POTENTIALS AND PROBLEMS OF RFID-BASED COOPERATIONS IN SUPPLY CHAINS

Nico Schlitter [*] **Florian Kähne** [*] **Stiefen T. Schilz** [*] **Holger Mattke**[*]
forename.name@iti.cs.uni-magdeburg.de

Group on Knowledge Management & Discovery
Institute of Technical and Business Information Systems
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
PO Box 4120, D-39016 Magdeburg, Germany

Abstract

This article deals with RFID potentials related to a shift from local to cooperative optimization. We present the resulting problems caused by different interests within the supply chain. Therefore we define four typical stakeholders, describe their relationships and interactions and figure out possible solutions. Finally, a practical application supports the described approaches. We present a RFID-based monitoring along the supply chain for prediction of manufacturing errors. The upcoming privacy concerns of the involved companies are smoothed out by new privacy preserving data mining methods.

Keywords: cooperative optimization, cooperation, supply chain, RFID, data mining, privacy preservation, manufacturing error prediction

---

[*] The work is partially done within the Ko-RFID project which is part of the program next generation media funded by the German Federal Ministry for Economics and Technology.

1 Introduction

Over the last few years, Radio Frequency Identification (RFID) technology has gained more and more influence on the area of supply chain management. Some pioneering companies such as Metro (see RFID Atlas 2006) or Wal-Mart (see Ostler 2004), have successfully integrated this promising technology in their business processes. Indeed, the capability of solid-state reading and writing of data on RFID tags gives raise to numerous useful applications. Thus far, the focus has been mostly on the optimization of internal processes. In this article, we advocate the investigation of cooperative optimization potentials across the supply chain. We discuss the shift from local optimization scenarios to cooperative optimization scenarios and elaborate on the potential and problems associated with this shift.

Here, local optimization refers to the improvement of a company's local business processes by using just the local knowledge. Whereas within the local optimization no other knowledge than the local one is used the cooperative optimization includes also knowledge of other companies in the supply chain. In this article we present the potential of this cooperative optimization by an example in the context of monitoring across the supply chain for the identification of manufacturing errors.

The remaining sections of this article are organized as follows. In section 2 we compare the local optimization with the cooperative optimization. We show the necessity of cooperation and how RFID supports it. The resulting problems caused by the different interests of the stakeholders are described in section 3. Furthermore we outline possible solution strategies for the resulting cooperation problems. In order to illustrate these theoretical approaches we give a practical application for such cooperative optimization scenario. The application, the resulting problems and existing solutions are presented in section 4. Finally, section 5 concludes the article.

2 RFID-based decreasing of warehousing costs

As mentioned before some pioneering companies have started to integrate RFID in their business processes. Nevertheless, most of the introduced RFID applications merely aim to local optimization. The following exemplary optimization scenario shows the advantages of cooperative optimization compared to a local optimization. The solutions presented below are already realized in practice and mentioned here in order to motivate the article.

As a result of the soaring globalization and the increasing competition companies are forced to optimize their cost structures. In the following we look into the decreasing of warehousing costs, which arise of the storage of goods. Amongst others the total warehousing costs consist of:

- the costs for the warehouse (e.g. rent or writing down)
- the costs resulting from the gap between stored inventory and customers' demand (e.g. missing interests of the money tied up in stocks)

In the local optimization scenario we use exclusively local information in order to minimize the warehousing costs per item. We decrease the costs per item by increasing the warehouse capacity whereas the total warehousing costs are kept constant. The example in section 2.1 shows an RFID-based *Random Location Storage* to achieve this aim. Afterwards we illustrate the additional advantage of using global information in a cooperative optimization scenario. We point to the *Vendor Managed Inventory* (VMI) system which is able to reduce the costs of oversized stock of inventory.

2.1 RFID-based Random Location Storage

The choice of the method which is used for assigning items to a physical storage space affects the system performance of a warehouse. Thus, many corporations handling with stock keeping units deploy a randomized storage, or floating slot system. Unlike the so-called dedicated storage system with fixed assignments for each item, randomized storage places items to different locations within the warehouse according to an assignment rule. That is, items are not categorized and can be allocated to each vacant space. A widespread and often used storage location assignment rule is to store incoming items at the closest open location possible. However, alternative rules such as storing at the location that can be reached most quickly are used in practice as well (see Altassan / Malmborg 1998). With randomized storage, total space requirements are tried to be reduce. It is often applied if demand tends to be non-stationary (see Pfohl 2004). Deploying randomized storage necessitates clearly identifiable storage spaces within the warehouse. Frequently, this is realized by zoning the entire warehouse (see Martin 2000). In addition to the extreme policies of randomized and dedicated storage, there are variations of these policies such as class-based storage and shared storage (see Altassan / Malmborg 1998).

