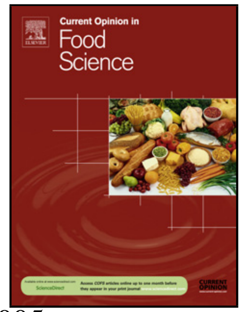
Accepted Manuscript



Title: Network analysis: A promising tool for food safety

Author: Declan P Naughton Tamás Nepusz Andrea Petróczi

PII: S2214-7993(15)00181-2
DOI: http://dx.doi.org/doi:10.1016/j.cofs.2015.12.005
Reference: COFS 113

To appear in:

Received date: 2-9-2015
Accepted date: 19-12-2015

Highlights

- A network analysis toolallows rapid interrogation of large databases
- The ‘live tool’is capable of automatic or manual data download
- It comes as a desktop application allowing regular data updates
- Data mining is by network analysis or descriptive statistics
- Outputs come as network maps or charts for reporting

Network analysis: a promising tool for food safety

Declan P Naughton[1], Tamás Nepusz[2] and Andrea Petróczi[1]

[1] School of Life Sciences and Pharmacy, Faculty of Science, Engineering and Computing, Kingston Univereisty London, Kingston upon Thames, London KT1 2EE

[2] Sixdegrees Ltd., Érd, Hungary

Corresponding author: Prof Declan Naughton, School of Life Sciences and Pharmacy, Faculty of Science, Engineering and Computing, Kingston University London, Kingston upon Thames, London KT1 2EE.

Email: D.Naughton@kingston.ac.uk

Phone: +44 208 417 7097

Abstract

Global challenges in food safety include a range of concerns: (i) persistent long-term issues such as heavy metals in seafood, (ii) emerging incidents such as a particular food contaminant (e.g. melamine), (iii) the apparent growth in food fraud,and (iv) the effectiveness of regulation and enforcement policies and their implementation acrossnations/borders.Food testing, a key unifying theme across these concerns, produces enormous databases of confirmed or suspected food and animal feed across the globe. In a resource-constrained environment, food safety officials would benefit from advanced data-mining applications to optimise use of the rich information these databases contain. This report overviews the network analysis approachwhich allows rapid interrogation of large databases to identify trends in nations detecting and producing faulty foods.

Introduction

World-wide food supply chains have become so complex that the term ‘food supply network’ is usually more appropriate. The volume of testing that is required - to secure safe food and feed supplies – produces an enormous level of information across the globe. Frequently, test results are combined into large national- or continental- level databases which are not user-friendly to examine for complex queries. Network analysis is particularly suited to studying such databases as itcaptures the relationships (termed ‘edges’) between food supplier and user nations (termed ‘nodes’), with edges having ‘weights’ (e.g. the number of reports filed/received). Network analysis can greatly surpass descriptive statistics in capturing the complexity by simultaneously taking the number of reports and number of countries involved into consideration. It provides a mathematical expression of the network properties with a good visualisation output for report generation. To date, network analaysis has been applied to study a range of phenomena with a view to understanding:connectivity in fighting terrorists [1,2], new approaches to effectively intervene in the growth of crime circles [3,4], examining the spread of infectious diseases [5] or studying the interactions of molecules in complex biosystems [6,7].

A limited number of publications have applied network analysis to the study of food safety issues. These include a report on the poultry-linked microbiome from farm to fork [8] and a concise analysis of meat traceability for Brazilian beef and pork [9].A further report employed network analysis to investigate dynamics of complex international food trade networks and the links to food poisoning outbreaks [10].From an overview of other fields [1-7], the potential for network analysis to assist enforcers and researchers on matters of food safety is considerable. This review aims to present an update on a bespoke network analysis tool that has been designed specifically for use with food safety databases.

