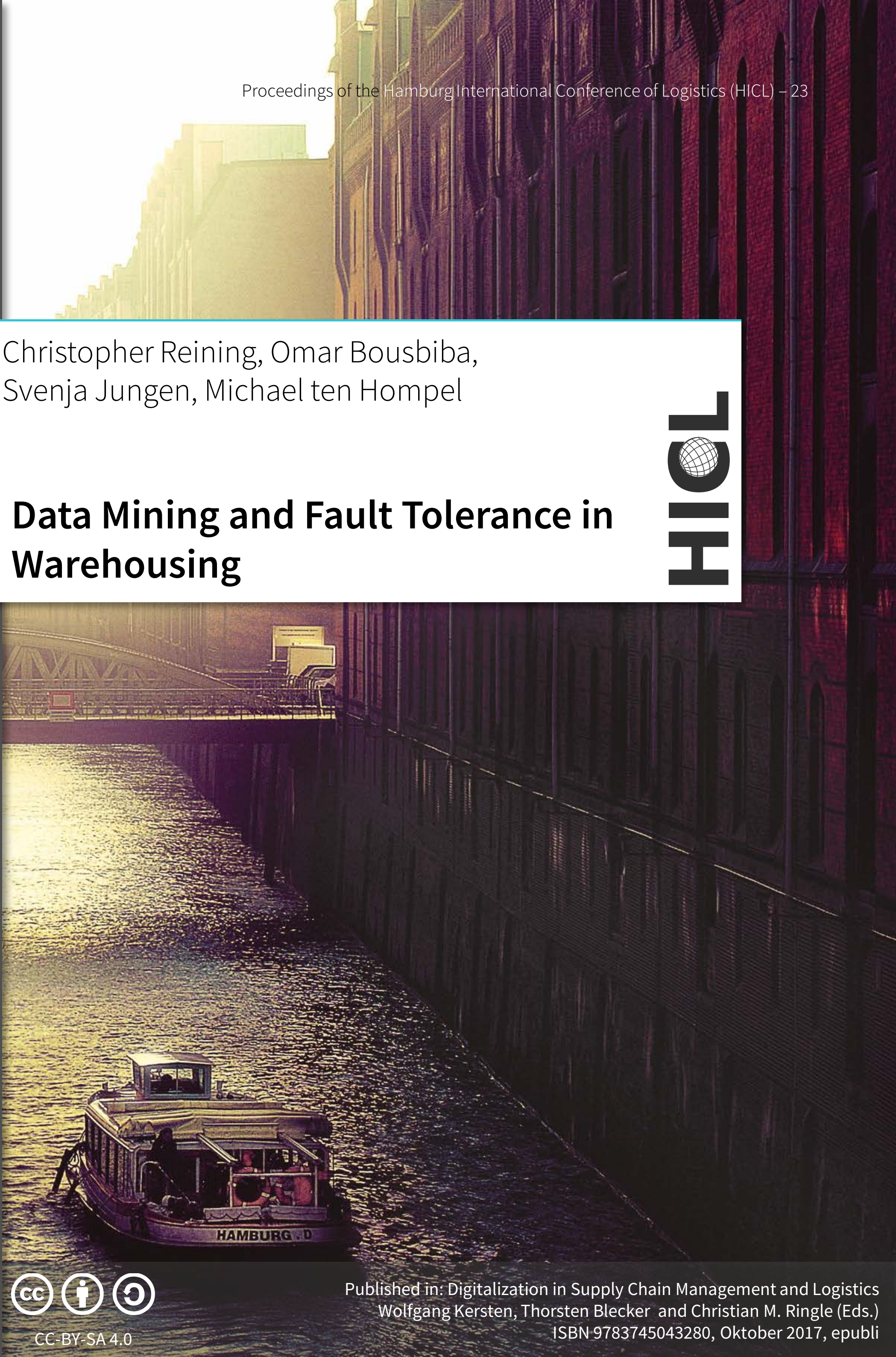

Proceedings of the Hamburg International Conference of Logistics (HICL) – 23

Christopher Reining, Omar Bousbiba,
Svenja Jungen, Michael ten Hompel

Data Mining and Fault Tolerance in Warehousing

Published in: Digitalization in Supply Chain Management and Logistics
Wolfgang Kersten, Thorsten Blecker and Christian M. Ringle (Eds.)
ISBN 9783745043280, Oktober 2017, epubli

Data Mining and Fault Tolerance in Warehousing

Christopher Reining[1], Omar Bousbiba[1], Svenja Jungen[1], Michael ten Hompel[1]

1 – TU Dortmund University

This paper surveys the significance of data mining techniques and fault tolerance in future materials flow systems with a focus on planning and decision-making. The fundamental connection between data mining, fault tolerance, and materials flow is illustrated. Contemporary developments in warehousing are assessed to formulate upcoming challenges. In particular, the transition towards distributed systems and the increasing data volume is examined. The significance of taking fault tolerance into account is emphasized. Ultimately, research issues are derived by conflating the previous findings. They comprise a holistic approach towards the integration of data science and fault tolerance techniques into future materials flow systems. Tackling these research issues will help to proactively harmonize the data representation to specific data mining techniques and increase the reliability of such systems.

