Data Mining in the Closed-Loop CRM-Approach for Improving Sustainable Intermodal Mobility

Thees Gieselmann¹, Marcel Severith², Benjamin Wagner vom Berg³, Jorge Marx Gómez⁴

¹University of Oldenburg, Department of Computing Science, Ulihornsweg 84, D-26129 Oldenburg thees.gieselmann@uni-oldenburg.de
²University of Oldenburg, Department of Computing Science, Ulihornsweg 84, D-26129 Oldenburg marcel.severith@uni-oldenburg.de
³University of Oldenburg, Department of Computing Science, Ulihornsweg 84, D-26129 Oldenburg benjamin.wagnervomberg@uni-oldenburg.de
⁴University of Oldenburg, Department of Computing Science, Ulihornsweg 84, D-26129 Oldenburg jorge.marx.gomez@uni-oldenburg.de

ABSTRACT
In this paper we introduce sustainability oriented Customer Relationship Management by taking Jinengo as an example. Jinengo is a mobile application assisting users with planning routes by offering mobility options that incorporate various means of transportation, thus implementing inter modal mobility. The offered alternatives are based on the preferences set by the user but rated by the route’s sustainability performance, thus pushing the user subtly to a more sustainable option. At present, the actual route choices of individuals in the past are not yet considered for compiling the options suggested by the system. Therefore we will discuss how to use data mining to provide even better suggestions, thus implementing the so-called closed loop of Customer Relationship Management.

Keywords
Business Intelligence, Data Mining, Intermodal Mobility, Sustainability, Customer Relationship Management, Closed Loop.

1. INTRODUCTION
The role of information and communication technologies in the context of sustainable development is discussed frequently. So far, most studies in this field are dedicated to issues like using resources or energy more efficiently during production processes [4] & [5], leading to the development of corporate environmental management information systems [12]. However, these efficiency gains may be lost by using the corresponding good or service to a greater extent, leading to rebound effects [4], [5] & [11]. Thus, for enabling sustainability by ICT, deep structural change is necessary, including radical changes in consumption patterns. In particular this includes the transition from property-based to service-based consumption habits [4]. Transportation contributes about 13% to all anthropogenic greenhouse gases worldwide and is one of the fastest growing sectors [8]. Thus, transportation is one out of many obstacles for ICT4S 2013: Proceedings of the First International Conference on Information and Communication Technologies for Sustainability, ETH Zurich, February 14-16, 2013. Edited by Lorenz M. Hilty, Bernard Aebischer, Göran Andersson and Wolfgang Lohmann. DOI: http://dx.doi.org/10.3929/ethz-a-007337628

us to overcome on the path to sustainability. Besides technical innovations like gas-efficient or alternative fuel cars, social breakthroughs are necessary as well. Solutions like public transport and car sharing already incorporate the idea of a service-based economy. Even though such solutions hold huge potential in regard to reducing environmental burdens, it is hard to change people’s behavior. Many people regard services like public transport as insufficient to their needs, for rational as well as emotional arguments. Besides necessary improvements in the availability and quality of offered mobility services, one has to change the normative decision making of each individual as well.

Furthermore there is also the chance to satisfy the need lying behind a specific travel decision in an alternative way. For example, recreational holidays could be spent on a closer destination and business meetings could be substituted by video conferencing. According to [10] such a cultural shift in satisfying needs enables the greatest sustainability prospects, because it offers a greater degree of freedom regarding which technology or system to choose. While a distant destination may only be reached by plane, a more proximate one could be reached by the technology “car” with the system even being “rent”.

