Uncovering Document Fraud in Maritime Freight Transport Based on Probabilistic Classification

Ron Triepels[1]([✉]), Ad Feelders[2], and Hennie Daniels[1]

[1] CentER, Tilburg University, Warandelaan 2, Tilburg, The Netherlands
{r.j.m.a.triepels,h.a.m.daniels}@uvt.nl
[2] Department of Information and Computing Sciences, Utrecht University,
Princetonplein 5, Utrecht, The Netherlands
a.j.feelders@uu.nl

Abstract. Deficient visibility in global supply chains causes significant risks for the customs brokerage practices of freight forwarders. One of the risks that freight forwarders face is that shipping documentation might contain document fraud and is used to declare a shipment. Traditional risk controls are ineffective in this regard since the creation of shipping documentation is uncontrollable by freight forwarders. In this paper, we propose a data mining approach that freight forwarders can use to detect document fraud from supply chain data. More specifically, we learn models that predict the presence of goods on an import declaration based on other declared goods and the trajectory of the shipment. Decision rules are used to produce miscoding alerts and smuggling alerts. Experimental tests show that our approach outperforms the traditional audit strategy in which random declarations are selected for further investigation.

Keywords: Data mining · Fraud detection · Freight forwarding · Global supply chains

1 Introduction

International trade is going through an impressive growth. Eurostat estimated the total value of goods imported by the member states of the European Union over 2014 at 1.6 trillion euros, which constitutes an increase of almost 64 percent compared to 2004 [8]. Increased trade is an indication of improved economic integration and world prosperity, but at the same time suggests that fraud is more likely to happen with more serious consequences. Particularly, freight forwarders face problems when trade increases. The focal position of freight forwarders in global supply chains restricts their visibility and control on shipment documentation while budget and manpower for customs brokerage are also limited. Freight forwarders must blindly trust externally generated shipping documents and declare goods they usually do not even see [12]. Without clear overview

or effective risk controls in place, fraudulent declarations easily vanish in the extensive volume that needs to be processed.

The rise of information technology (IT) proposes a shift towards automation, risk management and intelligence in the customs brokerage practices of freight forwarders [11]. IT allows more detailed supply chain data to be recorded. Take radio frequency identification (RFID) as an example. The small size of the RFID tags and their low production costs makes them useful for tracking and tracing international cargo flows [2] and reducing delays at customs clearance locations [14]. At the same time, IT also provides tools to share data among supply chain participants in a fast and reliable way. Several technologies have been proposed to connect shippers and freight forwarders in global supply chains, like electronic data interchange (EDI) [17]. The rich supply chain data that freight forwarders are able to collect can be analyzed to manage customs brokerage risks.

In this paper, we present a data mining approach to detect miscoding and smuggling from supply chain data. Our approach can be applied by freight forwarders for signaling and internal auditing purposes. Customs agents can consult our data mining models during customs brokerage to determine if the documentation of a shipment has high potential to involve fraud. Potentially fraudulent declarations can thereby be prevented from being sent to customs. In addition, our data mining models can improve existing auditing and risk management procedures. Usually, freight forwarders take random samples of declarations and audit these declarations to ensure procedural compliance. Instead of taking random samples, our data mining models can be consulted to specifically audit only those declarations that likely involve miscoding or smuggling.

In summary, the contributions of this paper are as follows:

1. We define decision rules for miscoding alarms and smugling alarms.
2. We present a data mining approach to automatically produce alarms for miscoding and smuggling from supply chain data.
3. We conduct a comparative study on the performance of different classification models.
4. We demonstrate that our data mining approach performs significantly better at detecting miscoding and smuggling than the random audit strategy.

2   Background

2.1   Document Fraud in International Shipping

One of the oldest and most common fraud schemes in international shipping is document fraud. Document fraud is the act of manipulating facts in contracts or agreements with the intent to benefit by commercial gain [13]. International logistics is an attractive target for document fraud because it heavily relies on the exchange of formal shipping agreements. Although shipment agreements are issued and checked by different actors in a supply chain, the exchange of agreements takes the form of a serial connection. Thus, fraudulent agreements at the beginning of a supply chain are simply adopted by others and are hard to detect

at the end [12]. According to the World Customs Organization, the incentive to commit document fraud in international shipping is to evade customs duties and tax payments, or to circumvent shipping restrictions and sanctions [21].