Keywords: Materials Flow System; Data Mining; Fault Tolerance; Survey

1 Introduction

1.1 Increasing Data Volume and Data-Driven Operations

Nowadays the amount of available data in materials flow systems grows faster than the performance of computers to process them. Due to the implementation of superior sensors, more data is available for analysis. The application of distributed systems and current trends such as Internet of Things (IoT), Cyber-physical systems (CPS) and Industry 4.0 reinforce this trend (Xu, He and Li, 2014; International Controller Association, 2014; Hofmann and Rüsch, 2017). Smart devices are expected to decide and act autonomously or in collaboration with each other (Schuh and Stich, 2014, pp.203–213). Decentralization is seen as necessary in order to increase flexibility, reduce reaction time, and adapt to unplanned scenarios (Wilke, 2008). Accordingly, the area of decision-making broadens, input factors diversify, and standardized processes are fragmented or eliminated. Then again, there is a desire for predetermination when it comes to planning and controlling materials flow systems. Paradoxically, further information gathering and processing do not necessarily facilitate decision-making (Günthner and Ten Hompel, 2010, pp.2–5). Emerging data analysis methods are expected to create valuable information for both humans and machines to enable enhanced cooperation (Klötzer and Pflaum, 2015). Superior data processing shall relieve human workers by taking on recurring decisions autonomously.

In contrast to other scientific fields, there is little knowledge about data analytics in logistics so that further research is crucial (Rahman, Desa and Wibowo, 2011). Firms consider data and its analysis as a relevant resource to ensure their future competitiveness (Mazzei and Noble, 2017). At the current stage, data in warehousing systems is mainly used for anomaly detection and process control. To achieve additional benefit from the potential offered by the available data, processing methods for proactive optimization and prediction have to be enabled. Data-driven operations require relevant information to be obtained from raw data (Manyika et al., 2015). Their implementation connects human employees and their creativity in solving problems with state-of-the-art technology so that both can act ideally in real-time. This results in an environment which actively supports the cooperation of man and technology. The transition from warehousing as an isolated task into digital social networks is proposed by “Social Networked Industry” (ten Hompel, Putz and Nettsträter, 2017).

1.2 Rising communication complexity

From the findings of the previous subsection it can be concluded that distributed systems will be more widely used in future logistics facilities. In such systems, tasks are spread and solved in smaller groups of autonomous cooperating computing systems (Becker, Weimer and Pannek, 2015), that are also referred to as nodes. A distributed system can be described as a collective set of nodes, which interact with each other through message exchange. In contrast to centralized systems, there is no so-called "master-node", which holds the full control over all components in the system (Tanenbaum and Steen, 2008). Moreover, the application of a distributed system is often predetermined by the physical distribution of its components.

Each node has to be able to cooperate with each other node, e.g. by exchanging messages. All tasks have to be distributed and processed by each participating node in the distributed system. If several individual tasks exhibit time causal dependency or are limited due to physical restrictions, nodes have to be able to coordinate their tasks with each other. Two autonomous intelligent forklift trucks that try to take goods from a narrow alleyway are an exemplary case. If they are unable to coordinate, a deadlock might occur. A deadlock describes a blocking state of a system in which each participating node N_i blocks each other node N_j ($i \neq j$) by holding a lock on a particular object O_i (object O_i is locked by node N_i) of interest. Suppose that a group of cooperating autonomous forklift trucks tries to gather goods from a narrow alleyway. Further suppose that one of the forklift trucks crashes. As long as the fault is not detected, all forklift trucks may remain in a blocking state.

The coordination problem can be simplified when all fault-free nodes of a distributed system share the same global view of open tasks as well as a global view of the progress of accomplished tasks. Depending on their malfunction, faulty components can hinder the fault-free nodes from reaching a global view.

Due to the complexity of future distributed systems and the increasing reliability requirement (e.g., autonomous cooperating forklift trucks working in the same area as a human worker), it is important to make these systems resilient against faults. Otherwise, corruption (e.g., flash corruption due to supply voltage faults) of subcomponents will result in high costs and/or safety risks. This may cause the distributed system to enter an unsafe state. The problem of ensuring that fault-free nodes of a distributed system always share the same global view (reaching an agreement or consensus) independent of any presence of faulty nodes is also known as Byzantine agreement problem. The Byzantine generals' problem was first introduced by Lamport, Shostak and Pease (1982). With the help of agreement protocols, reaching a global view among fault-free nodes becomes possible. Byzantine agreement protocols have a long tradition. Thus, the state-of-the-art in the design of agreement protocols compromised a huge number of solutions (Lamport, Shostak and Pease, 1982; Jochim and Forest, 2010). Therefore, the fundamental problem of reaching agreement in the presence of faulty components can be seen as solved. Its unsolvability for different fault assumptions and/or system models has been proven by both Fischer, Lynch and Paterson (1985) and Santoro and Widmayer (2007).

Now, the challenge has become an optimization problem of the communication complexity (e.g., reducing the total number of communication rounds, number of redundant nodes and message transmission per communication round) which depends on the number of tolerated faults (Jochim and Forest, 2010; Khosravi and Kavian, 2012; Bousbiba, 2015).

