Food fraud and prediction

Hans Marvin and Yamine Bouzembrak (RIKILT Wageningen UR)

12-14 September, Shanghai, China
Outline

- Food fraud
- Developing a food fraud early warning system using European media monitor (EMM)
- Food fraud prediction
Food fraud, GFSI definition

‘A collective term encompassing the deliberate and intentional substitution, addition, tampering, or misrepresentation of food, food ingredients or food packaging, labelling, production information, or false or misleading statements made about a product for economic gain that could impact consumer health’

Source: Saskia van Ruth (RIKILT), 2015
Identification of fraud: research at RIKILT

- Analytical methods (chemical, physical, microscopy)
  - Targeted approach: single marker
  - Untargeted approach: Multiple markers
  - Fingerprints

- Text mining approach (MedISys)

- Trend analysis (databases (RASFF, EMA, UPS etc) & prediction modelling (agent modelling, Bayesians network)
Identification of fraud: research at RIKILT

- Analytical methods (chemical, physical, microscopy)
  - Targeted approach: single marker
  - Untargeted approach: Multiple markers
  - Fingerprints

- **Text mining approach (MedISys)**

- Trend analysis (databases (RASFF, EMA, UPS etc) & prediction modelling (agent modelling, Bayesians network)
European Media Monitor (EMM)
The Europe Media Monitor (EMM) provides advanced analysis systems for monitoring of both traditional and social media.

- EMM applies text mining techniques to screen different types of media on the world wide web: websites, databases, blogs, ..etc.
- EMM is updated every 10 minutes, 24 hours per day.
- EMM gathers reports from news portals world-wide in 60 languages.
- EMM contain 3 portals: NewsBrief, NewsExplorer and MedISys

(http://emm.newsbrief.eu/overview.html)
MedISys: public health related topics

No collection of publications on food fraud
EMM Food Fraud Filter Design steps

Step 1: Definition of food fraud key words
- Scientific articles
- Food Fraud articles
- USP Database
- EMA Database
- RASFF

Step 2: Validation of the keys words by Food fraud experts
- Prof. Saskia Van Ruth (RIKILT)
- Dr. Hans Marvin (RIKILT)
- Dr. Karen Everstine (USA)

Step 3: The design of the food fraud filter in EMM
- EMM system

Step 4: Evaluation and improvement of the filter.
- Analyse the articles
- Relevance evaluation
- Key words improvement

Figure 1: Design steps
# Food Fraud keywords

<table>
<thead>
<tr>
<th>Generic Key words</th>
<th>Specific key words</th>
<th>Excluded key word</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Food Fraud</td>
<td>• food+fraud</td>
<td>• Fraud</td>
</tr>
<tr>
<td>• Mislabelled</td>
<td>• fake+food</td>
<td>• Artificial Food</td>
</tr>
<tr>
<td>• Counterfeit</td>
<td>• mislabeled+food</td>
<td>• Counterfeit Medicine</td>
</tr>
<tr>
<td>• Tampering</td>
<td>• economic+adulteration +food</td>
<td>• Goods fraud</td>
</tr>
<tr>
<td>• Misleading</td>
<td>• food+adultration</td>
<td>• Wine</td>
</tr>
<tr>
<td>• False</td>
<td>• counterfeit+food</td>
<td>• How to make fake alcohol?</td>
</tr>
<tr>
<td>• Intentional substitution</td>
<td>• olive+oil+dilution</td>
<td>• How to make fake milk?</td>
</tr>
<tr>
<td>• Smuggling</td>
<td>• olive+oil+fraud</td>
<td>• How to make fake Oil?</td>
</tr>
<tr>
<td>• Dilution</td>
<td>• counterfeit+olive+oil</td>
<td></td>
</tr>
<tr>
<td>• Substitution</td>
<td>• mislabeled+olive+oil</td>
<td></td>
</tr>
<tr>
<td>• ..etc</td>
<td>• honey+mislabling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• honey+fraud</td>
<td></td>
</tr>
</tbody>
</table>

**Total of included key words**  

64 x 8 = 512
Key Words Languages

- Spanish
- Dutch
- German
- Italian
- French
- Arabic
- Chinese
Food fraud filter efficiency:

More than 95 Food Fraud articles in last 2 months. Relevance index = 86%.
MedISys adapted to detect food fraud
Wine fraud in the UK is ‘massive’, says expert

By Nicholas Robinson+, 03-Dec-2014

Related topics: Food Safety, Beverages

Fraud in the UK’s wine industry is rampant because the sector has failed to implement the controls necessary to stamp it out, a food safety expert has claimed.

Wine fraud in the UK has been grossly underestimated because of the way the sector is structured, claimed the food safety expert, who did not want to be named.

Big wine companies were not placing enough emphasis on preventing fraud and, as a result, crooks were cashing in on the opportunity, he said.