2.2   Types of Document Fraud

Two types of document fraud are commonly committed in international shipping documentation:

Miscoding is the act of providing incorrect or incomplete information about the nature of goods that are being shipped. To keep track of the variety of goods, the World Customs Organization introduced the harmonized system (HS). The HS is an international system of codes to classify goods. Fraudsters commonly specify HS-codes of other goods with similar properties but which are not prohibited or require to pay lower customs duties.

Smuggling is the act of secretly shipping goods under conditions that are against the law by any of the countries that are crossed by a shipment. Usually, goods are secretly put inside a container somewhere along the supply chain and removed after goods have been cleared at the destination country. Drugs, weapons, cigarettes and alcohol are goods that are frequently smuggled because they are prohibited or require to pay higher customs duties.

2.3   Related Research

Research dedicated to the task of detecting document fraud from supply chain data uses supervised and unsupervised learning techniques. Supervised learning techniques assume that data about fraudulent cases is available a priori which can be used to predict fraud in future observations. Bayesian classifiers have been applied to detect miscoding of HS-codes [9]. The researchers built an hierarchical Bayesian model based on values of the consignee, country of origin and destination country that are listed on import declarations. Fraudulent behavior is learned by training the model on a sample of correctly and incorrectly classified goods. A comparable study constructed a classifier based on association rules [22]. The classifier aims to predict the overall risk level of shipments based on inconsistencies in product features like prices and weights.

Unsupervised learning techniques assume that there is no a priori knowledge about document fraud available. These techniques focus on observations that significantly deviate from a statistical norm. Outlier detection is applied to detect deviating product properties of goods marked for clearance [6]. The researchers developed an application that highlights the statistical distributions of product properties in a set of diagrams. Customs agents can use these diagrams to inspect how much goods on an import declaration deviate from others. In addition, ranking may help freight forwarders to prioritize outliers. A related study proposes a method to calculate a numerical ranking for price outliers in trade data [16]. Freight forwarders can use such a ranking to identify risky declarations that require further investigation.

2.4   Hidden Information in Shipment Trajectories

Existing research demonstrates the ability of data mining to detect document fraud from supply chain data. However, proposed applications mainly ignore the trajectory by which shipments find their way through the supply chain network. Including trajectory details in the analysis may improve the detection rate due to two recent changes in international trade. First of all, international trade is moving towards vertical specialization in which each country produces particular goods for the stages of a production sequence [15]. Second, logistic services are subject to optimization by which the trajectory of a shipment is chosen based on a trade-off between economic considerations, like price, flexibility and service level [4] [20]. These changes are expected to create distinct patterns in shipment behavior which can be used to highlight cases of document fraud. Our approach differs from existing research in that we specifically use these patterns in shipment trajectories to predict the goods that should be listed on the corresponding import declarations.

3   Fraud Detection in Supply Chain Data

3.1   Detecting Miscoding and Smuggling

We propose to detect miscoding and smuggling using an unsupervised learning approach. Fraudulent declarations are identified by learning shipping behavior from supply chain data, in terms of probabilities that HS-codes are declared on a trajectory, and determining the extent to which an import declaration deviates from this norm. We introduce some notation to formalize this problem.

Let L be the state space of all possible locations between which goods can be transported. A trajectory is defined as a sequence consisting of a set of random variables $X = \{x_1, \ldots, x_n\}$, where each variable $x_i \in L$ represents a location in a supply chain and n is fixed for all trajectories. Moreover, let G be a set of indicators of HS-codes, that is, $G_i = 1$ if the good with HS-code i is listed on the import declaration, and $G_i = 0$ otherwise.