1.3 Contribution and Outline

This paper contributes scientific hypotheses and questions to be evaluated in further projects in the field of materials flow systems, data mining and fault tolerance. It constitutes the necessary knowledge basis and sensitizes the ensuing issues from both an academic and an industrial point of view (BVL - Bundesvereinigung Logistik e. V., 2017).

The rest of the paper is organized as follows: The second section contemplates warehousing from a researcher's perspective. The link to knowledge discovery and the transition towards distributed systems are discussed. The third section further illustrates the significance of fault tolerance in this context. Finally, research issues are derived based on the previous findings.

2 Challenges in Materials Flow Systems

2.1 Warehousing Operations and Tasks

As shown in figure 1, warehousing employs three elements, namely systems, processes and management. Thus, a feasible way to formalize operations is to consider them as a combination of these elements. This approach implies a wide variety of configurations. In basic terms, warehousing processes are conducted within a system in accordance with an underlying management strategy. Typical planning and decision tasks interact with these operations. They comprise technology selection, capacity planning of personnel and equipment and packaging planning.

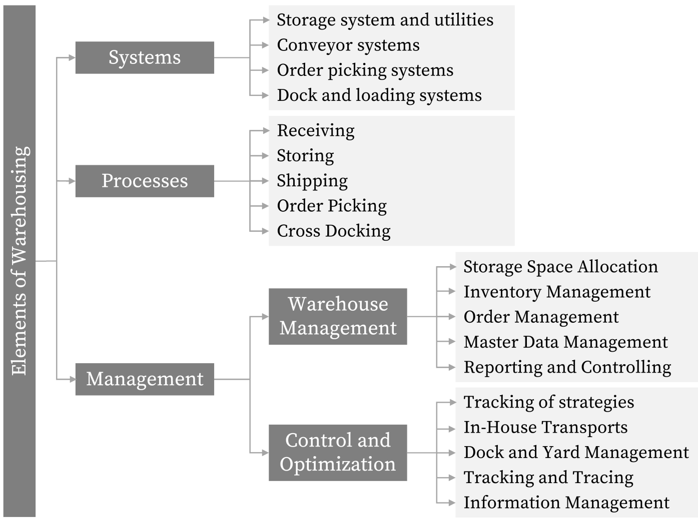

Figure 1: Elements of warehousing (Rohrhofer and Graf, 2013, pp.8–9)

As illustrated in figure 1, the storage space allocation of each article is defined by the warehouse management. It influences the effort of each order picking process, for example the travel time to corresponding storage locations. Likewise, the layout of the order picking system has a potentially major impact as it determines each shelf's position. More sophisticated warehousing principles cause more complex combinations of the elements. In a two-stage order picking system the interaction with the conveyor system would have to be taken into account. Then again, a conveyor itself is a electromechanical component. The aforementioned context emphasizes the interdependency of planning and decision tasks with warehousing. Other points of view, e.g. from maintenance or performance availability, reveal further ties.

On one hand, formalization of warehousing operations is desired in order to apply methods from other fields of science and represents a current research issue. On the other hand, task solving requires a high degree of domain knowledge. Running a warehouse is often characterized not only by interdependencies but uncertainties and rising task-complexity (Faber, de Koster and Smidts, 2013; Schieweck, Kern-Isberner and ten Hompel, 2016). These aspects hinder the application of analytical approaches. The required data, target figures and restrictions are depending on the operation's configuration and decisions already taken in related tasks.

Analytical approaches can be found in a wide range of technical standards and scientific works. They are limited to specific subsystems and use cases because warehousing is characterized by parallel and sequential operations that rigid formulas can hardly deal with. Materials flow simulation offers an alternative to deal with this issue. The commonly used discrete event simulation yet requires formalized operations and input parameters to obtain useful information (Timm and Lorig, 2015). Once the observed (sub-)system exceeds a certain level of complexity both analytical and simulative approaches tend to be infeasible to use in practical application (Roessler, Riemer and Mueller, 2015). They are either based on heavily simplified assumptions or they require a disproportional effort, e.g. specialized personnel or long preparation and running time of the simulation models (Frank, Laroque and Uhlig, 2013).

Applicable methods of data collection rely on observation and posterior analysis. In most cases, the conventional method of turning raw data into useful knowledge is done by manual analysis and interpretation (Bohács, Gáspár and Rinkács, 2012). Due to the necessary domain knowledge and insufficient formalization, the analysis and information extraction is conducted by specialists. While planning tasks usually offer a relatively long time frame, many decisions concerning warehousing are of short-term nature. For example, newly incoming goods in a distribution centre have to be assigned to one of several possible storage locations without a time lag. The assignment to a location is then performed based on a multitude of input factors of which some, e.g. the daily retrieval volume for the upcoming weeks, may not be available yet. Nevertheless, these figures may be predictable from past data. In such cases, warehousing operations can be expected to improve once they are data-driven.

2.2 Application of Knowledge Discovery and Data Mining

It is estimated that about 60% of data mining projects fail (Goasduff, 2015). Concerning warehousing, potential issues leading to this unsatisfactory situation are outlined.

