Computers & Industrial Engineering 106 (2017) 137–146

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Mining the relationship between production and customer service data for failure analysis of industrial products

Seokho Kang [a], Eunji Kim [a], Jaewoong Shim [a], Sungzoon Cho [a,b,*], Wonsang Chang [c], Junhwan Kim [c]

[a] Department of Industrial Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea
[b] Institute for Industrial Systems Innovation, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea
[c] Global Technology Center, Samsung Electronics Co. Ltd., 129 Samsung-ro, Yeongtong-gu, Suwon 16678, Republic of Korea

ARTICLE INFO

Article history:
Received 28 July 2016
Received in revised form 26 January 2017
Accepted 28 January 2017
Available online 7 February 2017

Keywords:
Data mining
Industrial product
Failure analysis
Product quality

ABSTRACT

Analyzing the causal relationships for failures of industrial products is necessary for manufacturers to prevent the occurrence of failures and enhance customer satisfaction. The data collected from each of the production and customer divisions can be a fruitful source for failure analysis. In this paper, we present a data mining process for efficient failure analysis of industrial products by a mashup of data collected from both divisions. The process consists of four main steps: problem definition, preprocessing, modeling, and visualization. Each step is designed to satisfy two constraints in order to be practically applied to industrial products. First, it has to be quick and incremental because the life cycle of most industrial products is not sufficiently long. Second, the insight derived from the process has to be easy to understand for domain experts since they are generally not familiar with data mining methodologies. A case study is conducted to demonstrate the effectiveness of the data mining process by using real-world data collected from a manufacturer in Korea.

1. Introduction

For industrial products such as consumer electronics and home appliances, failures often occur due to various reasons including insufficient inspection during production, low quality standards, and gradual changes of the products. This results in inconvenience and dissatisfaction of customers, which negatively affects customer loyalty (Reeves & Bednar, 1994). Furthermore, manufacturers would be subjected to the high costs of customer service including warranty and compensation. Therefore, preventing failures is directly associated with the profitability and survival of the manufacturer. In order to prevent defective products, it is important to examine the causes of failures that have occurred after production.

During the manufacturing process, the status of each product is monitored by a number of sensors, and the process parameters are adjusted accordingly. Once this process is completed, the final inspection is conducted to ensure the quality of the product. Only those that pass the inspection go on the market to be sold to a customer, while some of them are prone to be defective. When a defect is found in a product, the customer visits the customer service center to examine the symptom, determine the cause, and get it repaired depending on the symptom and cause. The repair history for each product is collected in the customer service center. Accordingly, plenty of data are accumulated in both the production and customer service divisions of manufacturers, which can be a fruitful source for failure analysis of the products.

Manufacturers have analyzed the data of their products in various ways (Choudhary, Harding, & Tiwari, 2008; Harding, Shahbaz, Srinivas, & Kusiak, 2006; Köksal, Batmaz, & Testik, 2011). The efforts to examine the cause of failures have been separately conducted by the production and customer service divisions. In the production division, the process measurements and inspection records generated during production are analyzed. A qualitative approach directly incorporates domain knowledge of experts regarding the failures (Venkatasubramanian, Rengaswamy, & Kavuri, 2003a), whereas a quantitative approach derives data-driven knowledge for predicting the inspection results using the process measurements and identifying crucial factors in the process (Köksal et al., 2011; Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003b). In the customer service division, statistical analyses of after-sales service records are conducted, such as the ratio of each failure type, trends in the occurrence of failures over time, and the distribution of the product lifecycle (Kalbfleisch & Prentice, 2011; Lawless, 1998; Mazhar, Kara, & Kaebernick, 2007).

* Corresponding author at: Department of Industrial Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.
E-mail address: zoon@snu.ac.kr (S. Cho).

http://dx.doi.org/10.1016/j.cie.2017.01.028
0360-8352/

These independent efforts are insufficient to provide meaningful insights of failures because the production and the customer service data have different information. The production data are related with the cause of failures, while the result of the cause is the occurrence of failures that are recorded in the customer service data. In order to enhance the significance of insights from failure analysis, the causal relationships of failures should be examined by a mashup of the production and customer service data. However, the mashup is difficult in practice because these data are independently collected and owned by each division. Moreover, identifying causal relationships from these data by a manual examination is an exhaustive and time-consuming task.

To address this problem, we present a data mining process for failure analysis of industrial products by analyzing the data from both production and customer service divisions. The causes of failures are explained by process measurements and inspection results from the production data, while the result of failures are recorded in the customer service data, as illustrated in Fig. 1. Thus, this process discovers causal relationships through data fusion to provide meaningful insights of failures. We focus on the following two aspects for an efficient analysis of the data with practical usability. First, quick and incremental analysis is required because the life cycle of most industrial products is not sufficiently long. Second, the analysis outcome should be easily interpreted by domain experts who are generally not familiar with data mining methodologies. Thus, the data mining process involves suitable guidelines for preprocessing and automation in modeling and visualization. In addition, an interface of bivariate visualization based on relative failure density is adopted to provide high interpretability. We conducted a case study using real-world data for home appliance products in order to demonstrate the effectiveness of the proposed data mining process.