Retrieval efficiency is the most important advantage of dedicated storage systems. That is, since each item is permanently located at the same place, complexity and, hence, retrieval costs are much lower. Yet, if savings in total space requirement in a randomized storage system are sufficiently high, retrieval costs can be even lower compared to using dedicated storage (see Malmborg 1996). As previously mentioned, warehouse capacity utilization is more efficient with randomized storage systems. Unlike in dedicated storage systems, warehouse space for each item is not determined by the maximum expected amount of items that ought to be stored at the same time. However, a drawback of randomized storage is that it cannot be applied with each type of commodity. For instance, various chemical products may not be stored very close to each other due to safety aspects (see Stadtler 1998). Moreover, for randomized storage systems being more complex and less concise, a highly sophisticated warehouse management system is required in order to control all incoming and outgoing items. That is, the assignment of an item to a vacant space is done entirely by the warehouse management system. It records which item is stored in which quantity at which location. It is obvious that a system breakdown or even a loss of all storage data costs tremendous time and effort to continue warehouse proceedings (see Pfohl 2004).

Common randomized storage systems are working with barcode techniques. The items equipped with a barcode are registered when entering the warehouse and then assigned to the next space available. However, identifying the desired item in the warehouse might be difficult due to the direct contact that must necessarily be established between the barcode reader and the barcode on the item. This is especially true if items are stored in a high bay racking (see FKI Logistex 2005).

The implementation of RFID tags on each item can lead to enormous improvements when storing items in randomized storage systems. Each item is equipped with a RFID tag instead of a barcode label. This enhances the process of locating items due to faster identification. RFID readers can be placed on lift trucks in order to rapidly identify products even if they are stored in a high bay racking. Additionally, failure rates in the identification process are reduced. The implementation of RFID chips improves accuracy. Hence, warehousing costs are lower with RFID than with barcode. According to a FKI Logistex study, these benefits can lead to a cost reduction of up to 28 % (see FKI Logistex 2005).

Several companies have already taken advantage of the RFID technology. At Adler Modemärkte GmbH, a German garment manufacturer and distributor, RFID improved the process of order picking since items can be identified faster and more accurate than doing it manually which was previously practiced (see Adler Case Study 2006).

In this section we have demonstrated the capability of RFID to optimize a local process by using local information. We show how RFID can be used to increase the warehouse capacity by realizing a space saving random location storage system. In this manner, the warehousing costs per item are decreased and in addition the efficiency and accuracy is increased. Nevertheless the main problem of an inappropriate high inventory is not solved, because the necessary information is not available locally. The next section presents a cooperative optimization where global information is used to solve this problem.

2.2 Cooperative optimization by Vendor Managed Inventory

Mostly, the amount of orders from a downstream member in a supply chain is the only valuable information to calculate the own production or inventory level. However, the amount of these orders distinguishes from actual sales. Moreover, this distortion of demand enlarges with each tier upstream in the supply chain. The increasing differences in demand in the form of orders lead to incorrect demand forecasts, high inventory, and a low service level. This phenomenon, termed *bullwhip effect*, was first analyzed by Forrester (1958) and experimentally illustrated by Sterman's "beer distribution game" (see Stadtler / Kilger 2002). The following passage is based on the remarks of Lee, Padmanabhan and Whang (1997).

The reason for the bullwhip effect is merely stationary information within the supply chain. Only the distributor which is directly connected to the ultimate customer knows the actual demand. Besides this lack of information four subcategories have been analyzed:

1. Demand signal processing forces the members of a supply chain to rely on past orders from downstream members in order to calculate their own amount of production or orders to the upstream member.
2. A second cause for a high volatility in orders is the rationing game that describes the strategic behavior of a player in case of stockouts. As a consequence all upstream members will overreact and orders at each stage exceed the actual demand.
3. The third reason is the practice of order batching. The combination of orders of different periods reduces transaction costs, but lengthens the time between orders and information will be transferred slowly. This results in an increase of the uncertainty of orders and the necessity to order more than is actually needed (see Axsäter 2000).
4. Price variations such as discounts also may induce members to order more than they need.

All these practices lead to excessive ordering and therefore extent variations in the amount of orders and increases inventory. In order to smooth fluctuations in orders all partners of the supply chain should be able to receive all the data they need for the purpose of calculating the optimal amount to order. This could be achieved via Electronic Data Interchange (EDI) which reduces, besides other costs, the amount of stock. However, enabling information interchange in order to eliminate uncertainties is not sufficient if partners are not willing to share their information.

A more centralized decision making such as Vendor Managed Inventory (VMI) could support the reduction of uncertainty in demand while reducing the lack of trust. VMI is a form of partnership where the customer within a supply chain does not have to order, because his supplier fills his stock up to a level which was formerly mutually agreed upon. While forecasting the demand and planning the partner's inventory the supplier also has to calculate his own inventory level (see Stadtler / Kilger 2002). If one company, ideally the manufacturer, centrally plans for all other members, VMI is a concept from which the whole supply chain will benefit (see Chopra / Meindl 2007). Obviously, reduction of uncertainty helps the manufacturer to produce the amount that fits with the ultimate customer demand and protects against shortages and other errors. The supplier's customer benefits from these decreased shortages and this results in an increase in sales due to more customer loyalty. Additionally, fewer inventory and less forecasting facilitates operations and brings costs down. In order to benefit from all these advantages it is important to build up a good partnership between the supplier and his customers supported by mutual trust as an encouragement for data interchange (see Kumar / Kumar 2007).

