

Using Unitary Traceability for an Optimal Product Recall

Thierno M.L. Diallo¹, Sébastien Henry¹, and Yacine Ouzrout²

¹ DISP Laboratory, University of Lyon, University Lyon 1, France
{Thierno.Diallo, Sebastien.Henry}@univ-lyon1.fr

² DISP Laboratory, University of Lyon, University Lyon 2, France
Yacine.Ouzrout@univ-lyon2.fr

Abstract. Product recall is a challenge which may have a significant financial impact. Incidents should be anticipated to improve responsiveness and reduce potential harm. In this paper, we propose a product recall approach following the detection of a critical fault. It is applicable to foodstuffs industry characterized by complex processes with high variability, high-speed manufacture and very large lots sizes. In such a case, usual strategy which consists of recalling entire lots is expensive and does not foster continuous improvement. The proposed approach in this paper allow to identify root causes and other products likely to present the same noncompliance in order to make a targeted recall. The root causes are searched based on an analysis of traceability data using a Bayesian model. A data model suitable for product and process traceability is also proposed. The originality of our approach lies on the reconstitution of the conditions of manufacturing of each item through the coupling of product and process unitary traceability data.

Keywords: Product Recall, Root Causes, Traceability, Bayesian Networks.

1 Introduction

Despite all the control means implemented in industries, the risk of shipping noncompliant products that do not meet consumer safety standard always exists [1-3]. When this nonconformity might cause serious and lasting health problems or death because of the use or exposure to the product then a recall is required. Due to globalization of exchanges (several suppliers and customers across all continents), product complexity (several ingredients and complex manufacturing processes) and regulations (accountability of the manufacturer on its product), product recall is nowadays a challenge that is facing more and more industries[4]. Although product recall may concern any type of product, but it is especially more frequent and critical in the field of food products. The causes of these recalls mainly come from materials, equipment or processes [1, 2, 4]. The usual strategy which consists of recalling entire lots generates direct and indirect costs especially on branding. These massive recalls are generally done without knowing the status (compliant or not) of recalled products. There are several examples where companies do very large recall because they cannot identify really defective items [2].

The product recall scenarios are quite varied depending on the type of product, the type and scope of the supply chain and regulation. These scenarios also vary depending on the lifetime of the product, its manufacturing process and the actors involved in its life cycle (manufacturer, distributor, and retailer). The common point of all these scenarios is the need to trace the product to be recalled [5]. All recall scenarios use more or less traceability data to achieve the recall procedure. The knowledge of product life cycle from the origin of raw material is a key factor for the identification of products to be recalled and the definition of an effective recall procedure [4-6]. That is why traceability plays a crucial role in product recall. The determination of the causes of a nonconformity will depend in large part on the quality of the traceability system.

The use of deterministic tools for causal analysis in product recall procedure as reported in [4, 6] is not always justifiable. In this work, we devise a framework for optimal product recall following a nonconformity finding based on a probabilistic causal model using a Bayesian network (BN). To find these causes, the knowledge of product records is necessary. To our knowledge, the few published data models dedicated to the unitary traceability (see for example [7, 8]) have some restrictions in terms of actual material and process data registration. The unitary traceability system (enabling a serialized unique identification at the item level) and the proposed data model allow to know accurately the process parameters of each item. From this history, causal analysis is performed to identify potential causes for the noncompliance and other products that may be affected. Determination of the list of potentially noncompliant products thus enables a targeted recall.

This paper is organized as follow. The overall recall procedure proposed is presented in Section 2. In Section 3, we propose a data model for collecting traceability data required for the developed causal analysis framework. This causal analysis framework is presented in Section 4. And last we conclude with the mains contributions and perspectives of this work.