We learn a model from data that estimates the probability of goods on an import declaration given all other declared goods and the trajectory of the shipment. Predictions of goods are compared with the goods listed on the import declaration to determine if the declaration involves miscoding or smuggling. An alarm is produced for miscoding when a good is declared but the probability of this good estimated by model M is very low:

$$G_i = 1 \wedge P_M(G_i = 1|G_{j\neq i}, X) \leq \alpha \tag{1}$$

Here, $G_{j\neq i}$ represents all goods other than the good that is currently evaluated on the import declaration and α represents a threshold. We define the alarm for smuggling in a similar fashion. An alarm for smuggling is produced when a good has not been declared but the probability of this good estimated by model M is very high:

$$G_i = 0 \wedge P_M(G_i = 1|G_{j\neq i}, X) \geq (1 - \alpha) \tag{2}$$

3.2 Construction of the Classification Models

We built probabilistic classification models to estimate the conditional probabilities in equation 1 and 2. Our modelling approach is similar to the one proposed in [23] for the construction of Bayesian chain classifiers, except that, unlike our model, Bayesian chain classifiers are intended for multi-label classification. For that reason the authors of [23] need to impose an order on the labels, where predicted values of labels that are prior in the order are used to predict subsequent labels. They use a Bayesian network to define the order, and use the parents of a node as predictors for the child node. Unlike in the multi-label classification scenario, we observe all labels (declared HS-codes) during prediction, and can use the Markov blanket instead of just the parent set. The construction of our models can be broken down into two main steps.

In the first step we define associations between goods in a shipment. We do this by learning an undirected graphical model (Markov random field) from a data table where each row lists the goods that are contained in a shipment, that is, $G_i = 1$ if good i is contained in the shipment, and $G_i = 0$ otherwise. To avoid overfitting, we score candidate models using Akaike information criterion (AIC) which penalizes model complexity. Unfortunately, finding the model with the best AIC score from data is NP-hard [3]. To find a good (locally optimal) model, we perform a hill-climbing search with forward selection. Still, it can take a considerable amount of time to find such a model when the algorithm needs to start from the mutual independence model (empty graph). Therefore, we first generate a minimal AIC forest [7] and then use this simplified model as initial model to speed up the hill climbing.

In the second step we evaluate two types of probabilistic classifiers: Naive Bayes (NB) and Tree-Augmented Naive Bayes (TAN). NB is a simple classifier that uses Bayes rule of conditional probability to compute the probability of a label given a set of features. This computation is simplified under the assumption that all features are independent of each other given the label. A NB classifier is constructed for each good by determining its Markov blanket in the undirected graphical model (Fig. 1A), and estimating the probability of the good given the goods in its Markov blanket and locations in the shipping trajectory. The resulting classifier is depicted as a Bayesian network in (Fig. 1B). We incorporate only dependent goods as additional features in the classification models to avoid features that have poor prediction power with respect to the target good. Therefore, term $G_{j \neq i}$ in equation 1 and 2 is substituted by $MB(G_i)$. Here, $MB(G_i)$ refers to the Markov blanket of good G_i in the undirected graphical model.

The main drawback of a NB classifier is that it fails to capture the dependencies between features due to its strong independence assumption. A TAN relaxes this assumption by specifying a tree structure on the feature set in which each feature only has as parent the label and at most one other feature. We use Chow and Liu's algorithm [5] to automatically generate TAN classifiers from data. The structure of a TAN is generated by creating a weighted spanning tree with mutual information weights. This spanning tree is transformed to a directed tree by picking a root node and pointing all edges away from this root. Finally, the
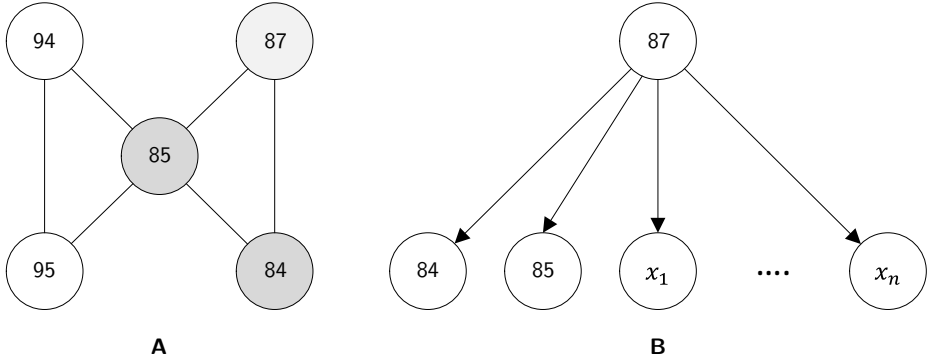
Fig. 1. An example graphical model with dependencies between HS-codes (A) and the corresponding Naive Bayes classifier for HS-code 87 (B).