Previous technologies are not sufficient to enable *real-time information sharing* between manufacturer and its customers. RFID can be a supportive tool for the retailer to collect all the required information. With the use of RFID retailers are less occupied with forecasting and purchasing and are more focused on service (see IBM case study 2007). In case of VMI products and pallets could be tagged with transponders so that the products can be identified directly when arriving or leaving the warehouse. It is also possible to install a reader at the transit between warehouse and sales room. Products are identified in the moment when they are pre-positioned to be stored in the shelves. This is a direct indication of customer demand, because they will be sold in the near future. Moreover, even the shelves can be equipped with RFID readers and the amount of stocks located in the shelves can be observed (see Thiesse / Fleisch 2007). Finally, the data obtained through RFID are sent to the manufacturer (e.g. via Extensible Markup Language interfaces) so that he can plan his production (see Hansen 2007). Consequently the manufacturer is enabled to decrease his inventory as a result of raising certainty.

Here, we have exemplary shown the advantage of cooperative optimization in a scenario of minimizing warehousing costs. In this case, it is possible to decrease the manufacturer's inventory and consequently the warehousing costs by using non-stationary, global information. However, it is obvious that this kind of optimization needs cooperation and trust between supply chain companies because the companies are reluctant to share sensitive information. Several reasons, why this necessary cooperation doesn't exist naturally, are presented in the next section.

3 Typical stakeholders within RFID-based cooperation along supply chains

In this section we analyze the different interests and strategies of stakeholders faced with the need to cooperate within supply chains. In general a certain stakeholders will participate in the cooperation process if expected benefits from participating are significantly higher than its costs. In the following we focus on different interests and intentions stakeholders may have and which in turn define the strategy they will follow.

Several approaches can be found in literature to describe the generic market strategies of companies. According to Miles and Snow (Miles/Snow 1978) four types of companies can be distinguished: defenders, prospects, analyzers and reactors.

Defenders base their strategy on existing markets and products and focus on improvements of such products.
Prospects consequently develop new products and try to explore new markets.

Analyzers try to combine both, defenders' and prospects' approaches.
Reactors have no clear strategy for their future development. They are affected by external forces resulting from environmental changes.

Even though the focus of this article is not on market strategies we see a strong relation to cooperation along the supply chain supported by the adoption of RFID as a new technology. This cooperative RFID adoption goes along with significant changes of business processes and interorganizational relationships and therefore enables the joint development of new kinds of services and products.

Recently a lot of research interest could be observed in the field of interorganizational systems (IOS). The development of standards for these IOS is also a big challenge in the context of managing RFID data collected over the entire supply chain. Löwer (2006) gives insights into how actors should coordinate the process of IOS standardization according to their special interests. He discusses ways in which an individual company could participate in standard developing activities depending on the benefits this company expects to obtain and the interests of other participants, that may be aligned or divergent.

To clarify the different forces in the process of cooperative RFID adoption, we slightly adapt the classification of Miles and Snow (1978) combined with Löwer's (2006) approaches in order to define the following four roles of stakeholders and their generic strategies: initiators, adopters, observers and defenders. Figure 1 distinguishes between the stakeholders regarding their interests and expected benefits.

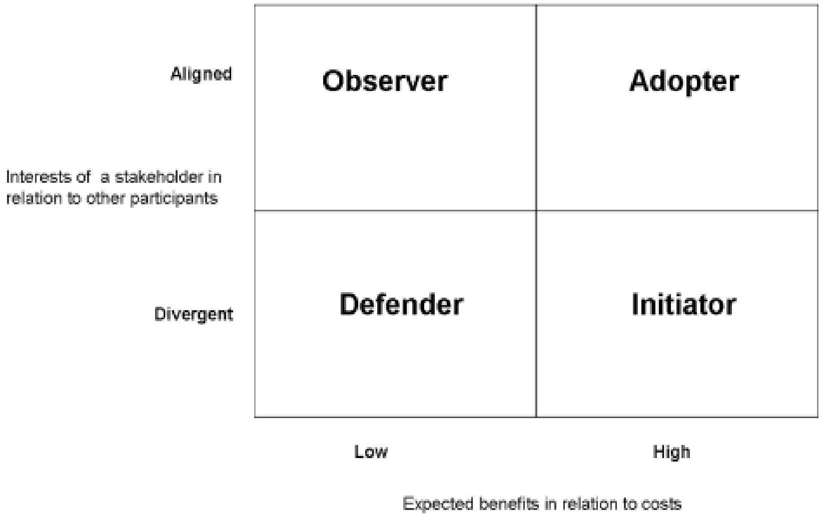

Fig. 1: Types of stakeholders

The actions a company may take to follow these strategies are described below:

The *initiator* is defined by a strong interest to drive the cooperation process in a certain direction aligned with his interests. This implies that different alternatives may exist and especially one he would prefer the most. For the initiator the expected benefits exceed the costs of implementation, even if he is faced with an extra effort of specifying the preferred solution and the challenge of convincing others to follow this specification.