2 Traceability and the Proposed Product Recall Procedure

Traceability is an effective means for the mastery of the supply chain and for production optimization (scheduling and resources optimization). It enables to cope with urgent and unforeseen situations such as product recalls. When the element under consideration is a food product, the important elements to trace are raw materials and ingredients making up the product, the history of transformation processes and distribution and location of the product after delivery [9]. The traceability unit can be an aggregation of several articles (e.g. a lot or a pallet) or can be thinner and correspond to an article (unitary traceability). Within the framework of product recall, both types of traceability are required: tracking or forward traceability and tracing or backward traceability. Forward traceability is used to determine, for example, finished products containing a particular ingredient or having undergone a specific process. Backward traceability offers the possibility to identify suppliers and processes involved in producing a particular article [5]. In terms of visibility and management policy, there are two levels of traceability; internal and external traceability [10]. Internal traceability

deals with private data stored within the company for internal use or in case of requests from the authorities. Among these data may be mentioned the process parameters, the origin of raw materials, quality records, etc. External traceability concerns public data that the different partners in a supply chain shall exchange between them to ensure end-to-end traceability. The recall procedure presented in this paper uses unitary and internal traceability data.

Unlike the conventional recall procedure, the aim of the proposed procedure is to restrict the number of items to recall by exploiting unitary traceability data. This strategy is especially relevant for products with high added value with large lot sizes (>1 million). The proposed recall procedure starts with the detection or the reporting of a noncompliance. Fig.1 illustrates graphically the main tasks of the proposed recall procedure.

Confirmation of the Noncompliance. This preliminary task enables the product manufacturer to establish unambiguously the noncompliance reported and to place the responsibilities between the different stakeholders (manufacturer, shipper, distributor, etc.). To do this, it may refer to historicized traceability data and tests. These traceability data must be reliable and usable.

Determine Possible Causes of the Noncompliance through Backward Traceability. Determine root cause (raw material /process / machine parameters) likely to be causing the noncompliance. This task corresponds to the backward traceability process. This is a challenging task due to the large number of parameters to be considered, their large variability and uncertainties. Causes other than process / machine parameters such as design defects may give rise to non-compliances. But in our study, we limit ourselves to causes related to process / machine parameters.

Determine the Duration of Abnormal Operation. Determines the range of time during which the manufacturing process has been in abnormal operation. This abnormal operation is characterized based on the identified root causes. Through the traceability data, the goal is to automate this task. Spite of the 100% control, the risk of shipping noncompliant products is nonzero especially in process industries.

Identify and Locate Items to Recall by Forward Traceability. The determination of the offending articles is crucial for an optimal recall. With a reliable unitary traceability, it is possible to restrict the number of recalled items. Through the unit traceability system, it is also possible to determine the position of each item in the supply chain by its identifier.

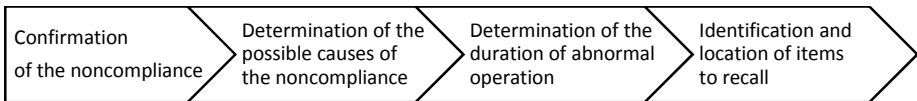


Fig. 1. Main tasks in our recall procedure

The implementation of this procedure depends largely on the traceability data and the causal analysis to determine root causes. This was two issues we have identified in this research work. The following two sections present our contribution to cope with these two challenges.

3 Data Model for Traceability

Through the development of automatic data collection tools and technologies (RFID, Data Matrix, etc.), it is possible to track each item and accurately determine its process parameters. The unitary traceability (unlike lot traceability) provides a more detailed knowledge of the process. It allows a finer search for the roots causes. To find the roots causes of noncompliant in the context of this research work, we mostly need the internal traceability data. A few data models dedicated to the unitary traceability have been published. Jansen-Vullers, M.H., et al. [7] and Khabbazi, M.R., et al.[8] develop an internal traceability data model with some restrictions in terms of actual material and process data registration. The traceability performed in the industrial area (including pharmaceutical and food) is usually managed by disconnected data models. The motivations are often compliance with regulatory in unusual situations. Root cause search is very tedious in these conditions. We aim to collect all the necessary data and facilitate the causal analysis to determine the root causes and facilitate data exchange. IEC 62264 [11] standard provides objects models and attributes of manufacturing operations. The GS1 EPC (Electronic Product Code) Global standards [12] allow end-to-end product traceability along a supply chain. We based on IEC 62264 and EPC Global standards to propose item-based traceability data model (Fig.2). The proposed model allows to know for each item, the process parameters of its manufacture. In our data model, the traceability data are organized by production order. The production order data is made up of data related to different process segments. For each process segment, production data, material consumed actual and material produced actual are recorded.