Food fraud costs UK firms £11bn a year

By Laurence Gibbons+
28-Nov-2014
Last updated on 28-Nov-2014 at 13:02 GMT

Fraud is costing the food industry £11.2bn a year
The horrors of food fraud: vodka which kills and rotting fish disguised with chemicals

UK: An illicit factory producing fake vodka was discovered in Ireland which was found to have more than 20,000 empty bottles. It was meant to look like real vodka, but the final product was found to contain high levels of isopropyl alcohol.

2,500 Tonnen gefälschte Lebensmittel beschlagnahmt

Von Mozzarella, frisch gemachten Meeresfrüchten und falschem Whisky: Mehr als 2,500 gefälschte oder verdorbene Lebensmittel haben die Behörden in 47 Ländern der Schwerpunkt-Operation „Opson IV“ beschlagnahmt. Auch Italien war betroffen.

Most organic food in market is fake: food safety official

A.K. Mini, Assistant Food Safety Commissioner (Kollam), inaugurating a seminar on food safety in Kollam on Saturday. Photo: C. Suresh Kumar.

Much of the organic foods available in the markets are fake, Assistant Food Safety Commissioner (Kollam) A.K. Mini has said.
EMM Food Fraud Filter Design steps

1. **News papers**, **Blogs**, **Databases**, **Websites**
   - Updated every 10 minutes.
   - 24/7
   - 60 languages

2. **Text mining tool**
   - 600 keywords.
   - 8 languages.
   - ...etc

3. **mongoDB** ↔ **elasticsearch**
   - Automatic retrieval of reports
   - Automatic data storage
   - Automatic data processing

4. **ArcGIS**
   - Data visualisation using

- 6000 websites
- ...etc

- 600 keywords.
- 8 languages.
- ...etc
Data visualisation using ArcGIS

Food fraud reports (overall) in MedISys (period September 2014 to December 2015; N = 1114)
4 Data visualisation using ArcGIS

Food fraud reports (milk) in MedISys (period September 2014 to December 2015)
Identification of fraud: research at RIKILT

- Analytical methods (chemical, physical, microscopy)
  - Targeted approach: single marker
  - Untargeted approach: Multiple markers
    - fingerprints

- Text mining approach (MedISys)

- Trend analysis (databases (RASFF, EMA, UPS etc) & prediction modelling (agent modelling, Bayesians network)
Food fraud is reported in Rapid Alert System for Food and Feed (RASFF)
Products related to food fraud in RASFF

- meat and meat product
- fish and fish products
- other food product / mixed
- poultry meat and poultry meat products
- confectionery
- milk and milk products
- fats and oils
- honey and royal jelly
- herbs and spices
- soups, broths, sauces and condiments
- non-alcoholic beverages
- gastropods
- food additives and flavourings
- molluscs and products thereo
- wine
Modelling of Fraud in RASFF using Bayesians network

Model developed with RASFF data up to 2013

Model validated with RASFF data of 2014

Model performance:
- 82% correct prediction of fraud type
- 51% prediction of new fraud combinations
- static

Higher performance with dynamic model approach
(Bouzembrak and Marvin, 2016) used BNs prediction model to aid the border inspector to detect the type of food fraud.
Bayesian Networks (BNs) modelling in Big Data research

BN is able to combine data from different nature and location
Application of BNs to predict food fraud

Prediction of food fraud using multiple data sources and expert knowledge

- Identify parameters relevant for food fraud (=> nodes in BN model)
- Determine the relationships between these parameters (expert knowledge, literature)
- Identify data sources of these parameters
- Develop the BN model
- Validate the BN model
Application of BNs to predict food fraud

Parameters relevant for food fraud (some examples)

**Economic parameters:**
- Prices of the fraudulent product at the time of detection
- Price spike around the period of detection
- Trade volumes of the product between the country of detection and country of origin
- Complexity of the food chain

**Parameters of the country of origin & detection**
- Indices: corruption index, food safety index, governance index, legal system index, press index, human development index and technology index.

**Detected food fraud cases**
- RASFF and EMA
BN model to predicting food fraud
linking 36 data sources (18 databases and 8 expert judgements)
BN model to predict food fraud

Results: 91.5 % correct prediction of food fraud type
The effect of country of origin on the type of food fraud for meat in the low price category originating from Brazil (dark blue), Argentina (red), Poland (green), Canada (purple) and Belgium (light blue).
Application of BNs to predict food fraud

- (Marvin et al, 2016) A holistic approach to food safety risks: Food fraud as an example.
Conclusions

- The food fraud filter in MedISys, as developed by RIKILT, is useful to collect publications in the media on food fraud.
- The food fraud coverage in the media differ greatly between countries.
- There is an increasing trend in the number of fraud incidents, as reported in RASFF.
- Meat & meat products and fish & fish products are the food categories with the highest number of reports.
- BN models are suitable to predict the food fraud category.
Thank you

Acknowledgement
We would like to thank the following RIKILT WUR colleagues:
Esmée Janssen,
Ine van der Fels Klerx,
Esther van Asselt,
Gijs A. Kleter.