In the context of this article an *adopter* is seen as someone who implements a given technology because it is in general aligned with his interests; therefore he sees no need to actively participate in the process of specification. This strategy is also known as “free riding”. The adopter sees at least a balance between costs and benefits.

Observers do not see a mandatory need to implement the technology mostly because they are uncertain about possible benefits or expected costs. Observers may fear the risk of locking-in into an immature technology. Their future decision to follow a certain specification may be of great interest, because the group of observers may form a critical mass within the partners of a supply chain.

The decision to act as *defender* of a given technology against a new one is seen, if expected costs significantly exceed the expected benefits of a company. This can often be found for smaller companies, which may be forced to adopt a new technology. As former proprietary technologies may be replaced by standardized new ones and therefore more detailed information becomes available to all partners, the defender may expect negative effects to current competitive advantages.

The interdependent technical, organizational and economical aspects make adoption of RFID technologies along the supply chain a complex task. In the following we analyze interests and possible strategies of stakeholders involved in the process of introducing RFID technologies along their entire supply chain. As stakeholders we distinguish the original equipment manufacturer (OEM), 1st-tier suppliers, nst -tier suppliers and other institutions with legitimate interest (e.g. governmental organizations). Examples such as Wal-Mart and Metro show, that there is often a single company at a leading position to act as an initiator for RFID implementation. Mostly this position is held by the OEM. He has a strong interest in information from all partners in the chain. Often his bargaining-power over suppliers gives him the ability to force the implementation process. Other tiers of the supply chain often consist of smaller companies with specialized products. In general, these companies tend to be highly dependent to their contractors.

However, especially in the automotive sector we can observe a shift of bargaining-power towards 1st-tier suppliers, because these suppliers get more and more involved in significant and complete parts of the value-creation process (see Hirthammer / Riha 2005).

Facts such as different production requirements and different percentages on the overall value-creation process on each tier of a value chain cause different expectations about costs and benefits for RFID implementation. The demand for information grows with every tier towards the OEM, but in general suppliers are reluctant to share more information as necessary with other partners.

Regarding the changes of business processes and supplier relationships that come along with RFID implementation we cannot only consider the direct monetary costs for investments (e.g. for hardware and software). There are also costs for the reengineering of business processes, especially because of newly emerging collaborative aspects. Furthermore these changes can negatively affect the current competitive advantages and market shares of a single company, as it may loose bargaining power if information is shared easier over the entire supply chain.

3.1 Solutions for the cooperation problems

In the following we suggest three solutions for the problems caused by the defenders concerns. The aim is to motivate the defender to become an adopter. Accordingly to the role definitions the initiators expect to increase their benefit by establishing a RFID based cooperation along the supply chain. On the other hand the defenders do not identify any advantage of participating in such cooperation. In a monetary view the initiators expect higher returns than costs and the defenders suspect higher costs than returns. However, there is not only a monetary view to take into account but also power, responsibility and companies' image.

The traditional possibility to motivate or better to enforce cooperation is to wield the own power on dependent supply chain members. Dapiran and Hogarth-Scan (2003) define power in the supply chain "as the ability of one entity in the chain to control the decisions of another chain entity." In general the initiators' power is based on dependencies between the initiators and the defenders. The basis of these dependencies can be economic, juristic, technical or based on patents. Independent of the underlying power structures a forced cooperation ends in an unmotivated cooperation and the results are worse than possible. Hirthammer and Riha (2005) state that supply networks are becoming less easy to coordinate by the use of power and pressure towards smaller suppliers. In this regard a new way to inspire cooperation has to be found. The question is: Why should a defender cooperate if it has to invest more than it can finally gain?

One solution is to balance the defenders' economic disadvantages. This can be done by announcing an incentive for cooperation. These incentives were realized by Cost-benefit-sharing (CBS). Hirthammer and Riha (2005) define Cost-benefit-sharing as "a systematic and system-oriented incentive-system that motivates companies in a network to participate in joint projects that do not benefit them directly." The authors also state that first companies accept the necessity for CBS-systems and begin to introduce it. Notwithstanding a lot of work has to be done in order to implement the theoretical CBS approach in practice.

Otherwise, the concerns of the defenders are not only caused by economic disadvantages. There are also strategic reasons to take into account. For instance, the sharing of sensitive information with other companies within the supply chain is not acceptable for some stakeholders. They are unwilling to share information, even if it conduces to a supply chain optimization. Such concerns can be eliminated by new technological advances.

The described cooperation problems and the theoretical solutions are illustrated by a practical example in the next section.

4 Forecasting of manufacturing errors

In this section we illustrate a manufacturing error prediction as a cooperative optimization problem. Accordingly to our definition of cooperative optimization we do not use only local process information but also global information. That means all necessary information along the supply chain is collected and brought together. The advantage of such methodology is obvious. The bigger amount of information is likely to produce a better forecasting model. Because of the capability of contact-free identification of product items RFID technology is able to support both the information collecting and integration. The resulting potentials and problems are figured out in the following.

In this error prediction scenario we work with a small and fictive supply chain shown in figure 2.