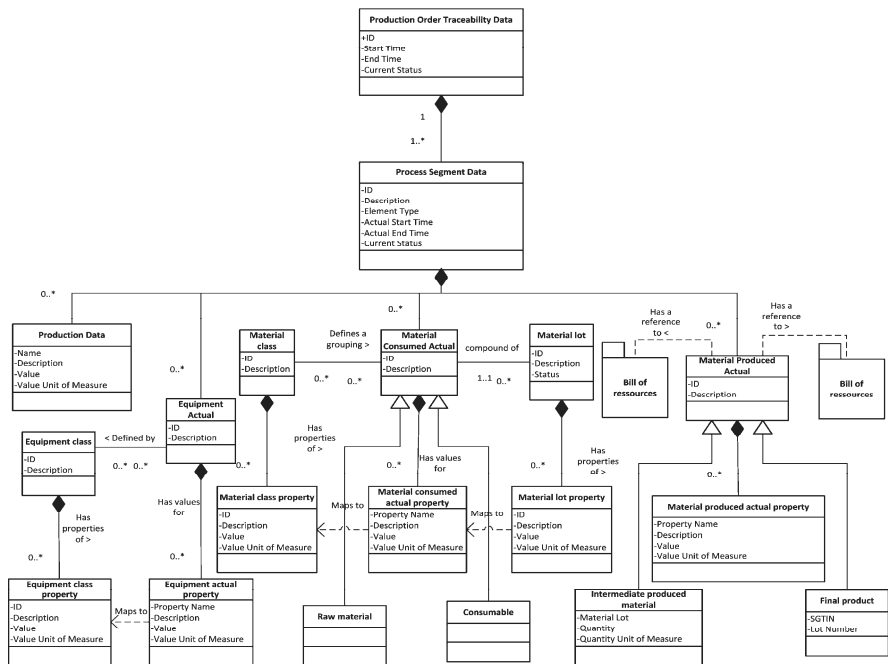


Fig. 2. Unitary traceability data model based on IEC 62264 [11] and GS1 EPCIS [12] standards

Usually, when a nonconformity is detected, one contents oneself with recalling the whole lot without seeking prior exact causes of this noncompliance. However, this search can promote an optimal recall focusing only on defective items and processes improvement. The benefits gained by a selective recall are considerable especially in the case of large lots sizes and a large dispersion of finished products. But identifying root causes of a recall is in some cases a complex task. The following section presents our contribution to the search for the roots causes related to process parameters.

4 Causal Analysis by Bayesian Modeling

Food industry often operate in a complex and difficult to model process. It is also characterized by a large number of parameters, a large variability and various types of failure. It is therefore difficult to diagnose a noncompliance. Various analytical tools have been used to determine the root causes of a noncompliance as part of a product recall procedure (see for example [4, 6]). Deterministic reliability engineering tools such as FMECA, HACCP, cause effect diagram and fault tree are often used. In the industrial context outlined herein, the use of deterministic methods with categorical decisions is not always justifiable [4]. In this paper, we propose a probabilistic causal model using Bayesian networks (BNs). BNs are graphical models for reasoning under uncertainty [13]. They allow to combine, on the one hand, certain and uncertain knowledge. On the other, they allow to exploit both data and expertise. A BN is Directed Acyclic Graph (DAG) represented by the pair (V, E) where V is a set of vertices and E a set of directed edges connecting vertices. It is associated with each node marginal or conditional probability distribution table of the corresponding variable. The main purpose of our model is to determine the parameters that might be responsible for the detected nonconformity.

We first present the construction of the model from the prior knowledge and historical data and then the use of the model with traceability data. Both historical and traceability data follow the data model developed in the previous section.

4.1 Structure of Our Model

The structure of the BN of our model is defined by background knowledge. Nodes and causal relationships are obtained from FMECA and fault tree produced by domain experts. The network is structured in 3 levels (see Fig.3): process parameters considered as root cause, product defect and nonconformity detected by end-users.