structure of the TAN is obtained by making all nodes in the directed tree dependent on the label. The advantage of modeling features in a tree structure is that it allows to make an optimal trade off between the ability to incorporate feature dependencies and the corresponding computational complexity [10].

4   Performance Evaluation

4.1   Sample Data

We extracted a large sample from the supply chain repository of an international freight forwarder. The sample contains data of 11,044 maritime shipments that were shipped to the Netherlands in the period of April 2012 and June 2013. Each shipment has data about the bill of lading that was issued by the sea carrier and the import declaration that was send to the Dutch customs authorities. Bills of lading related to the inland transportation were not available. For our application we focus on six features in the sample (Table 1).

From the bill of lading we extracted the bill of lading number (BLN) and four trajectory locations, the: origin (ORG), port of loading (POL), port of discharge (POD) and destination (DES). The origin represents the location where goods are produced. The port of loading is the first port in the shipment trajectory and the place where goods are loaded and secured on a cargo ship. The port of discharge is the last port in the shipment trajectory and the place where goods are unloaded for transloading. This is also the place where the goods in our sample are cleared by the Dutch customs authorities. Finally, the destination represents the location of the consignee. From the import declaration we extracted the HS-codes (HSC) that are declared for the shipments.

Table 1. Three example shipments in the sample data that transported goods to the port of Rotterdam and Antwerp. The goods are transshipped from these ports to consignees located in Amsterdam and Tilburg. The bill of lading numbers are anonymized for confidentiality reasons.

BLN	HSC	ORG	POL	POD	DES
01234567890123	69	Shenzhen	Yantian	Rotterdam	Amsterdam
01234567890123	94	Shenzhen	Yantian	Rotterdam	Amsterdam
01234567890123	39	Chicago	Montreal	Antwerp	Amsterdam
01234567890123	83	Delhi	Mundra	Rotterdam	Tilburg
01234567890123	94	Delhi	Mundra	Rotterdam	Tilburg
01234567890123	76	Delhi	Mundra	Rotterdam	Tilburg

4.2   Data Pre-processing

After collecting supply chain data, we performed pre-processing to transform the sample to the right format for analysis. We aggregated the shipments based on their bill of lading number and transformed the HS-codes to individual binary features. Then, we performed filter operations to ensure that we would be able to take stratified samples to partition the sample into training sets and test sets with equally distributed shipping trajectories. We used the following filters on the sample: each shipping trajectory must have been taken by at least 3 shipments, and each HS-code must have been declared for at least 15 shipments. Finally, we extracted the first two digits of the HS-codes. This was necessary given the small size of the sample. There are not enough observations to learn four digit or six digit HS-codes without over-fitting the classification models. The sample after pre-processing includes data of 10,154 shipments that transported 50 types of goods over 625 shipment trajectories. The bnlearn package [19] in R is used to construct the NB and TAN classifiers on this sample.

4.3   Evaluation Procedure

We evaluated the performance of the classification models by generating artificial declaration errors and counting how many times the models produced a correct alarm. The sample was first divided in a training set and test set based on a 75-25 percent split ratio. Then, artificial declaration errors were generated for ten percent of the declarations in the test set. Miscoding errors were generated by determining the joint Markov blanket of all goods listed on a declaration and adding a random good from this blanket to the declaration. If all goods from this blanket were already listed on the declaration, then a random good outside the blanket was added. Smuggling errors were generated by randomly removing a good from the declaration.