Prototype Network Tool

The authors developed a bespoke network tool to interrogate food safety databases. In its prototype form, it was used as a research tool to study combined global food safety data along with the European Rapid Alert System for Food and Feed (RASFF) database. Early work analysing worldwide food recall patterns showed that the majority of faulty foods originate in ten countries with four of the major producers making no reports [11]. The study highlighted the interactions between major reporters of faulty food (*detectors*) and suppliers (*transgressors*) revealing prominent roles for Iran, Turkey and China as transgressors. Analysis of food notification patterns for metal contamination in seafoods revealed the major nations acting as transgressors and detectors along with longitudinal variations [12]. This study revealed the presence of clusters of transgresssors / detectors, within the overall reporting patterns, which may be useful for monitoring and to isolate transgressors or enhance detector activity.Triangulation of these network analytical results with historical data provided support for the validity of the network approach.

A further report documented the activities within the RASFF database over a decade revealing the major reporting nations to be Italy, Germany, the UK and Spainwhich collectively made 60% of the reports [13].

Variations in reporting patterns over time were readily observed and the paperintroduced the publically-available interactive network tool designed for food safety with avariety of selectable features and an export function for report generation. The protoptye network tool was applied to study transgressor-detector interactions for mycotoxin safety reports [14] and in fuller form with a range of searchable sub-categories including (i) bacteria, (ii) metals, (iii) mycotoxins, (iv) all reports, (v) microbiological, as well as (vi) border rejections [15].

A further study which employed a combination of descriptive statistics and network analysis to investigate patterns of reporting to the RASFF database observed large variations in the contributions made by EU member states [16]. Using the detector index function, member state contributions over the period for all notifications versus border rejection notifications revealed some member states had high levels of activity in the key area of border rejections.For countries with high activity in border rejections, a medium to strong positive correlation exists between 'all notifications' and 'border rejections' suggesting a key role for border rejection notifications in influencing notification patterns.

Foodguard – a 'Live' Network Tool

Key design features

In contrast to adopting a commercially-available software package, the major advantage of desigining a bespoke network tool for food safety is the ability to incorporate key features along with road testing the programme. From this perspective the desirable major features of the network tool were selected to be:

- a desktop application;
- allow periodic automatic feeds from the RASFF database with an update facility upon request (e.g. daily);
- allow manual feeds from other food safety data which could be combined with the RASFF data
- contain a facility to automatically check and remove duplicates;
- contain a facility to validate and sort incorporated data where queries arise;
- offer combinable filter functions to allow data selection and focus;
- capability of data mining via a network analysis and/or descriptive statistics (chart) approach;
- capable of ready switch between network analysis and descriptive statistics functions;
- provide detector indices and transgressor indices to monitor the relationships between reporting (detector) and reported (transgressor);
- provide outputs as data files or as pictures (network maps and charts)for report generation.

Live Network Tool

Recently, the work progressed with the development of a 'live tool' which is capable of automatic or manual data download and comes as a desktop application. The facility for automatic feeds allowing the user to download dataat regular intervals was a key feature affording the latest available food safety data to be incorporated. In addition, it was considered vital to be capable of adding additional data via manual downloads into the network tool along with the ability to merge datasets.The tool, entitled Foodguard, is a user-friendly desktop application with the familiar drop-down menu lists as well as interactive page selection for validating imported data and switching between network analysis mode and descriptive statistics (chart)mode.It also allows the export of charts, data and network maps for report generation (Figure 1A). The tool was designed to incorporate network analysis (Figure 1B) but extends to include a ready switch to a descriptive statistics function with an export facility. Thus, searches under the broad scope of the network function can be exported as charts in a number of formats.

An example of the use of the live tool is via the search facility which can cover dates, countries, contaminants including source and destination of faulty foods. The example provided is for a study of horsemeat during 2014 within the RASFF database. The search term ‘horse’ (highlighted) automatically generates the network map (Figure 2A) revealing Italy in green as a major detector followed by Sweden and France and Cyprus. Romania (with self reporting), Poland and Argentina are the main transgressors shown in red. A range of the options exist to further refine the network map – for example by date or country or type of classification – which isuseful for large search subsets. In addition, the search results may be instantaneously portrayed as bar charts or line diagrams (Figure 2B). There are several options to draw bar charts as a combination of actions, agencies, notification types, classifications, notifier countries, countries of origin, and dates. The selection for Figure 2B was ‘actions’ and ‘product categories’ (highlighted). A key aspect is that this interrogation and report generation is conducted in circa one minute affording multiple searches on new (weekly or daily) updates.