However, changing mobility habits is a tough task. Choices of transportation are not solely motivated by reaching a specific destination, but express intrinsic social and cultural needs as well. The car used, for example, exposes social status, hence fulfills the need of social acceptance [6]. Customer Relationship Management (CRM) deals with the customer before, during and after the actual sale, influencing his behavior [9]. Therefore it is the system of choice when trying to offer more sustainable mobility options to the individual. Given this background, we apply data mining methods to the CRM system to improve understanding and knowledge of mobility behavior patterns. This approach is used alongside the route planning software Jinengo, which assists mobility choices of individuals. By analyzing day-to-day choices, the mobility preferences of a user can be discovered, segmenting them into homogeneous groups with similar habits. We believe that understanding these preferences is the key for unlocking peoples’ mobility habits and marketing sustainable alternatives.
2. CUSTOMER RELATIONSHIP MANAGEMENT

CRM describes the theory of pursuing the company’s goals by identifying and satisfying the customer before, during and after a sale [3]. Gathering data during CRM builds a knowledge base about the customer which considers not only internal but also external data sources.

CRM consists mainly of two phases: operative and analytical CRM. Operative CRM strengthens customer contact through marketing, sales and service. The information gathered during the operative phase is then processed in the analytical CRM, where raw data is analyzed in order to extract knowledge yet invisible. The information generated during analytical CRM can be used by the operative CRM to update its customer contact processes [3].

The term “Closed Loop” describes the flow of information between the operative and analytical CRM, which represents a never-ending loop of information being gathered, analyzed and processed. While customer-related data is collected in the operative CRM, it is transported to the analytical CRM. In the analytical CRM the data is stored and analyzed to extract meaning from it. This newly gained knowledge can be deployed by the operative CRM. Using this knowledge, more accurate marketing campaigns can be developed and targeted on specific customer groups [3].

3. TRANSFERRING CRM TO A SUSTAINABLE CRM

The traditional CRM approach does not meet the requirements to change individuals’ unsustainable behaviors. Therefore a Sustainable Customer Relationship Management (SusCRM) advances the original CRM by two dimensions: ecological and social. It thereby implements the triple bottom line of sustainable development [1]. Doing so, it adds a further goal to the CRM besides binding the customer. The SusCRM aims at achieving a learning relationship in order to enable the customer to change his consumption behavior to a more ecologically- and socially-conscious way. To meet these requirements, the closed-loop process has to be modified to not only include customer information but aspects like energy consumption and or CO₂ emissions as well [14].

CRM systems are able to influence customer behavior through targeted campaigns. For example, advertising an energy-efficient device could be supported with a campaign demonstrating to the customer how to use the device in an appropriate manner thus avoiding rebound effects. To apply a learning strategy, the SusCRM must be able to target the customer appropriately. This requires the appropriate data and the methods to analyze the broadened data scope. Increasing the SusCRM’s data source by adding ecological and social data fulfills the first requirements, while adding Data-Mining methods searching for links in the data meets the second requirement.

The knowledge discovered can e.g. be employed to organize customers in groups of likely behavior and identifying potential improvements within these groups. Revealing and targeting those potentials is one benefit of the SusCRM.

A SusCRM may be applied in businesses with a broad product spectrum, trying to offer sustainable products and services to those customers who could be interested. It may also help to accompany customers during the whole life cycle of products, thus strengthening customer loyalty. Customers may be supported with using the product responsibly, taking care and repairing it as well as recycling or disposing it properly.

A SusCRM may alternatively be applied by organizations not selling their own products and services, but mediating between customers and businesses instead. Such an organization could help customers find a sustainable way of satisfying their needs by appropriate products and services, while being independent from profit interests of individual businesses. Such application can be imagined for various aspects of human life, e.g. for food, housing or mobility. In the following we will focus on the latter.

4. JINENGO AS AN SUSCRM EXAMPLE

Jinengo is a practical application of a SusCRM in the field of intermodal mobility.

Intermodal mobility is the idea of combining different means of transportation in order to reach a given destination. With the shortcomings of public transport in regard to flexibility and the shortcomings of private transport in regard to economic, social and ecological high costs, intermodal mobility seems to be a viable option. A study carried out for the German railway company DB AG [7] found an increasing interest in intermodal mobility, especially in the group of early adopting city dwellers. However, most mobility options exist independent from each other and offer no real linkage.

