Food Control 61 (2016) 180–187

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Prediction of food fraud type using data from Rapid Alert System for Food and Feed (RASFF) and Bayesian network modelling

Yamine Bouzembrak, Hans J.P. Marvin*

RIKILT Wageningen UR, Akkermaalsbos 2, 6708 WB Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 8 April 2015
Received in revised form
16 September 2015
Accepted 22 September 2015
Available online 8 October 2015

Keywords:
Bayesian networks
Food fraud type prediction
RASFF

ABSTRACT

Because food fraud can harm human health and erode consumer trust, it is imperative that it is detected at an early stage. Therefore the aim of this study was to predict the expected food fraud type for imported products for which the product category and country of origin are known in order to target enforcement activities. For this purpose we used a Bayesian Network (BN) model that was developed based on adulteration/fraud notifications as reported in the Rapid Alert System for Food and Feed (RASFF) in the period 2000–2013. In this period 749 food fraud notifications were reported and were categorised in 6 different fraud types (i) improper, fraudulent, missing or absent health certificates, (ii) illegal importation, (iii) tampering, (iv) improper, expired, fraudulent or missing common entry documents or import declarations, (v) expiration date, (vi) mislabelling. The data were then used to develop a BN model. The constructed BN model was validated using 88 food fraud notifications reported in RASFF in 2014. The proposed model predicted 80% of food fraud types correctly when food fraud type, country and food category had been reported previously in RASFF. The model predicted 52% of all 88 food fraud types correctly when the country of origin or the product-country combination had not been recorded before in the RASFF database. The presented model can aid the risk manager/controller in border inspection posts in deciding which fraud type to check when importing products.

1. Introduction

The increasing complexity and globalisation of our food supply chains have posed new challenges to ensure the safety and quality of the food products that are available on the market. To cope with this challenge, a wide variety of hazards that may contaminate food need to be controlled by both food producers and authorities. Besides the microbiological, chemical, and physical hazards that enter the food supply adventitiously, an increasing concern is the introduction of hazards by food fraud actions. A historic point in case, for example, was the recent series of incidences involving the deliberate tampering of milk with melamine (Guan et al., 2009; Jia & Jukes, 2013). Food fraud is a collective term for fraudulent behaviour that is driven by economic gain and encompasses the deliberate substitution, addition, tampering, or misrepresentation of food, food ingredients or food packaging, or false or misleading statements made about a product (Spink & Moyer, 2011). Food fraud experts commonly refer to "the food fraud triangle", which is a concept that defines the fraud opportunity using three legs of a triangle (victim, fraudster, absence of a capable guardian). The area within the triangle represents the fraud opportunity which grows as the absence of detection and the number of fraudsters grow. The following examples of factors contributing to an increase in food fraud opportunities were presented in a recent NSF international report (Elliott, 2014): (i) increase of the complexity of supply chain networks, (ii) the rapid development of technology (internet, printing, mobile phone, etc.), (iii) the rapid grow of warehouse systems and refrigerated transport.

In the last 5 years, due to the 2013 horsemeat scandal in Europe, growing attention has been given to food fraud, which refers to activities dedicated to the collection and the analyse of food fraud historical data (Elliott, 2014; Everstine, Spink, & Kennedy, 2013; Tähkäpää, Maijala, Korkeala, & Nevas, 2015). The recent articles by Everstine, Tähkäpää and their co-workers (Everstine et al., 2013; Tähkäpää et al., 2015) encouraged much more research in this field and highlighted the need for creative and innovative methods for preventing and detecting food fraud.

Everstine and co-workers (Everstine et al., 2013) provided an overview of the EMA food fraud database (EMA, 2014), which these

* Corresponding author.
E-mail address: hans.marvin@wur.nl (H.J.P. Marvin).

http://dx.doi.org/10.1016/j.foodcont.2015.09.026

authors used to identify common characteristics of food fraud incidents which either allowed the adulteration to happen or allowed it to be detected. They reviewed the incidents since 1980 listed by the database (EMA, 2014), focussing mainly on incidents that impacted on US consumers. A similar review was done by Moore and co-workers (Moore, Spink, & Lipp, 2012), who reviewed and analysed the USP food fraud database (USP, 2014) to identify trends.

Recently, Elliott (2014) presented a fraud protection model to classify products based on fraud factors. Three factors were used in this model: the potential profit, the potential difficulty/cost and the likelihood of detection. The model can help food industrials and regulators anticipate the relative likelihood of fraudulent attack on many and varied product lines offered to consumers.