A further exemplification is a search of melamine in the RASFF database for 2014. The network map is instantly generated revealing China to be the major transgressor with a limited number of reports against Hong Kong (Figure 3A). The network map readily shows the reporting countries and the results can be rapidly exported as either a network map or as a bar- or line graph (Figure 3B). In this case the options were picked as ‘weeks’ and ‘actions’. The report wasgenerated in circa one minute.

Conclusions and future directions

This report provides an overview of network analysis and the potential to assist enforcement officials with interrogation of large databases. With the advent of Foodguard, a live feed tool, network analysis may be used in real time interrogation of the latest food safety data in order to analyse trends and identify emerging concerns (such as a cluster of transgressing countries). The same approach may be extended for use with suppliers within a large organisation in lieu of nations as portrayed in these examples. Company data could be used for this purpose in a confidential manner.

Further applications should address the following uses: (i) to assist with resource allocation – i.e. inform optimal testing and to reduce duplicate testing; (ii) to bring more 'profit' from RASFF and other databases – for example by combining them, and looking for clusters;(iii) to monitor the effects of changes in regulation; (iv) by companies to monitor supplier chains; and, (v) to identify emerging incidents such as underperforming nations in regard to food safety reporting.

References

1. Maheshwari S, Tiwari A: **A novel genetic based framework for the detection and destabilization of influencing nodes in terrorist network**.Smart Innov Syst Technol 2015, **31**:573-582.
2. Husslage B, Borm P, Burg T, Hamers H, Lindelauf R: **Ranking terrorists in networks: A sensitivity analysis of Al Qaeda's 9/11 attack.**Soc Networks2015, **42**:1-7
3. Rostami A, Mondani H: **The complexity of crime network data: A case study of its consequences for crime control and the study of networks**. PLoS ONE 2015, **10**:e0119309.
4. Bright DA, Greenhill C, Ritter A, Morselli C: **Networks within networks: using multiple link types to examine network structure and identify key actors in a drug trafficking operation**. Global Crime 2015, **16**:219-237.
5. Mehta SR, Wertheim JO, Brouwer KC, Wagner KD, Chaillon A, Strathdee S, Patterson TL, Rangel MG, Vargas M, Murrell B, *et al*.: **HIV Transmission Networks in the San Diego-Tijuana Border Region.** EBioMedicine2015, http://dx.doi.org/10.1016/j.ebiom.2015.07.024.
6. Nepusz, T., Yu, H., Paccanaro, A: **Detecting overlapping protein complexes in protein-protein interaction networks**. Nat Methods 2012, **9**:471-472.
7. Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, Wang P, Boutz DR, Fong V, Phanse S, *et al*.: **A census of human soluble protein complexes**.Cell 2012, **150**:1068-1081.
8. Oakley BB, Morales CA, Line J, Berrang ME, Meinersmann RJ,Tillman GE,Wise MG,Siragusa GR, Hiett KL, Seal BS:**The Poultry-Associated Microbiome: Network Analysis and Farm-to-Fork Characterizations.**PLoS ONE2013, **8**:e57190.
9. Nääs IDA, Neto MM, Vendrametto O, Canuto SA: **Comparative analysis of different meat traceability systems using multiple criteria and a social network approach**. Eng Agric2015, **35**:340-349.
10. Ercsey-Ravasz M,Toroczkai Z,Lakner Z,Baranyi J: **Complexity of the International Agro-Food Trade Network and Its Impact on Food Safety**. PLOS One 2012, e37810.
11. Nepusz T, Petróczi A, Naughton DP: **Worldwide food recall patterns over an eleven month period: A country perspective.**BMC Public Health2008, **8**:308.

12. Nepusz T, Petróczi A, Naughton DP: **Food alert patterns for metal contamination analyses in seafoods: Longitudinal and geographical perspectives.** Environ Int 2009, **35**:1030-1033.