Jinengo is a mobile application developed by a research group from the University of Oldenburg that assists users with planning such intermodal travel routes. After a user enters the starting point, destination and time the system offers alternative travel options, taking into account several means of transportation. The results are calculated and sorted according to priorly-entered user preferences, including comfort, flexibility, cost and sustainability. Because of this, results may vary from search to search. In order to encourage the user to change his habits, the routes are highlighted in shades of green according to their sustainability performance. The resulting choice of transportation remains with the user, but is influenced by the visualization of the different travel alternatives. Furthermore messages can be displayed in the user interface e.g. for giving hints or special offers for more sustainable transportation alternatives [13] & [14].

With its intermodal approach, Jinengo improves the accessibility to the plurality of mobility service providers for the user. Jinengo can take any kind of mobility service provider into consideration. Until now, considered mobility service providers include car sharing, rail and other forms of public transportation [13] & [14]. Additionally, it has the potential to change behavior by offering more sustainable solutions that may be in reach but haven’t been considered yet.

The system architecture of Jinengo consists of three layers, which are depicted in Figure 1.

---

1 The triple bottom line is quite popular but also a bit contradictory it itself. While ecology limits the extent of all human (societal) activities, economy is just one subsystem of human society. Because of this extraordinary importance, we will focus on the ecological dimension of a SusCRM in the following.
Figure 1. Jinengo System

All functional system processes, data collection and data storage are done in the back end layer. An operational CRM database stores all gathered data, in particular the personal user data and prior chosen routes of individuals using the software [13] & [14]. When a user requests a new route, the system composes a list of travelling alternatives, considering the personal preferences, from the operational CRM.

The front end layer contains a GUI, which serves as point of interaction with the user. The GUI provides a web-based simple graphical interface to plan routes and display reports with user’s statistics. The analytical layer consists of a data warehouse (DWH) and a reporting engine. The data stored in the operational CRM database is processed into the analytical DWH. The knowledge generated there is meant to fulfill two purposes. It is intended to feed reports and dashboards of the operational system as well as refine the user’s search results to better meet individual mobility requirements. However, this analytical part of Jinengo is not yet implemented.

At present, traveling alternatives are only derived from the priorly-entered preferences of a user. The actual route choices of individuals using the system in the past are not yet considered. However, storing past choices in the database enables Jinengo to segment the user accordingly and adjust future results correspondingly.

4.1 Motivating usage of Jinengo

In order to offer the user better fitting mobility options we apply Business Intelligence (BI) methodologies to Jinengo in the following. These will be introduced more detailed in chapter five. To do so, we need a great extent of data where we base our analysis on.

Interacting on a regular basis with the system, the user represents the most important source of data. The more users utilize Jinengo frequently the better data mining results become. Consequently the obstacles of signing up for the service at first and consecutively using it have to be overcome.

The main reason for using Jinengo is the unique intermodal approach pursued. With several means of transportation covered for a given route the user can choose the most fitting one for his current need, even if it is not a very sustainable one.

Additionally, incentives could support regular usage. Reports and interactive dashboards motivate sustainable conscious users optimizing their own sustainability performance and comparing it with others. This way every user can observe its own development, which provides satisfaction if positive or pressure if negative. As a prospect, sharing one’s own sustainability performances in social networks like Facebook could strengthen this incentive even further.

4.2 Interest groups

Besides the user there are other stakeholders which have a great interest in analyzing data of Jinengo.

The Jinengo executives are all people working at Jinengo, which are responsible for the strategic development of Jinengo. They are interested in analyzing the success of the platform and its ability to change user behavior. Therefore they need viable data about platform usage and the development of the user’s sustainability performances. This data serves as the basis for decision making in order to improve the utility of the platform. The executives for example may conclude which campaigns are to be launched.

Mobility service providers are also interested in the results of data mining. Analyzes can be used for finding the right audience for their marketing campaigns. Furthermore, mobility service providers may use it to refine their offerings. A car-sharing provider for example could find out in which areas potential car-sharers live and thus expand the business in this very area. Offering all this information to mobility service providers is a potential source of funding for Jinengo. Of course this must not harm the reputation of Jinengo being an independent route planning software with high data privacy standards.

