dangerous equation". In: *American Scientist* 95.3 (2007), p. 249.