13. Petróczi A, Taylor G, Nepusz T, Naughton DP:**Gate keepers of EU food safety: Four states lead on notification patterns and effectiveness**.Food Chem Toxicol2010, **48**:1957-1964.

14. Petróczi A, Nepusz T, Taylor G, Naughton DP: **Network analysis of the RASFF database: A mycotoxin perspective**. World Mycotoxin J2011, **4**:329-338.

15. **Nepusz T, Petróczi A, Naughton DP: **Interactive network analytical tool for instantaneous bespoke interrogation of food safety notifications**. PLoS ONE2012, **7**:e35652. {This repoprt provided the first bespoke publically-available network analytical tool for interrogation of food safety alerts. The tool incorporates a filter function allowing interrogation of notifications by contaminant or type of notification. The network tool captures complexity and aids identification of reporting patterns between countries – taking the relationships between reporting (detector) and reported (transgressor) nations into account.}

16. *Taylor G, Petroczi A, Nepusz T, Naughton DP: **The Procrustean bed of EU food safety notifications via the Rapid Alert System for Food and Feed: Does one size fit all?**Food Chemical Toxicol2013, **56**:411-418.{This study uses a combination of descriptive statistics and network analysis to report major variations in the contributions made by EU member states to the RASFF database. Network analysis was employed to compare the detector indices of member states over the period for all notifications versus border rejection notifications. The study highlights member states with high levels of activity in the key area of border rejections.}

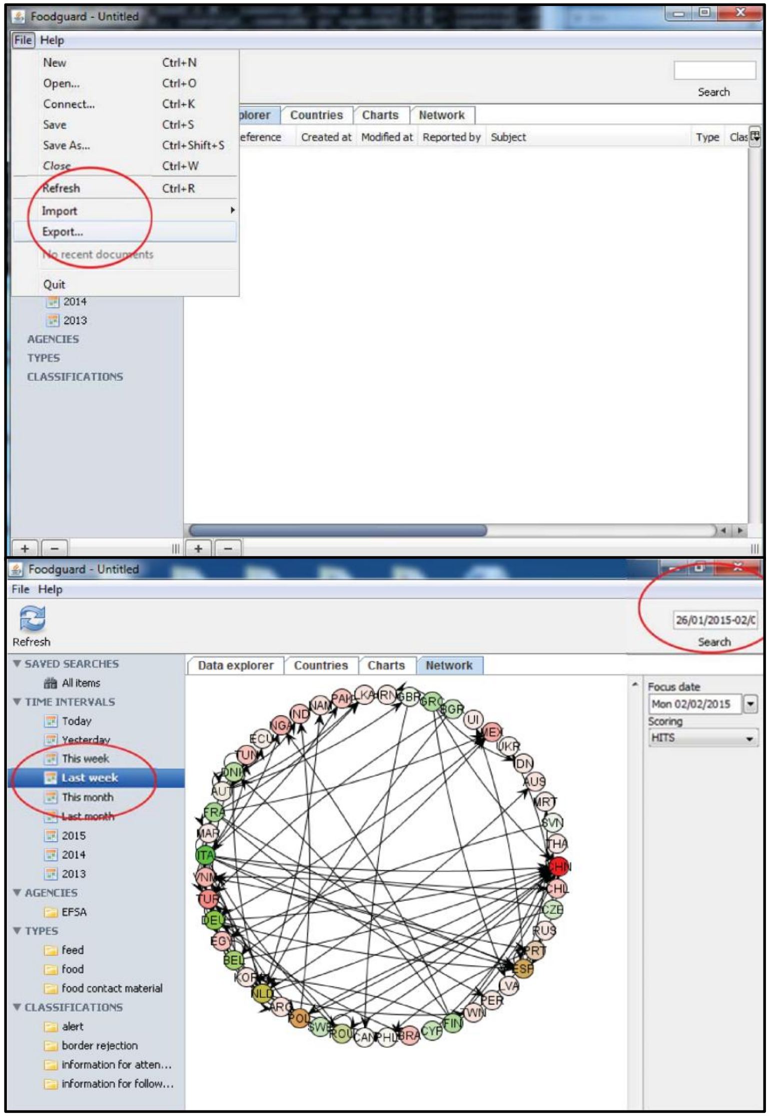
Figure legends

Figure 1: A–(top) showing file menu options for the desktop tool with data import- and result export-featureshighlighted; B - showing selection choices of duration, source, types and classifications – a network map for a one week period is shown. Nations coloured red are transgressors with those in green are detectors of faulty foods.