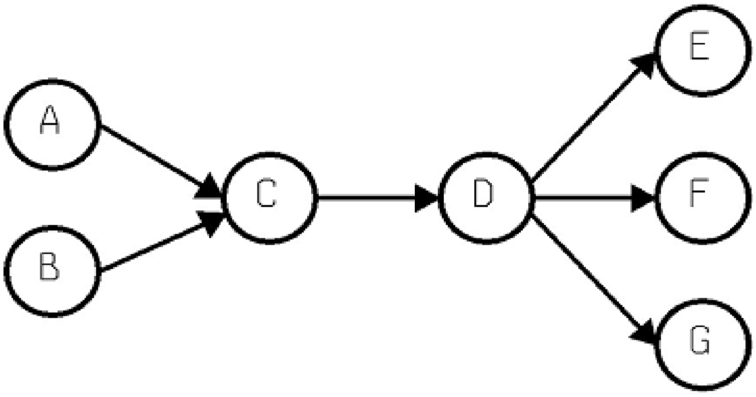

Fig. 2: Schema of a fictive supply chain

The manufacturers (companies A, B and C) produce new items or combine supplied components. The distribution center (company D) has the classic item commissioning and repackaging tasks. The retailers or wholesalers (companies E, F and G) receive the products from the distribution center and sell it to their customers. During this processes the companies collect a lot of data and store it in local databases. The data collecting would become significantly more efficient if all partners along the supply chain were cooperative. This cooperation could include for instance an agreement on using a standardized RFID technology in order to tag each produced item. Thus contact-free item identification will be enabled in the whole supply chain. Furthermore the companies are able to store the time and the location of the identified product in each phase of the manufacturing process. This data can be easily enhanced by putting additional information such as used machine, machine parameters, time and duration of manufacturing or means of transportation. Table 1 shows the local stored data of company A.

ID	Day	Machine	Parameter
A1	Wednesday	1	29,8
A2	Wednesday	2	31,1
A3	Thursday	1	30,5
A4	Tuesday	2	29,9
A5	Monday	1	30,0
A6	Monday	2	30,3
A7	Friday	2	30,7
A8	Thursday	1	30,8
A9	Monday	2	29,8
A10	Wednesday	1	30,7
A11	Tuesday	1	30,7
A12	Friday	2	30,2
A13	Wednesday	2	31,2

Tab. 1: Company A’s local stored data

After finishing the manufacturing process of company A the produced component are supplied to company C. Company C combines the components of its suppliers (companies A and B) and marks the resulting combination with a new RFID tag. In addition company C stores its own process information and the identifier of the built-in components. Thus a complete backtracking of the combined components is enabled. Table 2 shows the stored data of company C.

ID	Supplier’s IDs	Day	Shift	Dur.
C1	A1;B21	Thurs.	day	25
C2	A2;B22	Mon.	day	20
C3	A3;B23	Mon.	night	22
C4	A4;B24	Sun.	late	23
C5	A5;B25	Sat.	day	21
C6	A6;B26	Sat.	night	25
C7	A7;B27	Wed.	day	21
C8	A8;B28	Mon.	night	23
C9	A9;B29	Wed.	late	24
C10	A10;B30	Tues.	night	23
C11	A11;B31	Sun.	day	24
C12	A12;B32	Wed.	night	24
C13	A13;B33	Mon.	night	22

Tab. 2: Company C’s local stored data

The databases of the companies along the supply chain are jointly related by the product identifiers. Thus the producing history of each product can be recovered and the requirements for an intelligent data analysis are fulfilled. In this scenario we firstly apply common data mining techniques in order to analyze the data. The objective of this data mining process is finding rules for the prediction of manufacturing errors. However we need additional information to achieve this task. The retailers have to contribute customers' complaints of sold defective products. That way each product history record is enhanced by a class attribute consisting of the labels "defective" and "accurate". The entire history records are shown in table 3.

Company A			**Company B**			**Company C**			**class**
Day	**Machine**	**Param.**	**Day**	**Transport**	**Dur.**	**Day**	**Shift**	**Dur.**	**label**
Wed.	1	29,8	Wed.	air	1	Thurs.	day	25	A
Wed.	2	31,1	Thurs	train	4	Mon.	day	20	A
Thurs.	1	30,5	Thurs	truck	4	Mon.	night	22	D
Tues.	2	29,9	Wed.	train	4	Sun.	late	23	A
Mon.	1	30,0	Tues.	truck	4	Sat.	day	21	A
Mon.	2	30,3	Tues.	truck	4	Sat.	night	25	A
Fri.	2	30,7	Friday	train	5	Wed.	day	21	A
Thurs.	1	30,8	Thurs	truck	4	Mon.	night	23	D
Mon.	2	29,8	Tues.	air	1	Wed.	late	24	A
Wed.	1	30,7	Thurs	truck	4	Tues.	night	23	D
Tues.	1	30,7	Wed.	truck	4	Sun.	day	24	A
Friday	2	30,2	Friday	train	5	Wed.	night	24	A
Wed.	2	31,2	Thurs	truck	4	Mon.	night	22	A

Tab. 3: entire history records of products

This data collection is the basis for the following data mining process. We use the classical C4.5 (see Quinlan 1993) decision tree algorithm in order to identify the intrinsic rules. The result is the decision tree shown in figure 3.