Similar to mobility services providers, scientists with research in the field of mobility may have an interest in statistical data too. Obviously, especially when dealing with “external” stakeholders like mobility service providers, data privacy requirements play a crucial role. With the need of fulfilling both these “internal” and “external” information interests, different forms of data representation have to be chosen.

5. APPLYING BI METHODS TO JINENGO

Applying Business Intelligence methodology enables closing the CRM loop of Jinengo. The new knowledge gained in the DWH is employed in two ways to benefit Jinengo.

Reports and dashboards contain condensed information for the different stakeholders involved, like users, Jinengo executives, mobility service providers and scientists. The prepared reports and dashboards are aggregated in a way that on the one hand the demanded information are available but on the other hand personal data privacy requirements are ensured. These reports also can build the base for an incentive scheme for sustainable mobility.

Additionally, the gained knowledge is transferred back into the operational database to refine the user’s search results to better meet individual mobility requirements. This may help increasing Jinengo’s capability in offering qualitative results that are able to attract the user to more sustainable mobility choices. The analytical questions arising can be assigned to the previously defined stakeholders.

Users of Jinengo may ask: How sustainable is my own mobility behavior? How is my sustainability performance compared to other groups of people? Are there viable alternatives to my current form of mobility?

Jinengo executives may ask: How can the success of the Jinengo platform be assessed? Is the platform used regularly? How can users be motivated to use the platform regularly? How is the sustainability performance of users developing? How can we
motivate end users to a sustainable behavior? What incentives may the platform give for behaving more sustainable?

Scientists may ask: What motivates people to behave in a sustainable way? What are the obstacles? Which incentives do different audiences need for changing their habits?

Mobility service providers may ask: Which people are receptive to what means of transportation? What are the reasons for this interest? Which advantages of an individual means of transport have to be emphasized in special? How to best market a specific offer? Is there a demand for a product that is not yet covered by an according supply?

All these questions can be answered with the help of appropriate reports and dashboards, showing information in varying levels of detail. At this point, these reports and dashboards should not be addressed any further. We will focus on data mining instead.

Looking at data mining of past routes planned by Jinengo, the analytical questions can be condensed to the following:

1. What was the motivation of a user to go to his given destination?
2. Why did the user choose a specific route or set of means of transportation?

Dealing with these questions may support Jinengo to deal with new routing requests of users in the future. Answering the first question may help finding alternative destinations satisfying the same needs underneath. Answering the second question may help finding sustainable but still appropriate routing alternatives to suggest to the user.

These data mining tasks need an extensive data model to be based on. Thus, the main attributes and indicators of the data model are introduced in the following.

6. DATA MODEL

We designed the data model in Figure 2 especially for purpose of data mining. Therefore it is aggregated from the operational CRM database model of Jinengo.

The personal user data comprises all the master data of the users. Routes were once driven by the user and are differentiated into one to many sub-routes, each with one mean of transportation.

6.1 Personal user data

Personal user data are differentiated into preferences, personal attributes and available mobility options.

In Jinengo users can set their preferences for sustainability, flexibility, comfort and price intentionally. Users are able to change them over time and adapt them to their current attitude, so preferences might change quite frequently. As they do not interfere with protective private information they are easier to handle regarding privacy issues.

Personal attributes on the other hand connect directly to the user’s identity. In Jinengo age, gender, residence, income and family status are considered. Unlike preferences, personal attributes have to be protected heavily because of privacy reasons. Even though it is likely that users are cautious with entering personal attributes, this data is an important asset for data mining, thus need to be collected as broad as possible.

Available mobility options specify which means of transportation can be used easily by an individual. This includes the ownership of a car, an e-bike\(^2\) as well as the membership-based access to car sharing, public transport and intercity rail. As these mobility options relate to individuals like personal attributes do, they have to be protected regarding privacy issues as well.