Figure 2: A–(top) showing network map for a selection of ‘horse’ from 2014;B - showing the same data portrayed as a bar chart under the options of ‘actions’ and ‘product categories’.

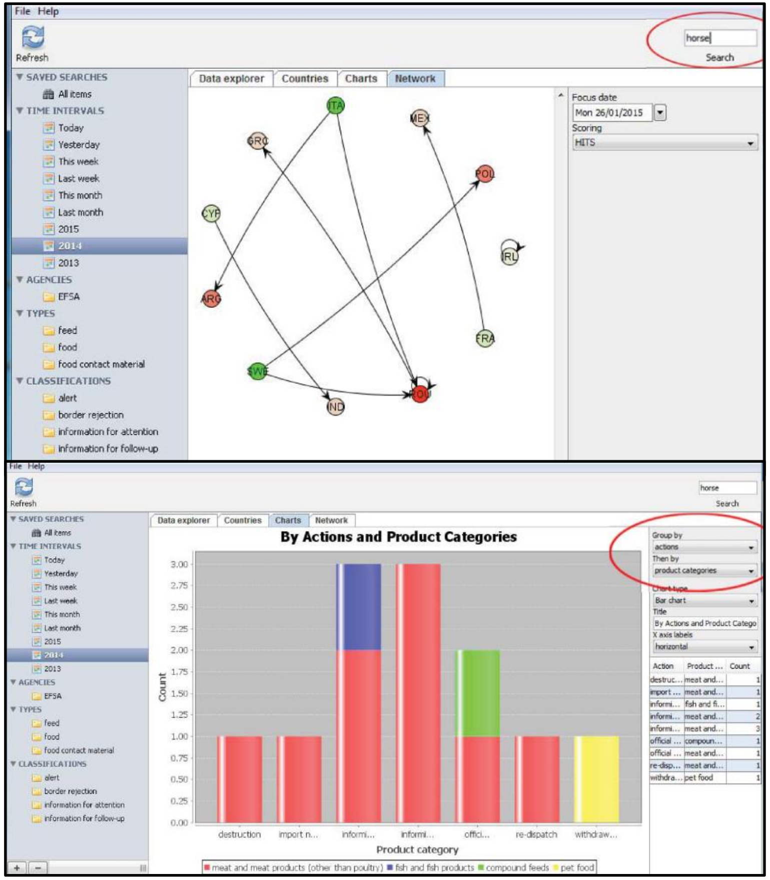
Figure 3: A–(top) showing network map for a selection of ‘melamine’ from 2014;B - showing the same data portrayed as a bar chart under the options of ‘weeks’ and ‘actions.