The figure 2 below illustrates a widely recognized approach towards Knowledge Discovery in Databases (KDD) as proposed by Fayyad, Piatetsky-Shapir and Smyth (1996). Its nine steps are considered more closely to assess the applicability of the KDD process in warehousing. An overview of alternative process models is provided by Mariscal, Marbán and Fernández (2010). In literature, there is a wide variety of coexisting terminologies for data science related terms. Still, the drawn conclusions are valid regardless of the used terminology. A potential distinction of related terms is provided by Mitchell-Guthrie (2014).

The first step in the process is to develop an understanding of the observed system, gather necessary domain knowledge and identify the goal of the KDD process. While the warehouse operators may provide the necessary domain knowledge, goals can be derived from different sources. In warehousing there is a wide variety of key figures to examine. The goals may be set by technical staff or stipulated by the senior management.

Next, a target data set has to be created. For example, data about the past throughput volume or technical properties of the stored goods and their assignment to a storage location may be of interest. This data can be taken from the Warehouse Management System (WMS), Enterprise Resource Planning System (ERP) or other sources of information. Most IT-systems in use have evolved historically without regards to Knowledge Discovery. Human practitioners often generate article master data with emphasis on easy comprehensibility. Redundancy or noise

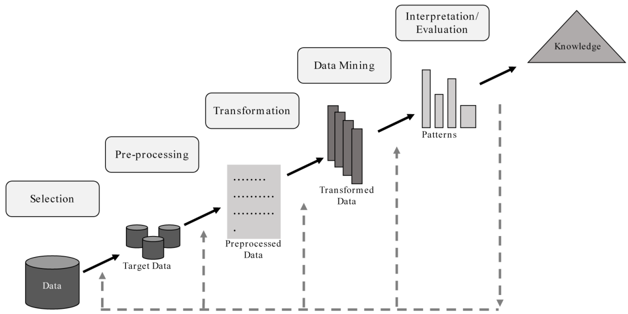

Figure 2: Overview of the KDD Process according to Fayyad, Pi-atetsky-Shapir and Smyth (1996)

avoidance has usually not been the priority when implementing the information systems.

The third step comprises data cleansing and pre-processing. This is necessary because the gathered data will very likely have inconsistencies, errors, out of range values, missing values and so forth. The goal of this step is to bring the data into a state which enables data mining methods to work as intended. While being time-consuming and labour-intensive, pre-processing facilitates the proceeding steps (Witten et al., 2017, pp.56–65). The same is true for the fourth step, data transformation. It comprises the search for useful features to represent the distinctive characteristics of the data. This procedure often goes along with dimensionality reduction. Data cleansing, pre-processing and transformation have to be conducted in accordance with the requirements of each data mining method. Accordingly, previous steps have to be repeated iteratively once an operation or an outcome turns out to be infeasible later on. The often given short-term nature of warehousing operations suggests that the utilization of data mining is vitally dependent on an efficient data preparation. An overview of pre-processing methods is provided by García, Luengo and Herrera (2016).

Data mining itself appears as the fifth step of the KDD process. A particular data mining method has to be matched with the predefined goal of the KDD process. The sixth step comprises the selection of a hypothesis and a data mining algorithm. The search for useful patterns and interesting information is conducted in the seventh step. It is crucial to recognize helpful data mining methods and algorithms for each warehousing task to exploit their potential. The chosen method and algorithm codetermine the optimal configuration of the overall KDD process. There are two thinkable scenarios but not limited to the following. Either the sheer amount of data is too big to be manually analysed, impeding data-driven decisions or the necessary information for making a valid decision is unknown beforehand.

While the first scenario is a topic of descriptive data mining, the second is a conceivable scenario of predictive data mining. Both scenarios may coincide in practice.

A common way to differentiate data mining methods is given by Fayyad, Piatetsky-Shapir and Smyth (1996):

— Regression
— Anomaly detection
— Association rule learning
— Clustering
— Classification
— Summarization

The data mining algorithm defines the model representation, the evaluation and the search method. For example, concrete algorithms of the regression method are, among others, the linear, polynomial and logistical regression. It is obvious that practitioners that choose to implement data mining methods are required to observe a multitude of possibilities. The question arises, which of the available methods and algorithms correspond to one or more warehousing tasks.

In the eight step the mined patterns are interpreted, e.g. by visualization. Previous steps may be repeated in further iterations of the KDD process. This is the case when the knowledge gain and its usability turn out to be insufficient. Often occurring short-term nature of warehousing operations suggests that time-consuming iterations are disadvantageous. They are ought to be avoided by a reasonable configuration of the KDD process in the first place. In the final step, the extracted knowledge is validated and made available for further processes.

To conclude this section, existing approaches to integrate KDD and data mining in warehousing are referenced. Ming-Huang Chiang, Lin and Chen (2014) presented an approach to improve the efficiency of order picking systems by finding associations between orders. The associations are the basis for heuristics to minimize the travel distance. Pang and Chan (2017) utilized data mining to find relationships among costumer orders. The identification of relationships helps to determine rational storage locations to minimize the picking effort. Gils et al. (2016) introduced an approach to predict the workload of order picking zones based on past data. The forecast can be used to determine the required number of employees and their assignment within the warehouse in a rational manner. Knoll, Prüglmeier and Reinhart (2016) presented an approach to use machine learning for predictive inbound logistics planning. It facilitates the system's adaptation to future scenarios. Their approach is based on the development of a logistics information structure (ontology).