6.2 Route attributes

Routes are distances travelled by a user during a specific period of time in the past. A sub-route is one part of the whole route, resembling one means of transportation used. Thus routes can consist of several sub-routes. The data model describes every route with its associated sub-routes individually. Sub-routes are connected with a n:1 relationship to routes.

Every sub-route includes the attributes described in the following.

**Point and time of departure and destination:** Specifies place and date of start and end of the sub-route.

**Ecological impact:** This attribute represents the accumulated ecological impact from the sub-route, consisting of resource and energy consumption. There are several options for assessing the ecological impact. We won’t go into more detail here, but propose a simple thus easily understandable method like the carbon footprint or ecological footprint.

\(^2\) The ownership of an ordinary bike can be postulated in Germany, although this does not imply regular use. The use of an e-bike however improves comfort and widens the range of cycling for a broad audience.
**Time:** The time needed to finish the sub-route.

**Effective time:** How much of the travel-time can be used for other activities like work.

**Distance:** The distance of this sub-route expressed in km.

**Costs:** The amount of money the user requesting the sub-route has to pay for it.

**Means of transportation:** Reflects the type of the transportation, e.g. train or car. This attribute reflects the vehicle specific attributes like seats or entry restrictions as well.

**Context information:** Until now, these attributes are not incorporated into the system due to their complexity. Considering traffic jams or weather information is very important and an interesting topic, especially regarding data mining. But there is a lot of research to be invested into this topic before it can be fully incorporated.

Routes have a different set of attributes defining their character, described in the following.

**Point and time of departure and destination:** Specifies place and date of start and end of the whole route.

**Luggage:** The amount of objects the user wants to transport on the route.

**Need:** The embedded personal need behind the travel. A need could be shopping, work or vacation.

**Passenger:** The number of people travelling together with the user.

**Ecological advantage & ecological disadvantage:** The ecological impact of all appendant sub-routes can be aggregated on the level of the whole route. Subtracting this overall impact from the impact of the worst and best alternative route provided by Jinengo leads to both an advantage and a disadvantage. This helps assessing the ecological impact of a route in comparison to other options for the same destination.

The calculation of an advantage and a disadvantage can be done for the sub-route attributes time, effective time and costs as well.

All attributes represent the planned values from Jinengo. The actual values occurring during the travel are not tracked afterwards, e.g. the time of travel may be affected by unexpected delays. As a result the actual travel values remain hidden. Tracking these differences could give insights for an extensive data mining but is not yet incorporated into Jinengo.

### 7. DATA MINING

Data mining represents a core feature within a SusCRM. In order to support the user’s sustainability it is important to understand his behavior, as described above. To complete the closed loop, the data gathered needs to be analyzed and further knowledge has to be extracted from it. The methods provided by data mining enable the SusCRM to solve this task.

One viable procedure to implement the analytical part is exemplary described in the Jinengo-Platform. In Jinengo the data stored in the analytical data warehouse is analyzed mainly with classification, cluster analysis and association methods. The route-attributes and the user-attributes represent the data pool on which data mining will be used.

After discovering new knowledge about the user, this information is being transferred into the operational database and can be used for further campaigns, thus closing the closed-loop. In the following the methods implemented in Jinengo are described.

#### 7.1 Association

Jinengo’s goal is to increase sustainable mobility. Association techniques help achieving this goal by analyzing day to day user choices. The method searches for rules in the user’s behavior in order to arrange the service in a way that accommodates the user’s behavior [2].

Although association is a descriptive data mining technique and normally is not used for forecasting, it is nevertheless of great importance. A typical application of association is the cross-selling use case, which can also be used in Jinengo. If a lot of people who used mobility service A also used mobility service B the chance that a customer will use mobility service B as well, if he already chose service A, is very high. The ambition behind association analysis is to uncover these rules in the user’s behavior.

#### 7.2 Cluster analysis

Another descriptive data mining technique is the cluster analysis, grouping different objects according to their similarity. In the Jinengo project, users are compared according to a specific set of attributes. If the accordance between two users is high enough in comparison to the accordance to other users, they are grouped into the same cluster. While comparing several attributes over a large number of users, groups of users can be identified that are very similar to each other [2]. For example, one cluster could contain all users with an age of 25-35, a mediocre income, and a medium sized car who are living in a suburb.

