Liquefied Petroleum Storage and Distribution Problems and Research Thesis

Marcin Gorawski[1,2], Anna Gorawska[1,2], and Krzysztof Pasterak[1,2]

[1] Silesian University of Technology,
Institute of Computer Science,
Akademicka 16, 44-100 Gliwice, Poland
{Marcin.Gorawski, Anna.Gorawska, Krzysztof.Pasterak}@polsl.pl
[2] AIUT Ltd.,
Department of Data Spaces and Algorithms,
Wyczolkowskiego 113, 44-100 Gliwice, Poland
{mgorawski, agorawska, kpasterak}@aiut.com

Abstract. The greatest threat to the environment and aquatic life is an uncontrolled fuel leakage, which is also extremely hazardous to health and safety of people. Guaranteeing the reliability of a leak detection system is probably the ultimate purpose of fuel management systems. However, there are more problems that ought to be solved before or simultaneously with detecting possible outflows of fuel products. In this paper we highlight major research opportunities consistent with wetstock management and statistical inventory reconciliation. The main goal is to outline thesis on the nature and impact of numerous phenomena on the inventory reconciliation methods. Issues considered in this paper include but are not limited to sensor miscalibration, data acquisition, and transmission problems as well as leak detection from both, tanks and connected pipeline.

Keywords: fuel leak detection, sensor miscalibration, petrol station, inventory reconciliation, statistical inventory reconciliation, leak from pipeline, data science, stream data warehouse

1 Introduction

Even a small inflow of liquid oil products or its water mixtures to the ground or aquifers, can easily damage the ecosystem in a manner that cannot be completely reversed [6]. In the event of an uncontrolled outflow, fuel can travel through underground rivers injuring the environment and aquatic life over a wide area. Moreover, it can jeopardize health and safety of people by contaminating drinking water supplies or through risks arising from its highly flammable nature. Therefore, to minimize the adverse effects of loss of containment, service stations must comply with numerous legislations governing operation and maintenance of each and every element in a facility infrastructure [2,3,17]. With costly clean-ups followed by legal penalties, the more accurate wetstock management and leak detection systems are, the more profits can be made. Consequently, earliest possible detection of fuel loss became central to successful storage and distribution of liquid fuels.

In modern service stations the problem of detecting an unaccounted fuel can be handled by built-in monitoring wells or sensors. Although purely mechanical solutions may indicate leak before any liquid enters the ecosystem [2], they are not risk-free. Another solution to the problem is performing manual leak tests periodically, which guarantee almost zero probability of leakage detection before an uncontrolled outflow occurs. Moreover, manually derived data is relatively imprecise due to human errors. Aforementioned approaches rely on mechanical or testing methods, while it is feasible to provide an efficient wetstock management system by means of constant statistical analysis of delivery, storage, and dispensing data. This base knowledge is a prerequisite for the inventory reconciliation [17], which this paper will be concentrated on.

The detection of an unaccounted fuel is not a trivial task – apart from abnormal fuel storage behaviours (e.g. leaks, sensors-related issues) there are several factors that have to be taken into account. Natural phenomena like evaporation or fuel thermal expansion may result in discrepancies between measured stock and expected values, which do not necessarily indicate leakage. Furthermore, numerous factors commonly overlap forming inventories of utterly complex data. Profound knowledge of all phenomena nature and impact on mentioned datasets is the key in distinguishing problems possibly resulting in disastrous effects from ones induced by the natural fuel behaviour.

All vessels storing liquids and pipelines may suffer from an uncontrolled outflow. The problem of anomaly detection was raised by researchers in the context of storage and distribution of liquids e.g. water, fuel, gas. In [15] classification and inventory reconciliation techniques were applied to leak control, while in [13] authors have used plastic optical fibre sensing. There is a group of pipe-leak oriented solutions [7], where e.g. fault detection clustering [14], fuzzy system classification [5], pattern recognition [20], support vector machine learning [4] or a frequency response diagram [12] were used to detect and locate leakage.

The issues of storage and distribution of liquid fuels are key challenges both from financial and environmental reasons. In this paper, anomaly detection challenges will be demonstrated with regard to operation of a petrol station infrastructure. However, the thesis and remarks in this paper can be easily transferred to all types of liquid fuel storing and/or dispensing facilities. Works carried out in this area are the subject of a project implemented by AIUT Ltd. in collaboration with the Institute of Theoretical and Applied Informatics of the Polish Academy of Sciences. The project is founded by the Polish Council of the National Centre for Research and Development within the *DEMONSTRATOR+* program [1].