The literature review shows that so far there is little effort spent on embedding data mining methods and algorithms into a holistic KDD process. They rather focus on finding suitable methods and algorithms for specific problems.

2.3 Distributed Systems in Warehousing

A central unit (master-node) controls the majority of current materials flow components. For example, a high-bay storage, a sorting system or a pick-by-light system are self-contained units in the sense that all information gathered and all decisions made within this system refer to a central unit. Therefore, warehousing operations are not triggered and conducted along several units. They are coordinated with help of the ERP, WMS or other central IT systems instead. To face the increasing flexibility demand, the distribution of warehousing operations among autonomous units is a current topic in both academic and industrial research. The underlying idea of realisation approaches is a plug-and-play behaviour of autonomous agents also referred to as ad-hoc-networking. Each agent is a modular unit (Lieberoth-Leden, Regulin and Günthner, 2016; Seibold and Furmans, 2017). In a distributed system, the warehousing system components are fluctuating. For example, the layout of the order picking zones cannot be assumed static once the facility is built. The system is repeatedly shifting its form in short terms. Agents may enter or leave the system for a multitude of reasons such as the required performance capacity, maintenance intervals and so forth. Furthermore, a single benign malfunctioning agent (e.g. fail-silent or crash) does not necessarily lead to a system crash, as it would be the case with a malfunctioning conveyor system in a high bay storage. Exemplary realisation approaches are referenced below.

Karis Pro (Colling et al., 2016) is an automated guided vehicle. Decentralized job creation without a prior teaching of transport connections between transfer points offers flexibility towards layout changes. Due to its modular construction and decentralized control, additional hardware like a central master computer are not necessary. Toru Robot (Kremen, 2016) is a warehouse robot that is able to pick items off shelves autonomously. In order to identify objects the robots create a map using laser sensors that scan the environment. This data is shared with other robots working in the same warehouse. The Grid Sorter (Colling, Seibold and Furmans, 2016) is a conveyor system which provides efficient and space-saving sorting of goods. The system's decentralized controllers and its structure consisting of identical rectangular conveyor modules allow flexible adaption to changing requirements. A reservation algorithm facilitates the transportation of goods, which are larger than a single conveyor module. The use of Cellular Transport Vehicles (Kirks et al., 2012) offers possibilities for areas where flexibility and changeability is required, planning reliability is not guaranteed or automation is not desired due to the lack of flexibility. The communication and control among the entities is realized by a decentral architecture. They are flexible in their topology meaning the formation of the entities is changeable at any time. Research and applications of distributed materials flow systems are neither geographically restrained to Europe (GreyOrange, 2017) nor solely focused on packaged goods (Serva Transport Systems, 2017). A further overview is given by Karabegović (2015).

The literature review shows that current research on distributed materials flow systems mainly focuses on technical issues regarding control and organization of the agents such as deadlock prevention or safety aspects (see e.g. (Kopecki, 2015)). So far, the data collection and processing, e.g. by the agent's environmental sensors, are subordinate to this purpose. Therefore, the data representation among agents from different suppliers is very likely not harmonized. As each agent is a source of data, the number and heterogeneity of data sources increases. Consequently, manual data analysis becomes increasingly impractical. The findings in this subsection further emphasize the urgency of proactively integrating knowledge discovery methods in future materials flow systems.

3 Fault Tolerance

The information in future distributed materials flow systems will be gathered from different locations and mobile systems. Using a centralized approach of data collection will be slow and expensive because all transferred information must be gathered by the master node. In other words, the entities and the master node need to be fully connected. Due to the communication bottleneck, costs and high communication overhead, it can be concluded that a distributed solution (i.e. a distributed knowledge discovery system) will be applied.

Data mining techniques in a distributed knowledge discovery system may help to improve the overall system by providing useful information to the group. This information (e.g., on several blocking areas) can be used by the autonomous robots to adapt their path planning in real-time in order to meet their constraints. Thus, the group will not reach a temporary deadlock.

In a distributed system, a global view has to persevered even if some of the components behave faulty. Otherwise, the extracted knowledge from the gathered datasets may result in faulty decisions. Another important aspect is the time in which a decision has to be made. Cooperating agents may rely on the knowledge extracted by the distributed knowledge discovery system in real time. This implies that the decision made by a distributed knowledge discovery system has a big impact on the effectiveness and efficiency of the overall logistics system (Dou, Chen and Yang, 2015).

Although the information provided by data mining can help to improve the effectiveness and efficiency of the overall logistics system, fault tolerance techniques must be applied nevertheless. The efficiency of large and/or complex distributed system heavily depends on the correct behaviour of its components and the decision made with the help of a distributed knowledge discovery system. For instance, Lee and Lin (2006) have shown that data mining techniques combined with fault tolerance techniques are able to cope with a certain degree of polluted data (containing a certain degree of noise, which – without using fault tolerance techniques – would lead to ambiguous conclusions).