The personal attributes and the preferences chosen are used to cluster the users into groups of high internal similarity and maximum distance to other groups. We assume that users within the same cluster will react in a similar way to proposals like choosing a different means of transportation.

An exemplary implementation in the mobility sector was done by [7]. Young city dwellers possibly interested in intermodal mobility were clustered into three groups:

- **Pragmatic:** People out of this group believe they have a good access to a variety of mobility options. They think that their usual targets are easy to reach without a car. Privacy during travelling is of no importance, ecological aspects are disregarded.
- **Car-affine:** People out of this group can hardly imagine being mobile without their private car. Less car usage implies a restriction of flexibility and freedom.
- **Eco-oriented:** People out of this group use bikes more frequently than the pragmatics but make less use of cabs or rented cars. Usually a car is available, but rated negatively because of ecological reasons.

The groups created serve as peer groups for benchmarking the user’s personal sustainability performance. The ecological impact of all routes driven by the individual is therefor set in relation to the overall ecological impact of the corresponding peer group. This comparison is showcased to the user. Comparing the user with a specific peer-group of high similarity instead of all users has two advantages. First it benefits those, who are not very sustainable yet, but try to improve. Comparing them to a person with a completely different personal background, e.g. a person not owning a car could demotivate them. Second, users with a high sustainability-score are still put into a competitive position. As a result the urge to a more sustainable behavior for those with low
sustainability is satisfied as well as the need for improvement even for a very sustainable user.

New users and changes of existing ones influence the clusters. After a certain amount of users entered the system for the first time or already existing users changed their preferences the cluster analysis needs to be updated, ensuring its quality.

7.3 Classification
In contrast to association and cluster analysis, classification is a data mining technique used for forecasting. It is used to classify a particular item according to the value of a specified attribute [2].

While clustering groups existing users, classification may help determining the affiliation of new or changed users to already existing groups on the basis of a set of attributes. An algorithm classifies each user into that group of users which has the largest similarities within the attributes consulted [2].

In contrast to the cluster analysis the classification is done automatically. Whenever a new user signs up for Jinengo the classification algorithm groups the user to a corresponding cluster depending on his initial information provided. The algorithm surveys personal information about the user like age, gender, family status and place of residence as well as all motor vehicles or local-traffic cards owned. In regard of these attributes the clusters available are searched for the one with the highest accordance and the new user is being grouped into that group.

Another application of classification in the context of Jinengo is the identification of favored sustainable behavior patterns. For example, the classification process could identify the “rules” leading to a personal ownership of an e-bike. Users with similar personal attributes and preferences may then be proposed to buy an e-bike as well, helping to reduce their kilometers driven with vehicles propelled by a combustion engine.

8. CONCLUSION AND OUTLOOK
In our work we describe the prospects of a Sustainable Customer Relationship Management and introduce Jinengo by example.

With Jinengo not offering own products and services but mediating between customer needs and intermodal mobility services instead, Jinengo is an independent and credible system assisting customers to find sustainable mobility options. We depicted a data model and possible ways of mining past mobility choices and thereby finding better suited mobility alternatives for individual users, thus completing the closed loop of a CRM system.

However, there are still many open challenges. Up to now, our data model only includes the planned data by Jinengo. The actual data of a route may vary, for example the time may change because of a delay. But incorporating this data is not that easy. With Jinengo being a planning platform the user would have to be motivated using the platform after the route has taken place. Because of these operational difficulties, tracking this data is not yet incorporated into Jinengo although it might give insights for an extensive data mining.

The biggest potential but also biggest challenge for Jinengo is the proposal of alternative ways to satisfy the same need as a user-given destination would have. Like travel portals offer holiday locations the user was not initially looking for, Jinengo could suggest alternatives likewise. Precondition for this is the automated recognition of needs behind a given destination or the possibility to enter them directly.

9. REFERENCES