Service Station Infrastructure

Each petrol station consists of the same elements: fuel tanks, pumping and distributing utilities, and dispensers with nozzles. During normal operation, when no malfunctions are considered, fuel is operating according to a following schema: delivery – storing – pumping – dispensing.

There are many types of connections between tanks and dispensers, more specifically – between tanks and individual nozzles. The basic criterion for division in accordance with [2] is how fuel is being pumped between infrastructure components. Vacuum or pressure systems cause suction or injection of fuel into the pipeline.

A pumping system, as well as the type of a storage, seemingly has no identifiable impact on inventory data. However, a pumping system reflects on the amount of fuel drawn from the tank into connected pipework. While station infrastructure schema available for the software vendor consists only of details on tanks and dispensers, length and configuration of pipework itself are not available. Therefore, piping should be considered a *black box*, with only inputs and outputs known. As a consequence, approximation of the amount of fuel present in piping is extremely difficult not only due to lack of aforementioned information, but because of the fuel thermal expansion phenomenon changing product volume, which was unknown from the very beginning.

Nevertheless, for a service station it is not always feasible to introduce modern, mechanical methods of fuel loss detection [17]. The adaptation to comply with legislations must be made without any changes to the petrol station infrastructure. Therefore, inventory reconciliation methods are popular alternative for the most common storage systems with underground, single-skin tanks and piping.

2 Data Analysis and Mining

As it was stated in the introduction, for most petrol stations it is not feasible to introduce modern, mechanical means of fuel management. Thus, measurements of stored, delivered, and dispensed product must become basis for surely statistical analysis. In this scenario, two main sources of data can be distinguished: tanks and dispensers, i.e. nozzles, which produce measurements carrying information about stored and sold fuel respectively. As a consequence, data representing a current state of mentioned facilities is commonly supplying a dedicated software application by two types of data streams. Third type of inventory records represents the amount of delivered petroleum. However, deliveries can be detected through analysis of tank volume – sudden and relatively large increase in value may indicate a delivery. Thus, such inventories are not shared with a wetstock management system vendor.

2.1 Tank Measurements

In modern stations Automatic Tank Gauges (ATG) [3] are increasingly used, thus, basic information on stored fuel ought to be drawn from product height. Then it is possible to transform measured fuel height to volume using an appropriate function $f \colon h \to v$, where h denotes height and v stands for volume. The

Table 1. Simulated tank measurements

Timestamp	TankId	Fuel Volume	Temperature	Water Valume
2014-01-01 07:00:00	3	28358,8912789761	15	0
2014-01-01 07:00:00	4	37536,4535629835	15	0
2014-01-01 07:05:00	1	9129,33206202403	15	0
2014-01-01 07:05:00	2	18850,602708773	15	0
2014-01-01 07:05:00	3	28309,2665066639	15	0

function is called *calibration curve* and it strongly depends on a shape and position of a particular tank and may change over time due to tank malformations. Therefore, from the analysis point of view, more unified data lies in the volume values.

The simulated measurements presented in Table 1 contain inventory data from four tanks. Fuel height was omitted in presentation, but temperature and water volume were taken into account. Obviously water presence in A tank is not desirable; therefore, commonly there are additional water level probes installed in tanks. Measuring water level is based on differences in liquids density. For simulation purposes a reference temperature of 15 Celsius degrees was used.

2.2 Nozzle Measurements

Records of dispensed product are represented mainly by the volume measurements with regard to a specific type of product, tank, and nozzles. Unlike in case of tanks, nozzle measurements made at a given moment of time are not simple to define. It has to be distinguished that values can be expressed as summarized volume of fuel dispensed through a nozzle (i.e. total counter) or information regarding current transaction only (i.e. transaction counter). Table 2 presents records obtained from the same simulation as mentioned in the previous section.