The accuracy of the miscoding alerts and smuggling alerts are reported by calculating corresponding confusion matrices. From these confusion matrices, we

derived three additional performance measures: precision, recall and F_1. Precision and recall are defined as [18]:

$$Precision = \frac{TP}{TP + TN} \qquad Recall = \frac{TP}{TP + FP} \qquad (3)$$

Here, TP denotes true positives, TN denotes true negatives and FP denotes false positives. Precision measures the fraction of alarms in which declarations indeed contain miscoding or smuggling. Recall measures the fraction of fraudulent declarations for which an alarm is produced. These two measures are combined to form the F_1 score [18]:

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \qquad (4)$$

The F_1 score constitutes the harmonic mean of precision and recall. It provides a single measure to score competing classification models. We compare the F_1 scores of our modes and the audit strategies and determine the model or strategy that scores the highest F_1 score.

The evaluation procedure of our models can be summarized as follows:

1. *Generate training set and test set* - take a stratified sample and generate a training set and test set based on a 75-25 percent split ratio.
2. *Construct the classification model* - construct a classification model on the training set. Estimate the parameters of the classifiers using Laplace smoothing [10].
3. *Generate miscoding and smuggling errors* - generate two separate test sets, one having a smuggling error in 10 percent of the declarations, and the other having a miscoding error in 10 percent of the declarations.
4. *Predict miscoding and smuggling* - use the classification model to predict whether the declarations in the test sets involve miscoding or smuggling.
5. *Calculate confusion matrices* - use the predictions to calculate corresponding confusion matrices for the miscoding and smuggling alarms.

5   Results

The results of the performance evaluations of the NB and TAN models are shown in the confusion matrices in Table 2. Alarms are produced at an alpha value of $\alpha = 0.05$. This means that the models must be: at most 5 percent certain that a good should be declared to produce a miscoding alert, and at least 95 percent certain that a good should be declared to produce a smuggling alert. Given these criteria, we performed 25 iterations of the evaluation procedure and calculated the mean prediction rates for the confusion matrices.

We compare the performance of our models with the post audit strategy that suggests to audit the same number of declarations as our models but choose declarations at random. We denote these strategies RNB and RTAN. The confusion

matrices for the strategies are shown in Table 3. In this table, prediction rates of RNB for miscoding are calculated as follows. Let A be the probability of choosing a declaration for investigation according the prediction rate of the NB model, and E the probability that a declaration contains a miscoding error. Given Table 2, the values for A and E are:

$$A = 0.0886 + 0.1297 = 0.2183 \quad \neg A = 1 - (0.0886 + 0.1297) = 0.7817 \quad (5)$$
$$E = 0.0886 + 0.0112 = 0.0998 \quad \neg E = 1 - (0.0886 + 0.0112) = 0.9002 \quad (6)$$

Because A and E are independent of each other, the confusion matrix can be calculated by multiplying the following combinations:

$$TP = A \cdot E \qquad TN = A \cdot \neg E \qquad FP = \neg A \cdot E \qquad FN = \neg A \cdot \neg E \qquad (7)$$

Here, TP denotes true positives, TN denotes true negatives, FP denotes false positives and FN denotes false negatives. After calculating the confusion matrices for the models and audit strategies, we derived their precision, recall and F_1 score and compared their scores, see Table 4.

Table 2. Confusion matrices for the alarms produced by the NB and TAN models with an alpha value of $\alpha = 0.05$. Rows indicate whether an alarm goes off, columns indicate whether something is actually wrong with the declaration.

	Miscoding				Smuggling			
	NB		TAN		NB		TAN	
	True	False	True	False	True	False	True	False
True	0.0886	0.1297	0.0944	0.0938	0.0441	0.1376	0.0500	0.0410
False	0.0112	0.7705	0.0054	0.8063	0.0558	0.7626	0.0498	0.8592

Table 3. Confusion matrices for the alarms produced by the RNB and RTAN audit strategies. Rows indicate whether an alarm goes off, columns indicate whether something is actually wrong with the declaration.

