Spare Parts Dispatch Fraud Detection Analysis

Shivangi Verma*, C. Rajendran

Shivangi Verma, Doctoral Student, University of Duisburg-Essen, Duisburg, Germany
C. Rajendran, Professor, Indian Institute of Technology Madras,India

Abstract

Along with warranty service abuse, dispatch frauds are a huge concern area for major IT companies. A spare part dispatch by a service provider to the customer is ideally complimented by a faulty part being returned back by the customer. Hence, every part that the service provider sends to a customer is followed by the customer replacing the non-functional part and sending it back. Unfortunately, the service providers lose over millions of dollars globally to non-returned parts and system exchanges. Neither are the companies tracking whether the parts or systems are being returned or not, nor are there processes in place to predict if a particular dispatch is fraudulent. A deep understanding of the dispatch process, with inputs from various stakeholders like Technical Support, Logistics Provider and Customer is required. The objective of the paper is to analyse dispatch frauds and develop a methodology that helps in avoiding and catching a fraudulent dispatch before it takes place, and hence, save millions of dollars in parts intake and sent. The steps followed to build this methodology are:

1. Recognizing and defining the metrics that help in identifying a machine on which a fraudulent dispatch has taken place
2. Carrying out primary analysis to gain a deeper understanding of the data and the fraud process as a whole.
3. Carrying out cluster analysis to group and characterize the features of such a machine and hence draw inference about possible machines on which fraud dispatch would have happened or can happen

The propensity, of a machine to commit fraud, is used to flag the machine and serves as an alert when they contact the service provider for another dispatch. Subsequently, very high risk machines can be blocked for no future dispatches.

Keywords: spare parts; fraud; dispatch; risk analysis; parts dispatch; replacement; field engineer; technical support; cluster analysis

1. Introduction

In today's world, loopholes and workarounds are easy to find even in the most secure of the systems and processes. Especially with the cheap and easy access to tools and technology, it has become far simpler for hackers and fraudsters to commit crimes. In turn, the same tools and technology can assist detectives to analyze and find these criminals.

One of the key challenges that most large IT organizations face in the times of a sluggish economy today is the balancing margins with customer satisfaction. Meanwhile, a number of entities, inside or outside the direct purview of the organization are investing concentrated efforts in duping the organization by millions of dollars by committing frauds – warranty, service or dispatch parts. Even though a large number or processes, manual or automatic, are usually in place to detect and prevent such frauds from taking place, it has still not been completely curbed.

In order to reduce such frauds from taking place, and at the same time, to maintain customer experience, proper checks at important points have to be implemented, metrics have to be defined, and continuous blocking of suspects and deeper investigation has to be carried out.

1.1. Parts Dispatch and Replacement Process

In a typical scenario of service parts dispatch, the customer contacts the service provider in a situation of a failed or malfunctioning part(s). The service provider's technical support wing communicates with the customer typically via email or telephone to diagnose the issue. In some cases, the technical support agent also tries to explain solution steps to the customer so that the part in the machine can be repaired by the customer himself in simple steps, hence avoiding the situation of sending an engineer to the customer's location.

In cases where the issue cannot be solved by the customer – because the technical support agent diagnosed the problem to be more complicated or requiring a replacement, the agent books a case or service request and sends the information across simultaneously to the field engineer team and the logistics team. The field engineer team allots a member for the case and passes on the customer and the issue details to him. Meanwhile, the logistics team sends the replacement parts to the customer location, where the field engineer can visit at a convenient time to repair the faulty part with these replacement parts. Sometimes, the field engineer picks up the replacement parts himself, and reaches the customer location. It becomes important that the agent has diagnosed the issue correctly, asking proper questions to the customer, since the replacement part that the logistics team sends to the customers is entirely dependent on this diagnosis exercise.

At the customer location, the field engineer examines the machine and the faulty part, carries out replacement with the new part, and takes back the faulty part. The field engineer has to send the faulty part to the repair center where the part is either repaired or scrapped.

Fig. 1. A typical parts dispatch process

1.2. Problem Statement

The service provider may face the following situations during a parts dispatch process:

- The customer contacts the service provider frequently, asking for replacement parts.
- The customer passes incorrect information about his name, address, phone number, machine number to confuse the service provider.
- The customer refuses to give back the faulty part to the field engineer, stating reasons like – the part has been thrown, the part has been already sent to the service provider etc.
- The machine number on which the replacement part has been ordered has not yet been sold, which implies, it is still in the retailer stock.
- The shipping address stated by the customer is different than the billing address.