Table 2. Simulated nozzle measurements

Timestamp	NozzleId	TankId	Transaction Counter	Total Counter
2014-01-01 07:02:00	24	4	0	207,638426564652
2014-01-01 07:03:00	13	1	0	223,688673885022
2014-01-01 07:03:00	14	2	0	64,6584545488741
2014-01-01 07:03:00	15	3	12,4583333333333	60,0984271005254
2014-01-01 07:03:00	16	4	0	132,636625676526

2.3 Inventory Reconciliation

The single measurement in terms of fuel station management are not reliable sources of information. Only when a greater dataset is created, complex trends and phenomena may be identified. As a consequence, measurements are aggregated over fixed time windows and each represent the total amount of stored, dispensed or delivered fuel within single time period.

In best case scenario, fuel volumes associated with current stock and purchased volume are equal. Unfortunately, it is purely hypothetical situation that serves as a reference for real data analysis. Achieving the balance between delivered and sold fuel can be handled by means of the inventory reconciliation techniques [17,18,19]. It can be performed monthly, daily or every few minutes, where each mode designates an adequate aggregation time. In literature the result of the inventory reconciliation is called *variance* or *error* [15,17,18]:

$$Var = V_s - V_p - V_d \tag{1}$$

where V_s stands for the volume of sold product and V_p for the volume of fuel pumped from the tank during sale transactions. The last symbol, V_d, represents delivered fuel, i.e. the amount of product poured into the tank during a delivery. In such case, when delivery occurs the volume of fuel pumped out from the tank is less than 0. In the ideal scenario variance equals 0.

The value of a single error represents state of a service station in a given period of time. Long term analysis requires unified presentation of summarized variance values, which is served by a *cumulative variance* (CV). The CV is the basis for detecting many problems, e.g. sensors miscalibration and leakage. Analysis of a single phenomenon occurrence has to be held under the assumption that other are absent. Specific problems will be described later in this paper.

3 Sensor-Related Issues

The accuracy of measuring devices affects measurement correctness and, as a consequence, quality of data. Each petrol station is equipped with a massive range of sensors, which can produce corrupted datasets which are subsequently processed giving invalid results, e.g. in reconciliation. In this section we will focus primarily on the volume sensors installed in tanks and dispensers.

Tank Sensors Miscalibration. The visible symptom of the fuel sensors miscalibration is either over- or underevaluation of real fuel volume which in consequence leads to emergence of virtual surpluses or leaks. Moreover, situation is further complicated by a partial distortion of the calibration curve. The curve usually changes only in some particular segments, thus causing *the height-to-volume* transformation incorrect only within certain ranges. Therefore, the result of reconciliation may be correct only in some volume ranges.

An original form of the calibration curve is created during an initial tank calibration – starting with an empty tank, the tank is successively filled with exactly known amounts (volumes) of fuel and every time product height is being measured with a dipstick or an ATG. While manual measurements made with a dipstick is performed by a petrol station worker, accuracy is vulnerable to a human error. Moreover, during the initial calibration the tank should not be used for any other purpose than the calibration. Later, when tank position or shape changes, it is essential to perform the recalibration without pausing any system – preferably by relying solely on data and software solutions.

As described in the previous sections, reconciliation of a tight and well-calibrated tank should result in an error equal to 0. When the probability of any leakage is known to be 0 too, it is possible to utilize the CV (cumulative variance) [15,17,18] as a calibration quality criterion.

Nozzle Sensors Miscalibration. Measurements of dispensed fuel are made on the dispensers site, thus, the amount of sold product is expressed in volume for each nozzle separately. Unlike with tanks, where height is a basic criterion, we can omit the height-to-volume transformation as well as recalculation of calibration curve. However, measuring devices installed in nozzles, i.e. fuel-flow sensors, can also introduce errors to measurements. Generally sensors placed in nozzles ought to be calibrated every six months to assure measurements correctness. Practice shows that, mostly due to financial reasons, it is performed less frequently or only when major problems occur. Nozzle sensor miscalibration may lead to misrepresentation of returned volume values. In order to eliminate that pejorative impact a *calibration coefficient* is included in all calculations involving sold fuel volume. Usually it is a factor that the volume value is multiplied by. The calibration coefficient must be calculated for each nozzle separately and should be actualized, while sensor miscalibration may worsen with time.