4 Identification of Research Issues

In section 2.2 it has been stated that an efficient design and quick application of the KDD process require the data mining method and algorithm to be determined beforehand. Consequently, it is necessary to define a framework which allows the assignment of data mining methods and algorithms on one side to warehousing tasks and the given time horizon on the other side. An approach to this issue is given in figure 3. While data science provides methods and algorithms, warehousing tasks need to be formalized and abstracted to be assigned to them. The time horizon possibly restricts feasible methods. As illustrated in figure 3, the long-term packaging planning (e.g., the necessary amount of pallets or packages in a specific time frame) is a possible use case of regression. In contrast, storage space allocation as a short-term decision could be a use case for classification. In such a case, formalization becomes an issue. The classification of articles is not static over time. The retrospectively identified best classification is not necessarily the actually applied classification at the time the goods had been examined for the first time. This discrepancy is usually not shown in the provided data. A high degree of domain knowledge is necessary to retrospectively identify the best classification of an article.

The overview is far from being complete and serves as a starting point for further research. An identified match of a warehousing task and a data mining method does not necessarily mean that data science provides the best solution. Analytical or simulative approaches may provide better solutions in regards to accuracy and/or expenditure of time. Determining the limits of these approaches may reveal a variety of feasible use cases of data mining in warehousing for further examination.

In distributed systems, the harmonization of data representation among the agent becomes an issue. A standardized framework can take the requirements of efficient data mining into account. An industrial standard to be used by all suppliers of materials flow components may be the final goal of this approach.

The question arises in what way the necessary computing capacity for short-term data mining applications is provided. The data could be gathered and processed by a central unit, by each agent on its own (with sufficient computing capacity) or cooperative by a multitude of agents (distributed). As stated in Section 3 any form of cooperation requires a certain degree of fault tolerance, such as in a distributed knowledge discovery system and/or cooperative multi-agent system.

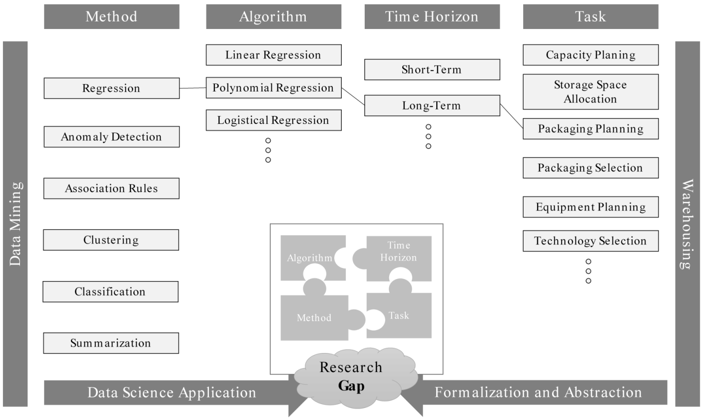

Figure 3: Research Gap between data mining methods and warehousing tasks

With respect to fault tolerant solutions in future distributed materials flow systems, the following research issues were identified. Designing an efficient Byzantine agreement protocol for different network topologies (Tanenbaum and Steen, 2008) is required as the complexity of materials flow systems will grow. In general, agreement protocols require a quadratic communication overhead to solve the Byzantine agreement problem. By relaxing the fault assumption (Jochim and Forest, 2010; Khosravi and Kavian, 2012) to simple faults (e.g., which cause fail silent behaviour) it is possible to solve the problem with linear communication overhead. However, such solutions cannot be applied in the presence of malicious faults (e.g., benign or malicious Byzantine faults).

One way to reduce the communication overhead is to design or deploy new efficient fault tolerant signature algorithms (Echtle, 1999). Many agreement protocols suffer from redundant message transmissions (Jochim and Forest, 2010). To provide fault tolerance, messages need to be retransmitted in recurrent communication rounds. It raises the question whether the design of a new signature algorithm may help to reduce the number of unnecessary retransmissions while simultaneously providing sufficient fault protection.

With the help of fault diagnosis algorithms and gossip-based membership protocols (Aljeri, Almulla and Boukerche, 2013) faults can be detected. Appropriate measures will be taken faster. As the detection of faulty behaviours without human interventions will play an important role in future logistics processes, such algorithms will help groups of cooperating nodes to remain functional in the presence of faults or breakdowns. For instance, if an intelligent forklift truck, which has reserved an area in order to complete its task, breaks down unexpectedly, it will block an area for a non-specific time. However, if the fault is detected by the other intelligent forklift trucks in the group, the area can be freed and the information about the faulty intelligent forklift truck can be further distributed to a sink, e.g. to a service engineer as well as to the knowledge discovery system.

5 Conclusion

The datasets being processed and analysed in future industry 4.0 warehousing applications will be tremendous. In order to extract useful knowledge from the gathered raw data, fault tolerant distributed knowledge discovery systems need to be applied. While there is a consensus that data science methods offer a great potential for planning and decision-making in warehousing, there is little knowledge about feasible use cases for data mining methods among typical warehousing tasks. This paper reviewed the status quo of data mining and fault tolerance. Crucial research issues have been derived and concrete approaches to address them have been outlined. In the Innovationslabor Hybride Dienstleistung, the Chair of Materials Handling and Warehousing at the TU Dortmund University plans to further work on this topic.

Acknowledgements

Parts of the work on this publication have been supported by Deutsche Forschungsgemeinschaft (DFG) in the context of the research project ”Human Activity Recognition in the Commissioning Process”.