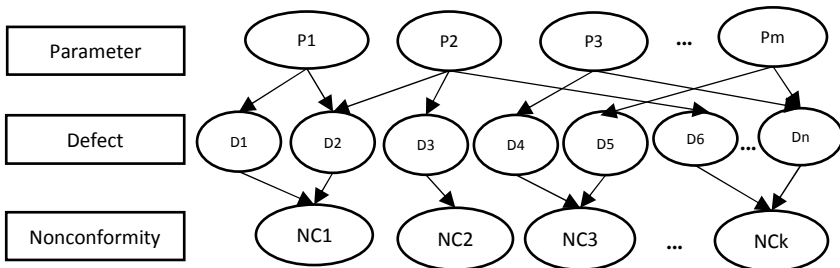


Fig. 3. The BN structure

The first level correspond to the different process/machine parameter considered that may impact the product quality. The product defect is any deviation from the specification of the product. The nonconformity represents a symptom or effect of one or many product defects.

Depending on its values (Normal or Abnormal), a process/machine parameter may generated (G) or not generated (NG) a product defect. As for the first two parameters, a nonconformity also has 2 states: observed or unobserved. Each of these nodes has a probability table.

4.2 Determination of Our Model’s Probabilities

Let X be a random or uncertain variable. We denote $p(X = x|\xi)$ or $p(x|\xi)$ the probability that $X = x$ with state of information¹ ξ . We assume that variable x is distributed according to f , where f is a parameterized probability distribution. This parametric model is noted $x \sim f(x|\theta)$ with θ the set of unknown parameters. The uncertainty on the parameters θ is modeled through a probability distribution $\pi(\theta)$ called prior distribution. The most critical point of Bayesian analysis is the choice of the prior distribution [14, 15]. Bayesian modeling has been adopted in several diagnostic and causal analysis applications. But the determination of f and the prior distribution of root causes has been little discussed. We start by addressing this issue.

We assume process parameters to be independent variables distributed according to a Gaussian distribution $f(x, \theta): f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ with $\theta = \{\mu, \sigma\}$.

With prior knowledge through the process data historian and in order to meet the requirements of objectivity, as the normal distribution belong to exponential families, we proceeded with the conjugate priors approach.

From $f(x|\theta) \sim \mathcal{N}(\mu, \sigma^2)$, we deduce $\pi(\mu)$ and $\pi(\sigma)$ (the following development is based on [14-16]).

$$\pi(\mu|\mu_0, \delta_0) = \frac{1}{\delta_0\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\mu-\mu_0}{\delta_0})^2}$$

$$\pi(\sigma|\sigma_0) = \frac{\sigma_0}{\sigma^2}$$

The hyperparameters $\mu_0, \delta_0, \sigma_0$ are determined by the Markov Chain Monte Carlo (MCMC) method and Maximum-Likelihood Estimation (MLE).

We use traceability data to update the prior belief by calculating posteriori probabilities. According to Bayes’ theorem, $\pi(\theta|x, \mu_0, \sigma_0, \delta_0) = \frac{\pi(\mu|\mu_0, \delta_0)\pi(\sigma|\sigma_0)f(x|\mu, \sigma)}{f(x|\mu_0, \sigma_0, \delta_0)}$ where $f(x) = f(x|\mu_0, \sigma_0, \delta_0) = \int_{-\infty}^{+\infty} \int_{\sigma_0}^{\infty} f(x, \mu, \sigma|\mu_0, \sigma_0, \delta_0) d\sigma$

¹ Unlike the statistical probability, Bayesian probability varies according to background knowledge. But to make the notation less cluttered, we will ignore this conditionality that accompanies all Bayesian probabilities.

The best estimates of μ and σ are: $\hat{\mu} = \left\{ \mu: \frac{d \log_e [\pi(\mu | x)]}{d \mu} = 0 \right\}$ and $\hat{\sigma} = \left\{ \sigma: \frac{d \log_e [\pi(\sigma | x)]}{d \sigma} = 0 \right\}$

Process/Machine Parameter or Root Causes. For each parameter, we divide the value space into two regions: Normal (N) and Abnormal (AN).

Let x_{up} and x_{low} the upper and lower bounds of definition range of parameter P_i .

$$P(P_i = N) = P(x_{low} \leq x \leq x_{up}) = \int_{x_{low}}^{x_{up}} f(x) dx \text{ and } P(P_i = AN) = 1 - P(P_i = N)$$

Product Defects or Intermediate Nodes and Nonconformities or Leaf Nodes. For each node representing a product defect or a nonconformity, its conditional probability table is leaned from historical data and background knowledge. These probabilities are updated based on new data or knowledge acquired on the process. These updates may also suggest a change in the structure of the Bayesian network.