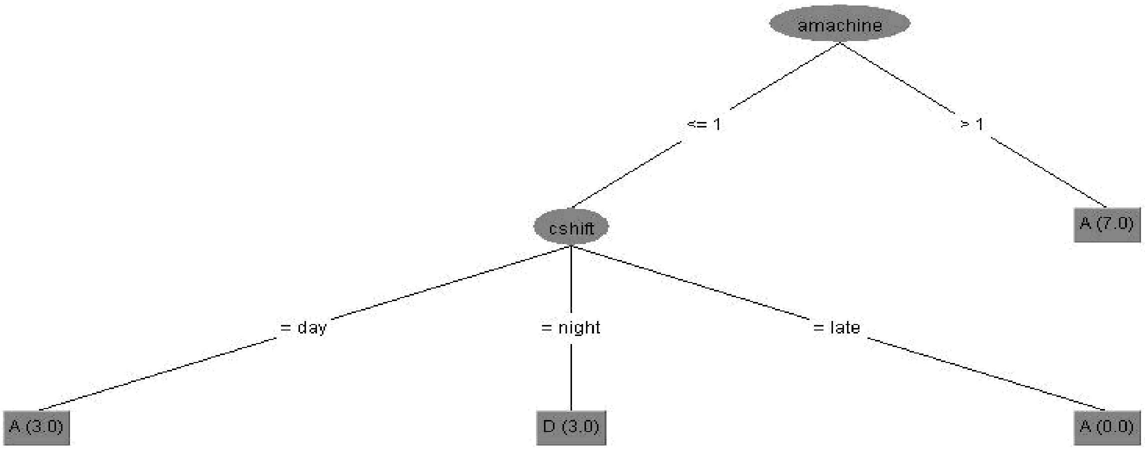

Figure 3: generated C4.5 decision tree

Each path from the tree's root to a leaf stands for a discovered rule: The internal nodes characterize a set of conditions and the class label of the leaf describes the consequence of these conditions. That way the following rules are represented in the tree above:

1. If the goods are not produced on machine 1 of company A then the goods are accurate.
2. If the goods are produced on company A's machine 1 and in the day or late shift of company C then the goods are also accurate.
3. But if the goods are produced on company A's machine 1 and in the night shift of company C then the goods are defective.

On the basis of the extracted rules the advantage of the cooperation is obvious. Because of the analysis of the *integrated data*, it is possible to discover rules with conditions located in different companies. In the example above the rules two and three would never have been found if company A or C did a separate data analysis. Company A is only able to find rule number one.

At least the discovered rules can be used to classify a new record or a new product respectively. The data record of a new product is matched with the extracted rules and if a rule fits the related class label is predicted. Because of the ability to classify a new product the tree and the included set of rules is termed as *classifier*. This classifier can be used to optimize several processes in the supply chain. In the following we describe such an optimization scenario where the presented classifier is used to predict a manufacturing error.

4.1 More efficient quality control

Production errors occur in each manufacturing process where components are produced or get combined. Before the finished goods will be delivered to the distributors they will usually be inspected in a company's internal quality control. There are many different possibilities for the configuration of such a quality control. Due to cost reasons especially at mass-produced items a complete inspection of all articles is impractical. Therefore random inspections are only examined regarding their quality criteria. Without previous knowledge of the products which will be examined the selection of the samples happens usually purely coincidentally (see Bleymüller / Gehlert / Gülicher 1996). Because of the product specific knowledge of the expected quality the selection process can be improved. For example, if it is known that products processed on a Monday on machine 1 show a quality below average then these products can be examined more focused in the quality control process. The following example describes this circumstance more detailed.

If there is no knowledge about the products which will be examined, they are selected randomly and examined afterwards. The goal of this sample is to fulfil the following, exemplary quality requirement:

At least 80 percent of the defective products should be identified with a certainty of 80 percent.

The accomplished sample follows in the mathematical sense a hyper geometric distribution (see Sachs 1978):

$$P(X = k) = \frac{\binom{m}{k} \cdot \binom{N-m}{n-k}}{\binom{N}{n}}$$

In this formula,

- n describes the sample size
- N describes the number of products
- m describes the number of defective products
- k describes the number of defective product which were found in the current sample

Therefore the necessary sample size n can be determined to fulfil the mentioned quality requirements. In the scenario described here, 100 products are to be examined regarding their quality. It is known that an average of 10 of these products is defective. The goal is to identify with minimum effort and a certainty of 80 % at least 8 defective products:

$$0.80 \leq \frac{\binom{10}{8} \cdot \binom{90}{n-8}}{\binom{100}{n}}$$

This inequation is fulfilled for $n \geq 84$. Therefore, on average 84 individual product examinations are necessary to fulfil the described quality requirements. If on the other hand previous knowledge exists over the products which will be examined the effort reduces enormously. Furthermore the above mentioned quality requirements will not be restricted.