References

Aljeri, N., M. Almulla, and A. Boukerche (2013). “An efficient fault detection and diagnosis protocol for vehicular networks”. en. In: ACM Press, pp. 23–30.

Becker, T., D. Weimer, and J. Pannek (2015). “Network structures and decentralized control in logistics: topology, interfaces, and dynamics”. In: *International Journal of Advanced Logistics* 4.1, pp. 1–8.

Bohács, Gáspár, and Rinkács (2012). “Adaptive simulation techniques for modeling material flow systems”. In: *2012 IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics*, pp. 559–564.

Bousbiba, O. (2015). “Reducing the Communication Complexity of Agreement Protocols By Applying A New Signature Scheme called SIGSEAM”. In: *DEPEND 2015 : The Eigth International Conference on Dependability*, pp. 29–32.

Bundesvereinigung Logistik e. V. (BVL) (2017). *Digitalisierung in der Logistik - Antworten auf Fragen aus der Unternehmenspraxis*.

Colling, D., S. Ibrahimpasic, A. Trenkle, and K. Furmans (2016). “Dezentrale Auftragserzeugung und -vergabe für FTF”. In: *Logistics Journal : Proceedings* 2016.10.

Colling, D., Z. Seibold, and K. Furmans (2016). “GridSorter – Dezentral gesteuertes Materialflusssystem zum Transport unterschiedlich großer Ladungsträger”. In: *Logistics Journal : Proceedings* 2016.05.

Dou, J., C. Chen, and P. Yang (2015). “Genetic Scheduling and Reinforcement Learning in Multirobot Systems for Intelligent Warehouses”. en. In: *Mathematical Problems in Engineering* 2015, e597956.

Echtle, K. (1999). “Avoiding Malicious Byzantine Faults by a New Signature Generation Technique”. In: *Dependable Computing ? EDCC-3*. Ed. by G. Goos, J. Hartmanis, J. van Leeuwen, J. Hlavi?ka, E. Maehle, and A. Pataricza. Vol. 1667. DOI: 10.1007/3-540-48254-7_9. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 106–123.

Faber, de Koster, and Smidts (2013). “Organizing warehouse management”. In: *International Journal of Operations & Production Management* 33.9, pp. 1230–1256.

Fayyad, U., G. Piatetsky-Shapir, and P. Smyth (1996). “From Data Mining to Knowledge Discovery in Databases”. In: AI Magazine, pp. 37–54.

Fischer, M. J., N. A. Lynch, and M. S. Paterson (1985). “Impossibility of Distributed Consensus with One Faulty Process”. In: *J. ACM* 32.2, pp. 374–382.

Frank, M., C. Laroque, and T. Uhlig (2013). “Reducing computation time in simulation-based optimization of manufacturing systems”. In: IEEE, pp. 2710–2721.

García, S., J. Luengo, and F. Herrera (2016). “Tutorial on practical tips of the most influential data preprocessing algorithms in data mining”. In: *Knowledge-Based Systems* 98, pp. 1–29.

Gils, T. v., K. Ramaekers, A. Caris, and M. Cools (2016). “The use of time series forecasting in zone order picking systems to predict order pickers’ workload”. In: *International Journal of Production Research* 0.0, pp. 1–14.

Goasduff, L. (2015). *Gartner Says Business Intelligence and Analytics Leaders Must Focus on Mindsets and Culture to Kick Start Advanced Analytics*.

GreyOrange (2017). *GreyOrange Butler*.

Günthner, W. A. and M. ten Hompel (2010). *Internet der Dinge in der Intralogistik*. ger. VDI-Buch. OCLC: 845595346. Berlin: Springer.