	Miscoding				Smuggling			
	RNB		RTAN		RNB		RTAN	
	True	False	True	False	True	False	True	False
True	0.0218	0.1965	0.0188	0.1695	0.0181	0.1635	0.0091	0.0819
False	0.0780	0.7036	0.0810	0.7307	0.0817	0.7367	0.0907	0.8183

Table 4. Precision, recall and F_1 scores for the NB and TAN models and their corresponding random audit strategies denoted as RNB and RTAN.

	Miscoding				Smuggling			
	NB	RNB	TAN	RTAN	NB	RNB	TAN	RTAN
Precision	0.4059	0.0998	0.5015	0.0998	0.2426	0.0998	0.5494	0.0998
Recall	0.8877	0.2183	0.9457	0.1882	0.4415	0.1816	0.5007	0.0910
F_1	0.5570	0.1370	0.6554	0.1305	0.3131	0.1288	0.5239	0.0952

6   Discussion

The F_1 scores show that the TAN model outperforms the other models and the random audit strategies on both miscoding and smuggling. The model produced close to 19 percent miscoding alarms and 9 percent smuggling alarms for the declarations in the test set. Miscoding alerts are produced with an average precision of 50 percent and a recall of 95 percent. This means that half of the time the model produced an incorrect miscoding alarm although almost each declaration with miscoding is detected. Smuggling seems more difficult to detect. Alerts for smuggling are produced with an average precision of 55 percent and a recall of 50 percent. Or in other words, the model produced less than half of the time an incorrect smuggling alarm and detected only half of the total declarations in the test set that contained smuggling.

Alarms produced by the NB model are of lower quality. The model produced more alarms for miscoding and smuggling, respectively 22 percent and 18 percent, while the precision and recall of these alarms are lower. Miscoding alerts are produced with an average precision of 41 percent and a recall of 89 percent, while smuggling alerts are produced with an average precision of 24 percent and recall of 44 percent. The NB model suggests to audit more declarations while it captures less fraudulent declarations than the TAN model. This inaccuracy can be attributed to the strong independence assumption of the NB model, which assumes that the goods on an import declaration and the locations in the shipping trajectories are independent of each other given the good that is being predicted, while there clearly seems to exist interaction between these features.

Although the TAN model also does not have an exceptionally high precision and recall, it performs substantially better than the corresponding random audit strategy. Using the random audit strategy leads in our experiment to anti-fraud investigations that are only worthwhile for 10 percent of the suspected declarations. Furthermore, the strategy is only able to detect 19 percent of the total miscoding errors and 9 percent of the total smuggling errors. Therefore, we argue that the use of our TAN model is a substantial improvement over the random audit strategy.

7   Conclusion

In this paper, we have proposed a data mining approach that freight forwarders can apply to detect miscoding and smuggling. We performed an experiment to determine the extent to which these types of document fraud can be automatically detected from supply chain data. The experiment shows that our approach outperforms the audit strategy in which the same number of declarations would be audited but declarations are chosen at random. Therefore, we conclude that our approach is a substantial improvement over this audit strategy.

We learned models from supply chain data that predict the presence of a HS-code on an import declaration based on other declared HS-codes and the trajectory of the corresponding shipment. Decision rules are defined to produce alarms for miscoding and smuggling when declared goods deviate too much from the expected norm. We proposed two different classification models for this task and evaluated their performance. Performance is measured by generating artificial declaration errors and counting how often the models produced a correct alarm. The classification model based on TAN achieved the highest F_1 score.

8   Future Research

Our approach to detect miscoding and smuggling from supply chain data could be further improved. Trajectories in our models have a fixed length and therefore ignore stops at intermediate ports or warehouses. Intermediate stops in a trajectory may reflect critical patterns of document fraud that, when including them in our classification models, may improve the quality of the miscoding and smuggling alarms. The decision rules that raise these alarms should thereby also consider prohibited goods and incorporate the magnitude of the potential fraud in terms of costs. Prohibited goods usually do not show up in the import declarations and therefore remain undetectable by our models. In addition, including costs is important because businesses need to find a balance between the costs to undertake anti-fraud investigation and the savings that can potentially be achieved [1]. We leave these issues open for future research.