This is not an exhaustive list. There might be various other such situations arising in front of the service provider. But the dispatch is still sent to the customer so as to avoid a hit on the customer satisfaction metric. Also, there is always a chance that the dispatch is a genuine dispatch and hence, to avoid a mistake with a genuine customer, the service provider does not hold any dispatch, and in turn loses millions of dollars on the frauds.

The current state and the desired state can be depicted as follows:

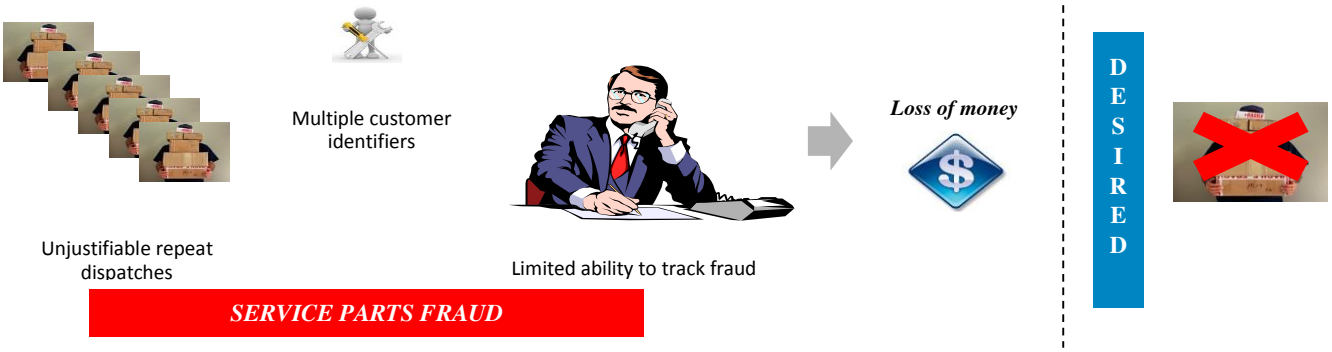

Fig. 2. A depiction of current and desired states in case on parts dispatch fraud

1.3. Defining the Metrics

The first step in the process of building an engine to predict dispatch frauds is to collect data. A suitable and trustworthy data source is chosen (which is specific to a service provider) to collect a list of all those dispatches which are confirmed as fraudulent. This dataset can be at the dispatch level, service request level or machine number level. It has to be noted that the machine number is a unique identifier for each machine that is produced. Hence, we now have a dataset with (for example) a machine number column and a flag column that indicates whether the machine, in its entire history, ever had a spare part dispatch that was fraudulent or not. For our analysis, we used the data at the machine number level.

This set of confirmed fraudulent machines is merged with a larger population of all the machine number that ever had a dispatch with them in the last 1 year. This gives us a large dataset to work with, and also compare the confirmed fraudulent machines with the non-confirmed machines and draw a contrast. So as to draw comparison between the two distinct sets of machines, it is required to quantify the process in the form of relevant and suitable metrics. The metrics decided to carry out the analysis are listed as follows:

- Number of dispatches: This refers to the count of distinct dispatches that have taken place on a machine, in its entire history. A high ‘number of dispatches’ corresponding to a machine indicates suspicious behavior.

- Rate of change of phone numbers: This metric is the number of distinct phone numbers by which a dispatch has been ordered for that machine divided by the total number of distinct dispatches for that machine. This metric indicates the frequency of change of phone numbers on a single machine. A high value of this metric indicates suspicious behavior.

- Rate of change of Email address: This metric is the number of distinct email addresses by which a dispatch has been ordered for that machine divided by the total number of distinct dispatches for that machine. Similar to the Rate of change of phone numbers, a high value of this metric indicates a suspicious behavior.

- Rate of change of postal code: This metric is calculated by dividing the number of distinct postal codes on which the dispatches have been delivered for that machine divided by the total number of dispatches. As mentioned above, a high value of this metric indicates suspicious behaviour.

- Frequency of dispatches: This is the average number of days between consecutive dispatches on a machine. A low number indicates a fraudulent behaviour.

- Age of the machine: This is the number of days from the time of purchase and use of the system till present (the day data was collected). Older the machine, higher are the chances of its parts failing.

- Major parts dispatched: This metric calculates the total number of high value parts dispatched on a machine in its entire history. The definition of high value is specific to the service provider.