Hofmann, E. and M. Rüsch (2017). “Industry 4.0 and the current status as well as future prospects on logistics”. In: *Computers in Industry* 89, pp. 23–34.
Hompel, M. ten, M. Putz, and A. Nettsträter (2017). *Social Networked Industry*. Whitepaper. Fraunhofer IML, p. 10.
International Controller Association (2014). “Dream Car of the ”Ideenwerkstatt” at the ICV 2014”. In: pp. 1–59.
Jochim and Forest (2010). “An Efficient Implementation of the SM Agreement Protocol for a Time Triggered Communication System”. In: *SAE Convergence 2010*.
Karabegović, I., E. Karabegović, M. Mahmić, and E. Husak (2015). “The application of service robots for logistics in manufacturing processes”. In: *Advances in Production Engineering & Management* 10.4, pp. 185–194.
Khosravi, A. and Y. S. Kavian (2012). “Reaching an agreement in a distributed environment in absence of omission faults”. In: *2012 8th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP)*, pp. 1–4.
Kirks, T., J. Stenzel, A. Kamagaew, and M. ten Hompel (2012). “Zellulare Transportfahrzeuge für flexible und wandelbare Intralogistiksysteme”. In: *Logistics Journal* 2012.01.
Klötzer, C. and A. Pflaum (2015). “Cyber-physical systems as the technical foundation for problem solutions in manufacturing, logistics and supply chain management”. In: *2015 5th International Conference on the Internet of Things (IOT)*, pp. 12–19.
Knoll, D., M. Prüglmeier, and G. Reinhart (2016). “Predicting Future Inbound Logistics Processes Using Machine Learning”. In: *Procedia CIRP*. The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016) 52, pp. 145–150.
Kopecki, L. (2015). “An Approach for Deadlock Prevention in Decentralized Controlled Pre-storage Zones with Dynamic Behavior and Time Windows”. In: *International Journal of Materials, Mechanics and Manufacturing* 4.3, pp. 208–212.
Kremen, R. (2016). *Watch this new warehouse robot pick individual items from shelves*.
Lamport, L., R. Shostak, and M. Pease (1982). “The Byzantine Generals Problem”. In: *ACM Trans. Program. Lang. Syst.* 4.3, pp. 382–401.
Lee, G. and Y. t. Lin (2006). “A study on Proportional Fault-tolerant Data Mining”. In: *2006 Innovations in Information Technology*, pp. 1–5.
Lieberoth-Leden, C., D. Regulin, and W. A. Günthner (2016). “Efficient Messaging through Cluster Coordinators in Decentralized Controlled Material Flow Systems”. en. In: *MATEC Web of Conferences* 81, p. 06005.
Manyika, J., M. Chui, P. Bisson, J. Woetzel, R. Dobbs, J. Bughin, and D. Aharon (2015). *Unlocking the potential of the Internet of Things | McKinsey & Company*. Tech. rep.
Mariscal, G., Ó. Marbán, and C. Fernández (2010). “A survey of data mining and knowledge discovery process models and methodologies”. In: *The Knowledge Engineering Review* 25.2, pp. 137–166.
Mazzei, M. J. and D. Noble (2017). “Big data dreams: A framework for corporate strategy”. In: *Business Horizons*.
Ming-Huang Chiang, D., C.-P. Lin, and M.-C. Chen (2014). “Data mining based storage assignment heuristics for travel distance reduction”. en. In: *Expert Systems* 31.1, pp. 81–90.
Mitchell-Guthrie, P. (2014). *Looking backwards, looking forwards: SAS, data mining, and machine learning*.

Pang, K.-W. and H.-L. Chan (2017). “Data mining-based algorithm for storage location assignment in a randomised warehouse”. en. In: *International Journal of Production Research* 55.14, pp. 4035–4052.

Rahman, F. A., M. I. Desa, and A. Wibowo (2011). “A review of KDD-data mining framework and its application in logistics and transportation”. In: *The 7th International Conference on Networked Computing and Advanced Information Management*, pp. 175–180.

Roessler, M. P., J. Riemer, and M. Mueller (2015). “Decision Support for Choosing an Appropriate Simulation Method for Dynamic Material Flow Analysis”. In: *Journal of Industrial and Intelligent Information* 3.4.

Rohrhofer, C. and H.-C. Graf (2013). *Weissbuch der Intralogistik und Logistiktechnologie: Systemüberblick, Implementierung, Optimierung*. ger. OCLC: 854726145. Aachen: Shaker.

Santoro, N. and P. Widmayer (2007). “Agreement in synchronous networks with ubiquitous faults”. In: *Theoretical Computer Science*. Structural Information and Communication Complexity (SIROCCO 2005) 384.2, pp. 232–249.

Schieweck, S., G. Kern-Isberner, and M. ten Hompel (2016). “Intralogistik als Anwendungsgebiet der Antwortmengenprogrammierung – Potenzialanalyse”. In: *Logistics Journal : Proceedings* 2016.05.

Schuh, G. and V. Stich (2014). *Enterprise-Integration: auf dem Weg zum kollaborativen Unternehmen*. ger. VDI-Buch. OCLC: 879580815. Berlin: Springer Vieweg.

Seibold, Z. and K. Furmans (2017). “Plug&Play-Fördertechnik in der Industrie 4.0”. de. In: *Handbuch Industrie 4.0 Bd.3*. Ed. by B. Vogel-Heuser, T. Bauernhansl, and M. t. Hompel. Springer Reference Technik. DOI: 10.1007/978-3-662-53251-5_2. Springer Berlin Heidelberg, pp. 3–20.

Serva Transport Systems (2017). *EVE - Effiecient Variable Ergonomic: The driverless transport system for intralogistics*.

Tanenbaum, A. S. and M. v. Steen (2008). *Verteilte Systeme: Prinzipien und Paradigmen*. ger. 2., aktualisierte Aufl. it-informatik. OCLC: 237201545. München: Pearson Studium.

Timm, I. J. and F. Lorig (2015). “Logistics 4.0 - A challenge for simulation”. In: *2015 Winter Simulation Conference (WSC)*, pp. 3118–3119.

Wilke, M. (2008). “Dezentral steuern, zentral kommunizieren - ein Steuerungskonzept für wandelbare Materialflussysteme”. In: *Logistics Journal : referierte Veröffentlichungen* 2008.April.

Witten, I. H., F. Eibe, M. A. Hall, and C. J. Pal (2017). *Data mining: practical machine learning tools and techniques*. Fourth Edition. Amsterdam: Elsevier.

Xu, L. D., W. He, and S. Li (2014). “Internet of Things in Industries: A Survey”. In: *IEEE Transactions on Industrial Informatics* 10.4, pp. 2233–2243.