The knowledge mentioned here is present in form of a classifier which assigns to each product on the basis of its characteristics either the class ”defective” or ”accurate”. The classifier splits the products into two disjoint subsets:

- The set A of products which are potentially accurate and
- the set D of products which are potentially defective.

However without doubt the classifier does not work perfectly - the consequence are false classifications. Thus, the set A will not exclusively contain accurate and the set D thus not exclusively defective products. Let us suppose that we manage to develop a classifier which produces the subsets shown below:

- The set A covers 60 of originally 100 products. Of these 60 products at most one is defective due to false classifications.
- Therefore the set D covers 40 products of which at least 9 are defective.

Because of our knowledge of the classification quality, it is suffices just to inspect the set D. The necessary sample size is calculated by:

$$0.80 \le \frac{\binom{9}{8} \cdot \binom{31}{n-8}}{\binom{40}{n}}$$

With a sample size $n \ge 37$ the inequation is fulfilled and 8 of the 9 defective products in set D are discovered with a certainty of 80 %. By using the classifier it is possible to reduce the necessary sample size of 84 to 37. The demanded quality requirements to discover 80 % of the defective products with a certainty of 80 % remain untouched. One the other hand it is also possible to keep constant the sample size. This way the certainty gets increased.

4.3 Data Mining related concerns of the stakeholders

Even though the illustrated scenario demonstrates huge optimization potential it shows also a huge deficit. In order to analyze the data of the several companies the data has to be integrated. Caused by different interests and several concerns the involved companies are divided in the roles defined in section 3.

In the presented scenario the OEM fills in the *initiator* role. The OEM is interested in improving the internal quality control without increasing the related costs. Furthermore the initiator aims for gaining the companies’ image by increasing the product quality and decreasing the amount of customers’ complaints. In total the OEM expects a higher benefit than the arising costs. Consequently the initiator is anxious to establish cooperation and to implement a RFID based data mining process along the supply chain.

The retailers affirm on the intention of the initiator because there are no additional costs for them but some advantages. Because of the better product quality the customers’ complaints and the cost involved are getting reduced. In addition there is a new selling potential because of the rising product quality. Consequently the retailers are *adopters* which are willing to support the initiators plans.

The *observer* role is cast by the distribution center. The data mining process do not touch the interests of the observer. The distribution center can identify neither advantages nor disadvantages of the integrated data mining process.

The defender role is filled by the manufacturers. The defenders are unwilling to support the initiators intention because their necessary costs rise above the expected benefits. But not only monetary reasons contradict the realization of the initiator’s plans because the companies are not willing to share their data. Therefore several reasons are possible. First of all, the companies' information is sensitive. It contains a lot of strategic knowledge and the defenders fear that competitors may use it against them. In addition there are also legal restrictions e.g. personalized data of customers are not allowed to be published or shared because of privacy protection laws.

But there is a solution for the mentioned privacy problem. The easiest way is to find a trustful third party. This trustful party gets all necessary private information in order to generate a classifier. The problem is obviously: The private information is not any longer private but rather shared with the third party. Nevertheless this method is used in practice (see Falcon fraud Manager 2007).

Caused by the problem of information disclosure in recent years huge progress has been made in the area of data mining. A new branch of research termed *Privacy Preserving Data Mining* deals with the upcoming challenges of privacy preservation. A lot of new data mining algorithms were developed (see Vaidya / Clifton / Zhu 2006) in order to enable a data analysis without disclosure of sensitive information. So Du and Zhan (2002) suggest a special protocol to analyze private data without the need for a trustful third party. However the presented method is limited to a two party scenario. The result is a decision tree generated without the previous necessary restriction of sharing the private data. But this tree also contains a lot of sensitive information. Therefore Vaidya and Clifton (2005) present a method to analyze private data of more than two parties and without publishing the built tree. The result is a distributed decision tree and each party keeps its own share of it. Using this algorithm it is possible to generate a classifier as described in section 4 but the involved companies do not obtain information on each other. Nevertheless the manufacturing error prediction still works as well as realized by the traditional classifier algorithm.

5 Conclusions

In this article we have shown the problems caused by a shift of the focus from local optimization to cooperative optimization in a RFID-based supply chain. We figured out the cooperation potential of RFID and describe some problems caused by cooperation along supply chains. Therefore we identify four typical roles of supply chain members which have different interests in such cooperation scenario: initiators, adopters, observers and defenders. We have analyzed the interactions between these roles, shown the resulting consequences and outline some appropriate solutions. Finally we have supported our theoretical view with a practical example.

The prediction of manufacturing errors was presented as a cooperative optimization problem and we have shown that it could be solved by using the information and knowledge of other supply chain members. We have figured out the interests of the different stakeholders and presented some solutions for their concerns. The result is an integrated, privacy preserving data mining algorithm for forecasting manufacturing errors.

Bibliography

Adler Case Study (2006), RFID – Anwenderbeispiel Adler Modemärkte / Motex, 23th May 2007, http://www.rfidatlas.de/images/stories/RFID_Fallstudien/adler_september2006.pdf.

Altassan, K. M. / Malmborg, C. J. (1998), Analysis of storage assignment policies in less than unit load warehousing systems, in: International Journal of Production Research, Vol. 36/12.