The remainder of this paper is organized as follows. In Section 2, we review the related work. In Section 3, we introduce the data mining process, and describe the details of its steps. In Section 4, we report the case study results from real-world home appliance products. Finally, conclusions and future work are discussed in Section 5.

2. Related work

As large amounts of data are being collected throughout the product life cycle including design, production, sales, and use (He, Wang, He, & Xie, 2016), data mining has been successfully adopted for various tasks from in manufacturing industry in regard to their products, such as failure analysis (Choudhary et al., 2008; Harding et al., 2006; Köksal et al., 2011), product design (Chien, Kerh, Lin, & Yu, 2016; Jin, Liu, Ji, & Liu, 2016; Lin, Chien, & Kerh, 2016), and decision support (Xu, Sun, & Xu, 2016). Each of these tasks is investigated in a data-driven way by defining a data mining process with considering three aspects: practical, graphical, and analytical (Snee, 2015). The practical aspect concerns the context and subject matter knowledge regarding the task. The graphical aspect involves exploration of the data and analysis of the results through graphical visualization. The analytical aspect refers to the adequate use of quantitative and qualitative models for the data with practical considerations. Successful application of data mining process with high quality data would offer useful information of interest to the target task (Kenett, 2016).

As for data-driven failure analysis of products, many studies have focused on predicting quality factors in the production process including intermediate inspection results of each product and the production yield (Kim, Kang, Cho, joo Lee, & Doh, 2012; Kumar et al., 2006; Shin & Park, 2000; de Toledo, Freitas, Colosimo, & Gilardoni, 2015). They sought to obtain highly accurate prediction models by utilizing additional variables and complex learning algorithms. Accurate predictions of these factors allow to be extended to virtual metrology by mimicking metrology equipment with prediction models in order to achieve efficient

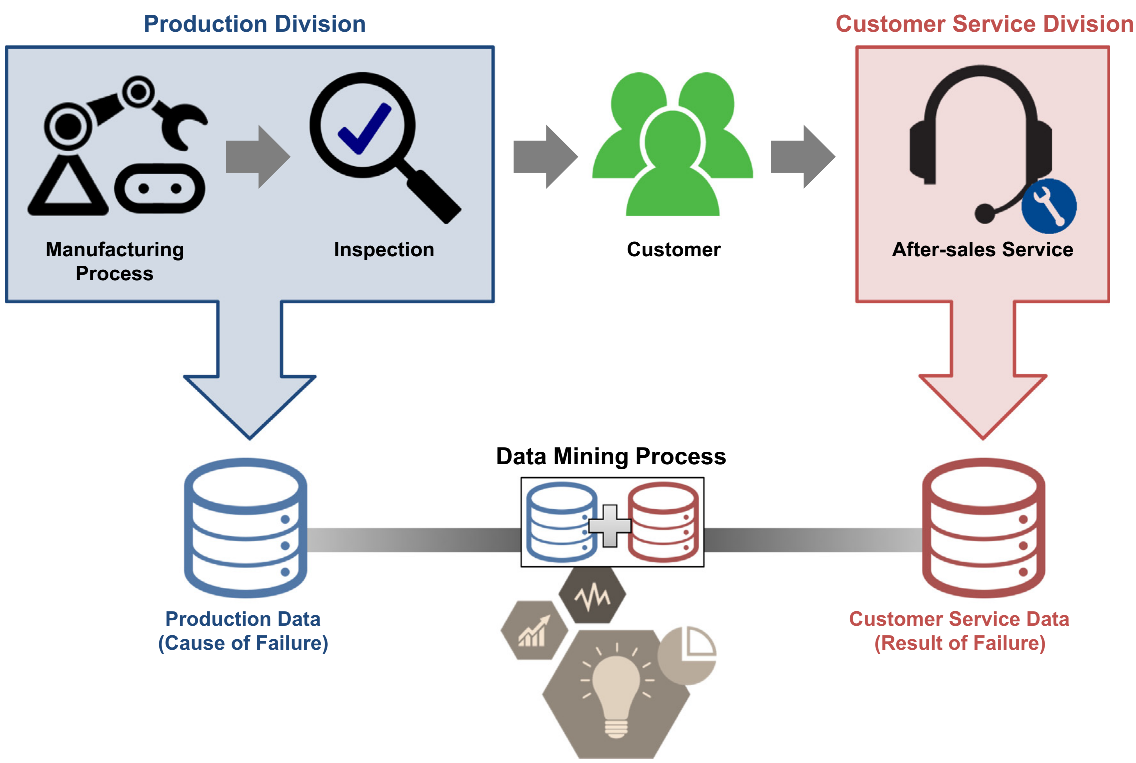

Fig. 1. Data mining process with fusion of production and customer service data.