In the European Union, as part of the implementation of the Regulation (EC) 178/2002, referred to as “General Food Law”, the Rapid Alert System for Food and Feed (RASFF) has been established to support the control and safety of food and animal feed on the European market. Under RASFF, members (i.e. EU-28 national food safety authorities, Commission, EFSA, ESA, Norway, Liechtenstein, Iceland and Switzerland) exchange information (e.g. notifications) regarding measures taken with regard to food safety, including food fraud. The notifications reported in RASFF are available through the RASFF portal, which features an interactive searchable on line database (RASFF, 2015). The notifications consists of two main types; (i) market notifications and (ii) border rejections. The former being subdivided into alerts, information notifications and news. The RASFF database includes both intentional food frauds such as fraudulent documents or adulteration cases as well as unintentional frauds such as improper, expired or even missing documents. In the period between 2000 and 2013, 749 notifications were reported in RASFF under the hazard category “adulteration/fraud”. This category contains food fraud notifications classified in 6 different fraud types: (i) improper, fraudulent, missing or absent health certificates, (ii) illegal importation, (iii) tampering, (iv) improper, expired, fraudulent or missing common entry documents or import declarations, (v) expiration date, and (vi) mislabelling. The RASFF database has a very different fraud type classification when compared to the US databases (EMA, 2014; USP, 2014). The EMA database proposes 9 types of food fraud, of which the most important are: substitution, artificial enhancement, dilution, transhipment, counterfeit, and misbranding. In the USP database, the following 3 types of food fraud are distinguished based on the definition proposed by Moore and co-workers (Moore et al., 2012): addition, replacement and removal.

In a few studies fraud notifications in RASFF have been analysed, such as by Kleter and co-workers for the period 2003–2007 (Kleter, Prandini, Filippi, & Marvin, 2009) and by Tähkäpää and co-workers for the period 2008–2012 (Tähkäpää et al., 2015). In the period 2003–2007, 248 notifications of fraud related issues were reported in RASFF, the majority pertained to illegal imports and lack of authorization of establishment and of transits. The predominate products were meat (23%), seafood (19%) and composite and mixed products (17%) and the products originated mainly from Asia (44% of the notifications) and EU (22% of the notifications). An increase in the number of reports was noticed after mid-2005 (Kleter et al., 2009). Similar results were observed by Tähkäpää and co-workers (Tähkäpää et al., 2015) in the following period (2008–2012). Fish (and fish products) and meat (and meat products) were the most reported product categories, 16% and 11%, respectively and Asia (45%) the most common region of origin. Although food fraud does not result in a high number of notifications (2%) (Tähkäpää et al., 2015), it has the potential to become real health risks for human and animals and therefore control and mitigation strategies should be developed in order to increase the efficiency of border control.

In this paper we explore the use of Bayesian Networks (BNs) modelling as a tool to assess the food fraud notifications in RASFF and to predict the type of fraud based on the country of origin and product category. BNs are a class of probabilistic models originating from the Bayesian statistics and decision theory combined with graph theory (Bonafede & Giudici, 2007; Nielsen & Jensen, 2007). BNs have the ability to model dependencies between variables, manage non-linear interaction and integrate different kinds of information about the system such as expert knowledge, measurement data, feedback experience and information regarding the system behaviour (Buriticá & Tesfamariam, 2015). BNs have been applied in a number of diverse problem domains such as medical diagnosis (Wiegerinck et al., 1999), image classification (Malka & Lerner, 2004), financial fraud detection (Kirkos, Spathis, & Manolopoulos, 2007; Ngai, Hu, Wong, Chen, & Sun, 2011), nuclear waste disposal (Lee & Lee, 2006) and electrical power systems (Huo, Zhu, Zhang, & Chen, 2004).

In this study, the food fraud notifications in RASFF from 2000 to 2013 were used to develop a BN model for the prediction of food fraud type. The notifications reported in 2014 were used to validate the model.

2. Materials and methods

2.1. RASFF data analysis

All notifications reported in the RASFF database under the hazard category “adulteration/fraud” were extracted from the period 01/01/2000 to 31/12/2013. Each RASFF record contains the following information: Notification ID, date of notification, country notifying, classification (alert, border rejection, information, information for attention, information for follow-up, news), notification type, action taken, distribution status, product, product category, substance/hazard and country concerned (distribution, origin). Based on the description provided in the subject section of the RASFF notifications, we divided the fraud/adulteration notification into 6 different fraud types (see Table 1). All food fraud notifications filed in RASFF during the period analysed (N = 749) were stored in a database for further analysis.

2.2. Bayesian network model

The data file with 749 food fraud records from RASFF was used to build the BN model. Generally, when sufficient data is available (such as in our case), machine learning techniques can be used to develop the model (Alameddine, Cha, & Reckhow, 2011). In this study we used the machine learning technique “expectation-maximization (EM)-algorithm” to construct the BN model and to determine the optimal configuration of the model from a set of all possibilities (Denœux, 2010, 2011). In addition, learning algorithms of a BN aims to compute the conditional probabilities (CPTs) between variables. The learning step was performed using the learning algorithm of the Hugin 8.0 software (http://www.hugin.com/).

2.3. Bayesian network model validation

The food fraud notifications in RASFF reported in 2014 were used to validate the BN model. To this end, these records were retrieved from the RASFF database (88 notifications) and stored in Excel. The validation cases were not included in the learning dataset. The variable “country” and “food category” of each individual RASFF notification were used as input in the constructed BN model to predict the “fraud type” as mentioned in the description of the particular RASFF notification. A score of 1 was given when the fraud type predicted by the BN model was similar (e.g. highest