Axsäter, S. (2000): Inventory Control, Norwell, Massachusetts.

Bleymüller, J. / Gehlert, G / Gülicher, H. (1996), Statistik für Wirtschaftswissenschaftler. München: Vahlen.

Chopra, S. / Meindl, P., (2007): Supply Chain Management – Strategy, Planning, & Operation, Third Edition, Upper Saddle River, New Jersey.

Du, W. / Zhan, Z (2002) Building decision tree classifier on private data, in IEEE ICDM Workshop on Privacy, Security and Data Mining, volume 14, pp. 1-8, Japan.

Falcon fraud Manager (2007) http://www.fairisaac.com/fic/en/product-service/product-index/falcon-fraud-manager/, 23th May 2007.

FKI Logistex (2005), RFID for the real world – challenges and opportunities in the warehouse and distribution center environment, 23th May 2007, http://www.rfidjournal.com/whitepapers/download/33.

Forrester, J.W.: Industrial Dynamics, A major Breakthrough for Decision Makers. In: Harvard Business Review. Jg. 36, Nr. 4, 1958, S. 37-66

Hansen, W.-R., (2007): „Wir sind noch nicht so weit“, RFID im Blick – Das Medium für den Kontaktlosen Datentransfer, 03/2006, http://rfid-im-blick.de/index.php?option=com_content&task=view&id=152&Itemid=84, 24th May 2007.

Hirthammer, K. / Riha, I.V. (2005). Framework for cost-benefit-sharing in logistics networks. Presentation at Frontiers of e-Business Research eBRF 2005 Conference. Tampere, Finland.

IBM case study (2007): The path to a successful RFID-enabled store environment: integrating processes to create value, IBM Corporation, http://t1d.www-03.cacheibm.com/industries/retail/doc/content/bin/rfid_path.pdf, 30th May 2007.

Kumar, P./ Kumar, M., (2007): Vendor Managed Inventory in Retail Industry, TATA Consultancy Services, http://www.tcs.com/NAndI/default1.aspx?Cat_Id=7&DocType=324&docid=430, 25th May 2007.

Lee, L. H./ Padmanabhan, V./ Whang, S., (1997): Information Distortion in a Supply Chain: The Bullwhip Effect, Management Science, Vol. 43, No. 4, Frontier Research in Manufacturing and Logistics.

Lee, L. H. / So, K. C.; Tang, Ch. S., (2000): The Value of Information Sharing in a Two-Level Supply Chain, Management Science, Vol. 46, No. 5.

Löwer, U (2006) Strategies for the Development of Interorganisational Standards. Physica-Verlag HD.

Malmborg, C. J. (1996), Storage assignment policy tradeoffs, in: International Journal of Production Research, Vol. 34/2.

Martin, H. (2000), Transport- und Lagerlogistik – Planung, Aufbau und Steuerung von Transport- und Lagersystemen, 3rd ed., Vieweg: Braunschweig.

Miles, R. E. / Snow, C. C. (1978). Organizational strategy, structure and process. New York: McGraw-Hill.

Ostler, U. (2004), RFID-Einführung: Auf die Middleware kommt es an, Silicon 23.9.2004. http://www.silicon.de/enid/business_software/10498 (letzter Zugriff: 23.01.2007).

Pfohl, H.-C. (2004), Logistiksysteme – Betriebswirtschaftliche Grundlagen, 7th ed., Springer: Berlin.

Quinlan, J. R. (1993), C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.

Quelle Case Study (2006), RFID – Anwenderbeispiel Quelle GmbH, 23.05.07, http://www.rfidatlas.de/images/stories/RFID_Fallstudien/quelle_januar2007.pdf.

RFID Atlas (2007), Metro AG. Projektübersicht. 2006. http://www.rfidatlas.de/index.php?option=com_content&task=view&id=43&Itemid=57 (letzter Zugriff: 23.01.2007).

Sachs, L. (1978), Angewandte Statistik. 5th ed. Springer Verlag.

Stadtler, H. / Kilger, C. (2002), Supply Chain Management and Advanced Planning – Concepts, Models, Software and Case Studies, Second Edition, Berlin, Heidelberg, New York.

Stadtler, H. (1998), Gestaltung von Lagersystemen, in: Isermann, H. (Ed.) (1998), Logistik – Gestaltung von Logistiksystemen, 2nd ed., Verlag Moderne Industrie: Landsberg/Lech.

Thiesse, Frédéric / Fleisch, Elgar (2007), Zum Einsatz von RFID in der Filiallogistik eines Einzelhändlers: Ergebnisse einer Simulationsstudie. In, 2007.- 8. Int. Tagung Wirtschaftsinformatik.- Karlsruhe.

Vaidya, J. / Clifton, C. (2005) Privacy-Preserving Decision Trees over Vertically Partitioned Data, in Proceedings of the 2005 IFIP WG 11.3 International Conference on Data and Applications Security.

Vaidya, J. / Clifton, C. / Zhu M. (2006) Privacy-Preserving Data Mining, Springer, New York.