- Segment: This depicts the segment of industry to which the customer belongs to – if it is a store retailer, an online retailer, direct sales etc.

For all practical purposes, we have restricted our analysis to a single country and a single commodity and the time frame taken for the analysis is 1 year.

1.4. Primary Analysis

After arranging the data in the form of rows and columns, with the metrics as the field items, primary analysis is carried out. This exercise is to understand the data in a comprehensive and detailed manner. This can be done in Microsoft Office Excel spreadsheet.

As a part of the primary analysis, 2 steps are followed:

- Create a 'Fraud Map': This is done by calculating the percentage of confirmed fraud cases in each city/region out of the total known fraud cases, and arranging the percentages in decreasing order. This gives a clear picture of the high-suspect regions and helps in catching a network that might be operational in the top 2-3 regions.

- One on one comparison of dispatches per machine – threshold setting: This step incorporates calculating the percentage of machines with confirmed fraud cases arranged in the increasing order of dispatches per machine. The same arrangement is carried out for the overall population all the non-confirmed machines. The dataset looks like below:
  The difference between the percentage values in each category gives an idea about the deviation in the characteristics of the confirmed fraud data and the overall population. As seen in the above figure, 20% of confirmed fraud machines had a single dispatch, while among the overall population of machines, 63% had a single dispatch. 17% of the overall confirmed fraud machines had 2 dispatches which is more than the corresponding percentage in the overall population (the difference being 4%). Moving on to the next row, 3 dispatches per machine, the percentage in the confirmed fraud is more than that of the overall population, and so on. It can therefore be inferred that the fraudulent machines are more likely to have equal to or more than 2 dispatches in such a case. Hence, 2 dispatches per machine is set as the threshold, and the cluster analysis is carried out for all the cases where the number of dispatches on each machines is equal to or more than 2.

1.5. Segmentation, Prediction and Inferences: Cluster Analysis

With the metric data of all the machines that ordered 2 or more than 2 dispatches, and the list of the confirmed fraud machines, the next step is to bucketize every machine into the risk segments. This is done using the cluster analysis techniques.

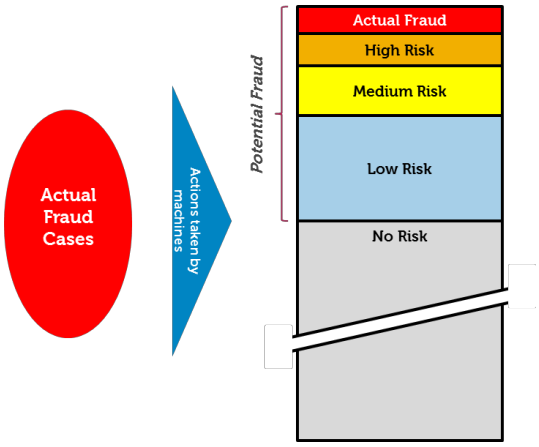

Fig. 4. Segmentation of overall population of machines into risk buckets using confirmed fraud cases

The data is fed into suitable software (R or JMP can be used to the purpose) and cluster analysis is run. The model gives results as follows:

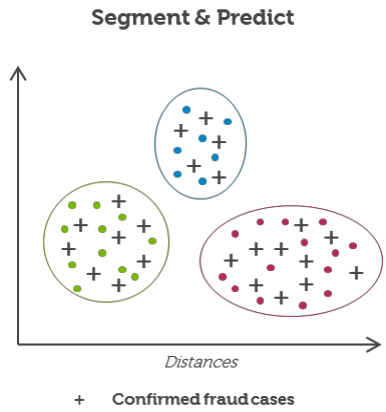

Fig. 5. Cluster Analysis for predicting fraudulent dispatch fraud cases – A depiction

As depicted in the above figure, the "+" sign denotes all the machines that are confirmed fraudulent cases, while the dots represent the machines in the overall population. There are two major points that can be inferred from the cluster analysis:

- The different segments will have different numbers of confirmed fraudulent machines. The segment with the largest number of fraud machines (the red coloured segment in Fig. 5) represents the high risk bucket, since this segment has the highest number of confirmed fraudulent machines.
- The propensity of a machine in the population to commit (or to have committed) a fraud can be calculated by measuring the distance from the nearest confirmed fraud case. Lesser the distance, higher is the probability of it being a fraud case. The distance can also be measured using the centroid of all the fraud cases in that segment

In this manner, all the machines that have already contacted the service provider can be analysed for their propensity to carrying out a fraudulent dispatch. These machines can be flagged and rated so that the next time they contact the service provider, further and deeper scrutiny can be implemented on a case by case basis. Machines with extremely high propensity can be blocked or banned from taking any more dispatches. Furthermore, any new machine ordering a parts dispatch from the service provider can be included in the analysis and the cluster analysis can be run, making this methodology more robust. Also, not only does this analysis cover the machines that have already contacted the service provider, it also makes way for the new machines.