Figure 1



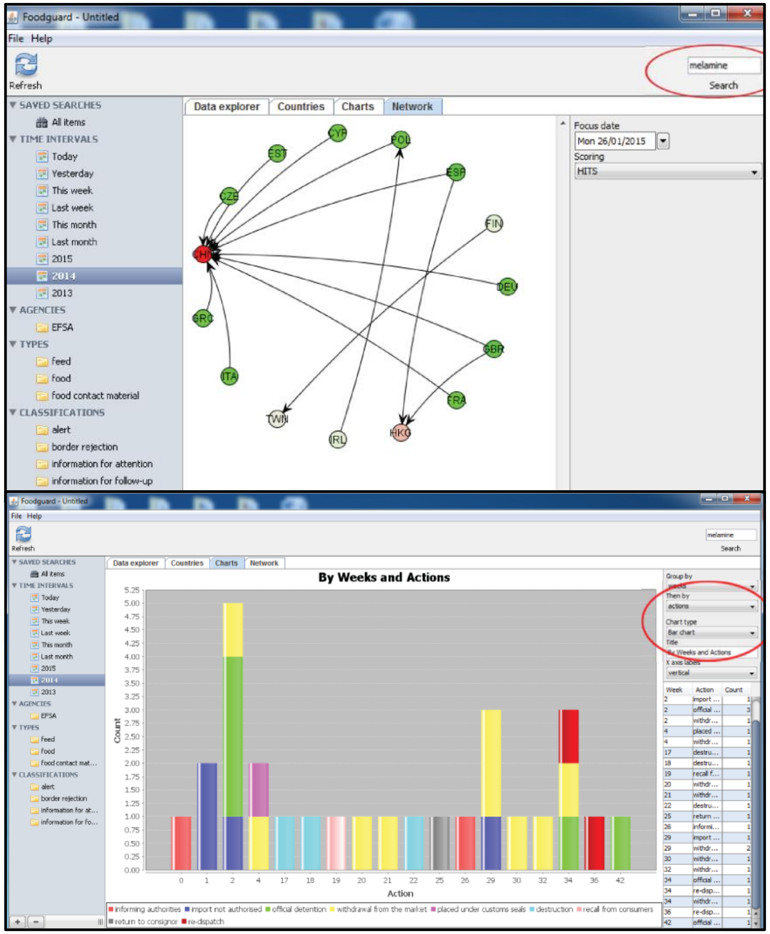
Foodguard - Untitled
File
Help
New
Ctrl+N
Open...
Ctrl+O
Connect...
Ctrl+K
Save
Ctrl+S
Save As...
Ctrl+Shift+S
Close
Ctrl+W
Refresh
Ctrl+R
Import
Export...
No recent documents
Quit
Search
Countries
Charts
Network
Created at
Modified at
Reported by
Subject
Type
2014
2013
AGENCIES
TYPES
CLASSIFICATIONS
Foodguard - Untitled
File
Help
Refresh
Search
SAVED SEARCHES
All items
TIME INTERVALS
Today
Yesterday
This week
Last week
This month
Last month
2015
2014
2013
AGENCIES
EFSA
TYPES
feed
food
food contact material
CLASSIFICATIONS
alert
border rejection
information for atten...
information for follow...
Data explorer
Countries
Charts
Network
Focus date
Mon 02/02/2015
Scoring
HITS

Figure 2



File Help
Refresh
horse
Search
SAVED SEARCHES
All items
TIME INTERVALS
Today
Yesterday
This week
Last week
This month
Last month
2015
2014
2013
AGENCIES
EFSA
TYPES
feed
food
food contact material
CLASSIFICATIONS
alert
border rejection
information for attention
information for follow-up
Data explorer
Countries
Charts
Network
Focus date
Mon 26/01/2015
Scoring
HITS
ITA
MEX
GRC
POL
CYP
IRL
ARG
FRA
SWE
ROU
IND
By Actions and Product Categories
Count
Product category
destruction
import n...
informi...
informi...
offici...
re-dispatch
withdraw...
meat and meat products (other than poultry)
fish and fish products
compound feeds
pet food
Group by
actions
Then by
product categories
Chart type
Bar chart
Title
By Actions and Product Catego
X axis labels
horizontal
Action
Product ...
Count

Figure 3


Foodguard - Untitled
File Help
Refresh
melamine
Search
SAVED SEARCHES
All items
TIME INTERVALS
Today
Yesterday
This week
Last week
This month
Last month
2015
2014
2013
AGENCIES
EFSA
TYPES
feed
food
food contact material
CLASSIFICATIONS
alert
border rejection
information for attention
information for follow-up
Data explorer
Countries
Charts
Network
Focus date
Mon 26/01/2015
Scoring
HITS
CYP
POL
EST
ESP
CZE
FIN
DEU
GRC
GBR
ITA
FRA
TWN
IRL
HKG
By Weeks and Actions
Count
Action
informing authorities
import not authorised
official detention
withdrawal from the market
placed under customs seals
destruction
recall from consumers
return to consignor
re-dispatch
Group by
Then by
actions
Chart type
Bar chart
Title
By Weeks and Actions
vertical
Week
Action
Count