4.3 Use of the Model

In the previous two subsections we have built the Bayesian network modeling the causal relationships of the studied system. Its use consist in updating the prior belief and inference calculations. When a nonconformity is detected, this causal analysis model is used to found the root causes. This is a diagnostic reasoning problem. In general, probabilistic inference in Bayesian networks is NP-hard. Despite the potentially large size of this type of graph, this inference problem can be addressed by techniques that are custom tailored to particular inference queries [17]. Once the root causes behind the nonconformity have been determined, a search in the traceability data is conducted to determine the duration of abnormal operation and then the other items likely to be noncompliant.

5 Conclusion

In this paper, we have presented a holistic approach to product recall. This approach allows to optimize the recall procedure by a search of the roots causes. It thus allows to limit the recall solely to incriminated products. The proposed recall procedure is based on unitary product / process traceability. In order to facilitate the collection and exchange of traceability data, we have proposed a data model based on IEC 62264 and GS1 EPCIS standards. The search for causes of a recall is done by analyzing these data by means of a Bayesian model. This model allow to combine certain and uncertain knowledge and to exploit both data and expertise. In future research works, this approach will be implemented and evaluated in industrial context. Experimental results will be analyzed and discussed. An extension of variables' (parameter, defect, nonconformity) value spaces is also envisaged.

Acknowledgements. This work was supported by OSEO (French organization for innovation support and funding, Ministry for Economy, Finance and Industry, and Ministry for higher education and research) through the FUI Traçaverre project.

References

1. Potter, A., et al.: Trends in product recalls within the agri-food industry: Empirical evidence from the USA, UK and the Republic of Ireland. *Trends in Food Science & Technology* 28(2), 77–86 (2012)
2. Hora, M., Bapuji, H., Roth, A.V.: Safety hazard and time to recall: The role of recall strategy, product defect type, and supply chain player in the U.S. toy industry. *Journal of Operations Management* 29(7–8), 766–777 (2011)
3. Magno, F.: Managing Product Recalls: The Effects of Time, Responsible vs. Procedia - Social and Behavioral Sciences 58, 1309–1315 (2012)
4. Kumar, S.: A knowledge based reliability engineering approach to manage product safety and recalls. *Expert Systems with Applications* 41(11), 5323–5339 (2014)
5. Wynn, M.T., et al.: Data and process requirements for product recall coordination. *Computers in Industry* 62(7), 776–786 (2011)
6. Kumar, S., Budin, E.M.: Prevention and management of product recalls in the processed food industry: a case study based on an exporter’s perspective. *Technovation* 26(5–6), 739–750 (2006)
7. Jansen-Vullers, M.H., van Dorp, C.A., Beulens, A.J.M.: Managing traceability information in manufacture. *International Journal of Information Management* 23(5), 395–413 (2003)
8. Khabbazi, M.R., et al.: Data Modeling of Traceability Information for Manufacturing Control System. In: *International Conference on Information Management and Engineering, ICIME 2009* (2009)
9. ISO, E., NF EN ISO 22005: Traceability in the feed and food chain - General principles and basic requirements for system design and implementation. 2007.
10. AISBL, G. GS1 Traceability Standard (2013), <http://www.gs1.org/gsm/kc/traceability> (cited 2013 16 février 2013)
11. ISO/CEI, IEC 62264-2, in *Enterprise-control system integration – Part 2: Model object attributes*, p. 96 (2004)
12. GS1, *The GS1 EPCglobal Architecture Framework*, GS1 (2013)
13. Korb, K.B., Nicholson, A.E.: *Bayesian Artificial Intelligence*. Taylor & Francis (2003)
14. Armstrong, N., Hibbert, D.B.: An introduction to Bayesian methods for analyzing chemistry data: Part 1: An introduction to Bayesian theory and methods. *Chemometrics and Intelligent Laboratory Systems* 97(2), 194–210 (2009)
15. Robert, C.: *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Government Printing Office, U.S (2001)
16. Sivia, D., Skilling, J.: *Data Analysis: A Bayesian Tutorial*. Oxford University Press, USA (2006)
17. Heckerman, D.: *Bayesian Networks for Data Mining*. *Data Mining and Knowledge Discovery* 1(1), 79–119 (1997)