References

1. Bolton, R.J., Hand, D.J.: Statistical fraud detection: A review. Statistical Science, 235–249 (2002)
2. Chang, Y.S., Son, M.G., Oh, C.H.: Design and implementation of rfid based air-cargo monitoring system. Advanced Engineering Informatics **25**(1), 41–52 (2011)
3. Chickering, D.M.: Learning bayesian networks is np-complete. In: Learning from-data, pp. 121–130. Springer (1996)
4. Choi, T.Y., Hartley, J.L.: An exploration of supplier selection practices across the supply chain. Journal of Operations Management **14**(4), 333–343 (1996)
5. Chow, C., Liu, C.: Approximating discrete probability distributions with dependence trees. IEEE Transactions on Information Theory **14**(3), 462–467 (1968)

6. Digiampietri, L.A., Roman, N.T., Meira, L.A., Ferreira, C.D., Kondo, A.A., Constantino, E.R., Rezende, R.C., Brandao, B.C., Ribeiro, H.S., Carolino, P.K., et al.: Uses of artificial intelligence in the brazilian customs fraud detection system. In: Proceedings of the 2008 International Conference on Digital Government Research, pp. 181–187. Digital Government Society of North America (2008)

7. Edwards, D., De Abreu, G.C., Labouriau, R.: Selecting high-dimensional mixed graphical models using minimal aic or bic forests. BMC Bioinformatics **11**(1), 18 (2010)

8. Eurostat: EU trade since 1995 by HS6 (2015)

9. Filho, J., Wainer, J.: Using a hierarchical bayesian model to handle high cardinality attributes with relevant interactions in a classification problem. In: IJCAI, pp. 2504–2509 (2007)

10. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian network classifiers. Machine Learning **29**(2–3), 131–163 (1997)

11. Gordhan, P.: Customs in the 21st century. World Customs Journal **1**(1), 49–54 (2007)

12. Hesketh, D.: Weaknesses in the supply chain: who packed the box. World Customs Journal **4**(2), 3–20 (2010)

13. Hill, H.: Nolo's Plain-English Law Dictionary. Nolo (2009)

14. Hsu, C.I., Shih, H.H., Wang, W.C.: Applying rfid to reduce delay in import cargo customs clearance process. Computers & Industrial Engineering **57**(2), 506–519 (2009)

15. Hummels, D., Ishii, J., Yi, K.M.: The nature and growth of vertical specialization in world trade. Journal of International Economics **54**(1), 75–96 (2001)

16. Kopustinskas, V., Arsenis, S.: Risk analysis approaches to rank outliers in trade data. In: Advanced Statistical Methods for the Analysis of Large Data-Sets, pp. 137–144. Springer (2012)

17. Murphy, P.R., Daley, J.M.: Edi benefits and barriers: comparing international freight forwarders and their customers. International Journal of Physical Distribution & Logistics Management **29**(3), 207–217 (1999)

18. Olson, D.L., Delen, D.: Advanced data mining techniques. Springer Science & Business Media (2008)

19. Scutari, M.: Learning bayesian networks with the bnlearn r package. arXiv preprint arXiv:0908.3817 (2009)

20. Tongzon, J.L.: Port choice and freight forwarders. Transportation Research Part E: Logistics and Transportation Review **45**(1), 186–195 (2009)

21. World Customs Organization: Illicit trade report (2012). http://www.wcoomd.org/en/topics/enforcement-and-compliance/~/media/WCO/Public/Global/PDF/Topics/Enforcement%20and%20Compliance/Activities%20and%20Programmes/Illicit%20Trade%20Report%202012/WCO%20REPORT%202013%20-%20BR.ashx

22. Yaqin, W., Yuming, S.: Classification model based on association rules in customs risk management application. In: 2010 International Conference on Intelligent System Design and Engineering Application (ISDEA), vol. 1, pp. 436–439. IEEE (2010)

23. Zaragoza, J.H., Sucar, L.E., Morales, E.F., Bielza, C., Larranaga, P.: Bayesian chain classifiers for multidimensional classification. IJCAI **11**, 2192–2197 (2011)