A Probe Hang. Another issue is mechanical suspension of a probe floating on a fuel surface. In such condition, the fuel probe can be located either below or above the fuel level. In terms of data analysis, it results in production of invariable measurements, even when stored fuel volume is changing due to deliveries or sales. When reconciliated, such data creates virtual surplus – despite the normal sale, measurement from probes shows that fuel level tends to remain constant.

A Fuel Thermal Expansion. A natural factor complicating data analysis is a temperature. Given the same mass of stored fuel, its volume may differ due to temperature discrepancies. With the rise of a fuel temperature, the volume increases, thus, comparison of fuel volumes in different locations or time periods must be associated with a temperature compensation. A fuel expansion phenomenon is even more significant when liquids are mixing during a delivery – the product temperature changes rapidly before it compensates. The thermal expansion should also be considered in accordance with seasonal weather changes.

4 Leak detection

According to norms [2] and [3] five classes of leak prevention systems can be distinguished, where each concerns leaks from tanks and connected piping. Solutions consistent with classes from I to III and V use physical devices and a special infrastructure to detect anomalies consistent with an uncontrolled fuel outflow, whereas class IV systems rely sorely on the data analysis. That is why in older petrol stations sensors are used only to gather data during normal operation and adequate statistical methods are detecting leak-related incidents. Although our primary focus is set on class IV systems, as most petrol stations cannot be enhanced by an automatic leak detection, remarks on the leakage nature are present for any means of wetstock control.

Leaks from tank and piping. The difference between an outflow from a tank and a leak from the connected pipework is not only based on the physical localization of the leak, but it is also noticeable in the datasets collected from the station. When a leak from a tank is considered, the measured fuel volume decreases gradually during the whole time frame the leak was presented. A leak from piping results in an observable loss only in fuel sale intervals – such operation involves pumping and thus the fuel is present in the pipeline only when sale transactions occur.

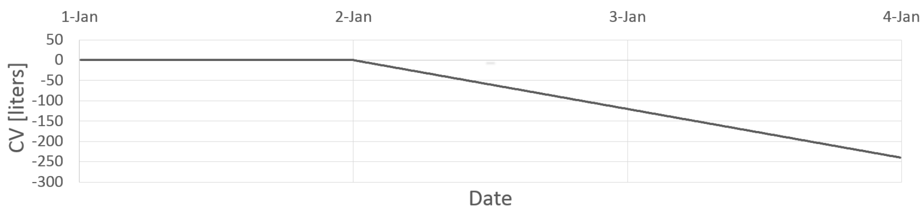

Fig. 1. Cumulative variance for leak introduced to the tank

Figure 1 presents a cumulative variance chart for a well-calibrated tank with a tightness problem. A leak of 5 litres per hour was introduced on the 2^{nd} of January and was present till the very end of the simulation. The cumulative variance chart is constant until that specified day, after which linear decrease in value is observed. In inventory reconciliation techniques negative balance between sold, dispensed, and delivered fuel may indicate leakage.

Leaks from piping are more difficult to detect – fuel loss is observable only when product is being sold. Therefore, on Figure 2 from the 3^{rd} of January, when leak (5 liters per hour) was applied, decrease in value is observable but it is not linear. Periods when cumulative variance is stable are consistent with no fuel being dispensed from connected tank, while decreasing trend denotes fuel loss during dispensing.

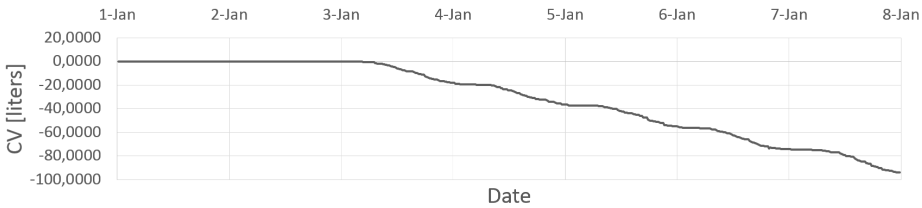

Fig. 2. Cumulative variance for leak from pipeline

More importantly, when intensity of customer arrivals is relatively low, leak from pipeline can remain undetected for a longer period of time. Moreover, even if such leakage is detected, it is not possible to specify its exact location. Structure and length of connected pipework is unknown, therefore with only statistical data it can be only indicated whether it occurs or not.