This methodology can be replicated to all the regions/countries and for all the commodities that the service provider sends as spare parts dispatch.

References

(1) BDEC Limited : http://www.bdec-online.com/bd-codes/bd-p.cfm

(2) Benito F., David O., Christopher W.(1992): Judgmental adjustment of forecasts: A comparison of methods; International Journal of Forecasting, 1992, Vol.7(4), pp.421-433

(3) Boylan J., Syntetos A. (2009): Spare parts management: a review of forecasting research and extensions; http://imaman.oxfordjournals.org/content/21/3/227.full.pdf+html

(4) Eaves AHC., Kingsman B (2004): Forecasting for the ordering and stock-holding of spare parts; Journal of the Operational Research Society, 2004, Vol.55(4), p.431

(5) Effectively managing the multi-billion dollar threat from product warranty and support abuse, PricewaterhouseCoopers LLP, 2009.

(6) Fliedner and Lawrence (1995): Forecasting system parent group formation: An empirical application of cluster analysis; Journal of Operations Management, 1995, Vol.12(2), pp.119-130

(7) Fliedner E., Mabert V (1992): Constrained Forecasting: Some Implementation Guidelines; Decision Sciences, 1992, Vol.23(5), pp.1143-1161

(8) Flores B., Whybark D (1986): Multiple Criteria ABC Analysis; International Journal of Operations & Production Management, 1986, Vol.6(3), p.38-46

(9) Ghodrati B, Akersten P, Kumar U (2007): Spare parts estimation and risk assessment conducted at Choghart Iron Ore Mine; A case study; Journal of Quality in Maintenance Engineering, 2007, Vol.13(4), p.353-363

(10) Guvenir A., Erdal E. (1998): Multicriteria inventory classification using a genetic algorithm; European Journal of Operational Research, 1998, Vol.105(1), pp.29-37

(11) J. Ai, P.L. Brockett, L.L. Golden, Assessing Consumer Fraud Risk in Insurance Claims: An Unsupervised Learning Technique Using Discrete and Continuous Predictor Variables, North American Actuarial Journal, Volume 13, Number 4, 438-458.

(12) Klim T., Griffin M. (2005): Spare parts inventory management; http://www.google.com/patents/US7266518

(13) Moon S., Hicks C. and Simpson A. (2013): The development of a classification model for predicting the performance of forecasting methods for naval spare parts demand; International journal of production economics, 2013, Vol.143(2), pp. 449-454

(14) Ng I., Maull C., Nick Y (2009): Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science : evidence from the defence industry; European management journal : publ. twice a year for the Scottish Business School, 2009, Vol.27(6), pp. 377-387

(15) Shlifer E., Wolff R. (1979): Aggregation and Proration in forecasting; Management science : journal of the Institute for Operations Research and the Management Sciences, 1979, pp. 594-603

(16) Schultz C. (1987): Forecasting and Inventory Control for Sporadic Demand Under Periodic Review; http://www.jstor.org/discover/10.2307/2582735?uid=2&uid=4&sid=21104218776637

(17) Smith BT (1994): Focus Forecasting: Using computers for inventory control; Materials management in health care, 1994, Vol.3(9), pp.40, 42-5

(18) Strasheim JJ (1992): Demand forecasting for motor vehicle spare parts; A Journal of Industrial Engineering, Vol 6, No 2, December 1992, pp 18-29

(19) Widiarta H., Viswanathan S and Piplani R (2009): Forecasting aggregate demand: An analytical evaluation of top-down versus bottom-up forecasting in a production planning framework; International Journal of Production Economics, 2009, Vol.118(1), pp.87-94 [Peer Reviewed Journal]

(20) Willemain R., Smart N. and Schwarz F (2004): A new approach to forecasting intermittent demand for service parts inventories; International Journal of Forecasting, 2004, Vol.20(3), pp.375-387

(21) Zotteri et al (2005): The Impact of aggregation level on forecasting performance; International Journal of Production Economics, Jan 8, 2005, Vol.93-94, p.479(13)