The other criterion of leak division is associated with an intensity of a leak. Two cases can be distinguished: constant and variable leaks. The former occurs when a leak rate is constant over time, while the latter is characterized by a variable rate. Usually it is a decreasing function of fuel volume as the amount of fuel stored in a tank has a direct impact on the pressure causing a leak. It is worth noticing that leaks with variable rate can only occur in tank because in pipes there is a constant pressure forced by a pump.

5 Data Acquisition and Transmission Problems

Besides previously mentioned problems, associated with different phenomena and anomalies that can occur on petrol stations, data can be affected with variety of distortions and malformations created during transmission or acquisition of the data. Most common issues of that type, which increase analysis complexity, are:

- missing data,
- data anachronism,
- data corruption.

Aforementioned issues are consequences of linking data sources with proper outputs through a transmission medium. In addition, all devices that generate data can also suffer from various disorders and produce corrupted data.

The first problem to consider is lack of certain data packages, which can be caused by sensor-related or networking issues. Such phenomenon can be quite effortlessly repaired, using interpolation techniques. It is applicable only when the missing part of data is relatively small.

The next problem is consistent with data packets being sent via multiple channels. It is possible for some of these packages to experience serious delays whereas the other (transmitted via a different channel) are delivered on time. This leads to a data anachronism. To neutralize potentially unsafe effects of this anomaly, two different approaches can be considered: wrong measurements can be rejected or input module can wait for a certain time interval to ensure that all delayed data has arrived. In the second method all measurements suffer delay equal to this interval.

The last phenomenon described in this section is data corruption. Due to different disturbances some bits of particular data packages can be changed causing malformation of the whole measurement. It can happen either on the data source site or during transmission. Nevertheless, wetstock management systems must be able to detect corrupted data and exclude it from its computations while they can violate further results significantly.

6 Proposed solution's architecture

The data from fuel stations are usually collected every few seconds and their amount depends on the number of tanks and nozzles. Complex wetstock analysis is often performed on data gathered from many stations at once, even the whole country area can be potentially considered. The stream nature [9,16] of fuel station data sources implies requirement of special tools capable of native processing of data streams. In terms of wetstock management following features are the most important when managing the stream source data:

- *real-time processing* – analysis oriented on the detection of mentioned storage and distribution problems ought to be performed as fast as possible for environmental and financial reasons,
- *the ability to process large data sets* – storing historical data along with relatively high frequency of generation; however, causes fast growth of size, it enables to perform advanced analysis concerning already determined issues as well as detecting periodic, subtle procedures, e.g. theft during deliveries,
- *handling data sources located within a large geographical area* – collecting data streams from a distributed measurement system [11],
- *quick access to data* – an accurate analysis may require data from any of time periods and on different aggregation levels.

When big datasets and complex analysis are considered, a Data Warehouse system is an obvious choice. Therefore, we propose stream-dedicated solution – *Stream Data Warehouse* (StrDW) [8,10]. An architecture of this specific system fully satisfies aforementioned requirements and adds many features supporting processing and maintaining of stream data, e.g. a stream ETL [8], a specialized processing engine StrSOLAP. Moreover, data streams supplied by the service station sensors have to be stored in a specialized data structure adapted to their continuous arrival specification. An StrMAL [10] engine is intended for fast access to current and historical data with regard to a selected aggregation level. This has a significant importance during data analysis involving trend and feature mining, as in the leak detection case.

7 Summary

Data discrepancies presented in the fuel volume inventories may be an indication of serious problems. Possible disastrous effects of issues highlighted in this paper can cause variety of economical and environmental losses. Discussed phenomena have either direct or indirect influence on the quality of data. The situation is further complicated by the fact that these factors are frequently independent from each other and may overlap. Therefore, the biggest challenge is the determination of problems source by a proper classification of the mentioned data discrepancies nature.

In this paper variety of research areas were discussed, i.e. tank and nozzles miscalibration, fuel probe hang, the thermal expansion impact on the volume measurements, the calibration curve malformation, leak detection as well as determination of water occurrence in a tank. The nature of each phenomena was presented with possible impact on the inventory reconciliation. This paper serves a purpose of a state-of-art support, while presented pool of knowledge is a very basis for successful management of storage and distribution of variety of liquids, especially fuels. To the best of our knowledge, there is no commercial solution providing unified environment for wetstock management with leakage detection and complex trend analysis that links current and historical data. Our future goal is to provide such unified system in a form of the Stream Data Warehouse with highly specified decision support and reporting services.

References

1. DEMONSTRATOR+ program, The Polish Council of the National Centre for Research and Development. `http://www.ncbir.pl/en/domestic-programmes/demonstrator/`.
2. EN 13160-1. Leak Detection Systems - Part 1: General Principles, 2003.
3. EN 13160-5. Leak Detection Systems - Part 5: Tank Gauge Leak Detection Systems, 2005.
4. H. Chen, H. Ye, C. Lv, and H. Su. Application of support vector machine learning to leak detection and location in pipelines. In *Instrumentation and Measurement Technology Conference, 2004. IMTC 04. Proceedings of the 21st IEEE*, volume 3, pages 2273–2277. IEEE, 2004.
5. H. V. Da Silva, C. K. Morooka, I. R. Guilherme, T. C. da Fonseca, and J. R. Mendes. Leak detection in petroleum pipelines using a fuzzy system. *Journal of Petroleum Science and Engineering*, 49(3):223–238, 2005.
6. S. Erkman. Industrial ecology: an historical view. *Journal of cleaner production*, 5(1):1–10, 1997.
7. M. Ferrante, B. Brunone, S. Meniconi, B. W. Karney, and C. Massari. Leak size, detectability and test conditions in pressurized pipe systems. *Water Resources Management*, 28:4583–4598, 2014.
8. M. Gorawski and A. Gorawska. Research on the Stream ETL Process. In *Beyond Databases, Architectures, and Structures*, volume 424 of *Communications in Computer and Information Science*, pages 61–71. Springer International Publishing, 2014.

9. M. Gorawski, A. Gorawska, and K. Pasterak. A survey of data stream processing tools. In *Information Sciences and Systems*, pages 295–303. Springer International Publishing, 2014.
10. M. Gorawski and R. Malczok. On efficient storing and processing of long aggregate lists. In *Data Warehousing and Knowledge Discovery*, volume 3589 of *Lecture Notes in Computer Science*, pages 190–199. Springer Berlin Heidelberg, 2005.
11. M. Gorawski, P. Marks, and M. Gorawski. Collecting data streams from a distributed radio-based measurement system. In J. R. Haritsa, K. Ramamohanarao, and V. Pudi, editors, *Database Systems for Advanced Applications, 13th International Conference, DASFAA 2008, New Delhi, India, March 19-21, 2008. Proceedings*, volume 4947 of *Lecture Notes in Computer Science*, pages 702–705. Springer, 2008.
12. P. J. Lee, J. P. Vtkovsk, M. F. Lambert, A. R. Simpson, and J. A. Liggett. Leak location using the pattern of the frequency response diagram in pipelines: a numerical study. *Journal of Sound and Vibration*, 284(35):1051–1073, 2005.
13. M. Morisawa and S. Muto. Plastic optical fibre sensing of fuel leakage in soil. *Journal of Sensors*, 2012, 2012.
14. P.-S. Murvay and I. Silea. A survey on gas leak detection and localization techniques. *Journal of Loss Prevention in the Process Industries*, 25(6):966–973, 2012.
15. M. Sigut, S. Alayón, and E. Hernández. Applying pattern classification techniques to the early detection of fuel leaks in petrol stations. *Journal of Cleaner Production*, 80:262–270, 2014.
16. M. Stonebraker, U. Çetintemel, and S. Zdonik. The 8 requirements of real-time stream processing. *SIGMOD Rec.*, 34(4):42–47, Dec. 2005.
17. United States Environmental Protection Agency. Standard Test Procedures For Evaluating Leak Detection Methods: Statistical Inventory Reconciliation Methods. Final Report, 1990.
18. United States Environmental Protection Agency. Introduction to Statistical Inventory Reconciliation For Underground Storage Tanks, 1995. `http://www.epa.gov/oust/pubs/sir.pdf`.
19. United States Environmental Protection Agency. Straight Talk on Tanks - Leak Detection Methods for Petroleum Underground Storage Tanks and Piping, 2005.
20. J. Zhang. Designing a cost effective and reliable pipeline leak detection system. *Pipes and Pipelines International*, 42(1):20–26, 1997.