Data and information utilisation in waste management systems

T. Nuortio, G. van den Dool, T. Hiltunen, V. Matikka, T. Räsänen & M. Kolehmainen
Environmental Informatics, Department of Environmental Sciences, University of Kuopio, Finland

Abstract

Within a waste management company new data is accumulating in the different data management systems, but comprehensive utilisation of available data is too often neglected. New software products and systems, like driving guidance systems, are collecting additional information daily about the service areas. Extensive use of the gathered data as well as the data available in public and private domains can provide essential information for the planning and decision making of regional waste collection.

In this study we explored the possibilities of refining, modelling and utilising the data stored in the operational databases of the Finnish waste management company Jätekukko Ltd. The focus of this study was especially on modelling and predicting waste formation with MLP (multi-layer perceptron) and giving an insight into future possibilities of utilising the modelled data.

Keywords: solid waste management, information utilisation, data modelling, neural networks, multi-layer perceptron.

1 Introduction

The changes in the Finnish waste management sector have increased the need for more comprehensive use of collected data and information systems. Information technology (IT) enables automatic data collection, data transfer and knowledge refinement, IT also enables implementing new services and technologies in waste management and provides the basic elements for network design to construct new sophisticated data management systems.
Information systems and software products are commonly used in the modern waste management in order to support the daily action of different parties. Within a waste management company new data is accumulating in different data management systems; still comprehensive utilisation of available data is too often neglected. New software products and systems, like driving guidance systems, are collecting additional information daily about the service areas. Also public and private domains are providing more accurate background information which could be utilised in the waste management information system. Extensive use of the gathered data as well as the data available in public and private domains can provide essential information for the planning and decision making of regional waste management (for example planning of waste collection).

2 Case and data description

Jätekukko Ltd. is a Finnish waste management company which is responsible for organising waste management in 18 municipalities (City of Kuopio and the surrounding areas) serving approximately 180 000 people in eastern Finland. Jätekukko is responsible for organising of municipal solid and hazardous waste collection, waste disposal site activities, waste consultation and other customer services. The actual waste collection and transport is outsourced to haulers which are selected by the tenders once in five years.

In June 2003 Jätekukko embedded a new Transport Control System (TCS) provided by Ecomond Ltd. in their entire service area. Each truck is equipped with special designed hardware, software and GPS-devices in order to guide the driver through the workload (routes). The offices of Jätekukko and haulers have the software tools for controlling and monitoring the collection process. TCS enables almost real-time data transfer and control of collection because after each pick-up the driver is updating the time and possible exceptions in the system as well as additional information about the location.

Other data management systems in use are a business control system (JHL), designed and maintained by Tietomitta Ltd, which is mainly used for billing purposes and a weighting system (Skalex Eko), provided by Pivotex Ltd, in waste disposal sites; storing data about all the movements in the waste disposal site (Figure 1).

In this study we explored the possibilities of refining, modelling and utilising the data stored in the operational databases of Jätekukko Ltd. The link between the databases (TCS and Scalex) was established in order to track the routes for each weighting action in the disposal site. The statistical information (including all socioeconomic information about the study area in 250 m × 250 m geographic grids) maintained and provided by the Statistics Finland was also used in this study.

The data structures available could be used in several domains e.g. to cluster analyse the waste disposal site data using data mining techniques but this study is limited to analysing only the data about the collection process. Section 4 focuses on modelling and predicting waste formation although the same data could also
be used to model and simulate other stochastic components e.g. the travelling
times of waste collection. Section 5 is giving an insight to future possibilities of
data and information utilisation in the waste management sector.

Figure 1: Data sources and data management systems.

3 Pre-processing the data

The data obtained from the TCS databases provided the necessary information of
clients including container volumes, contract agreements, collection locations,
and the total amount of emptied containers between two disposals. The statistical
database could be used to obtain information about the population characteristics
in an area in order to calculate the distribution (of waste indicators) per container
or location. Also other statistical parameters of consumers could be achieved
from age distribution, income levels and social patrons inside independent areas
or grid cells.

In the TCS the drivers are entering constantly new data about collection
events which is stored as historical data (upon this moment approximately
630 000 data lines, collected in half a year) to the server. Like in every human
action errors are inevitable and therefore the validation of the entered data is
important. In the waste disposal site the data about vehicles is automatically
saved but the data about waste types could still be typed in incorrectly.

The data pre-processing was carried out and the incorrect data was deleted
from the dataset. Also the correlation between the weight of the load and the
total volume of containers was checked. Data mining techniques were used to
explore the data structures and to find unexpected relationships from the data in
order to find the best data for the modelling and validation of the models.

4 Calculating and predicting regional waste formation

The rate of waste formation varies over time and place depending on waste type,
the amount of consumers, socio-economic factors, time of year, attitudes etc. [1,
2, 3, 4]. Exact waste formation (weight/volume) in the containers can only be known by measuring it e.g. with the collection vehicle with weighting system. To implement a weighting system is a large investment and not many waste management companies have made this effort; therefore the exact weight of containers is not often available. By using existing data the regional waste formation can quite reliably be estimated as well as the evaluation of the needs (and profitability) for measurements.

4.1 Multi-layer perception model

The multi-layer perceptron (MLP) is probably the most widely known and successful neural network [5]. It consists of system of interconnected neurons which is a model representing a nonlinear mapping between an input vector and an output vector [6]. MLP provides a flexible and non-linear tool for tackling regression problems and because non-linear relationships exist in the real world it is not sufficient to model these problems using traditional regression [7].

![Diagram](https://example.com/diagram.png)

Figure 2: A. The input data for the model, B. Model for predicting regional waste formation.

In the first phase of this study a “MLP” model was constructed in order to determine the waste formation in statistical grids (250 m × 250 m). The collection data was known to include container volumes and locations, vehicle waste amount in the time of disposal and the emptied containers between two disposals. From the statistical database the information about population per container was calculated. Also the consumer habits were extracted from age distribution and income data. (Figure 2)

The known statistical information and container capacity were used as input variables in training. The output of the network was selected to be waste formation (kg) for 30 days. A 20 % random sample of the complete data was used in validation. In the second phase the waste formation in particular time
(the time of collection) was predicted. In the future, if necessary, the model can be improved by individual weight measurements.

4.2 Waste formation in containers

The data in waste disposal sites can be combined with the collection data (TCS) using vehicle IDs. The Scalex Eko weighting system saves data of every weighting action in the landfill. Consequently, the mass of the load and for that the route or the pickup places can be tracked. To process this information in further the total waste is divided to private and business clients (containers) in proportion to container capacity. The waste of business clients is reallocated to containers relative to container volume. The waste of private clients is reallocated to the grids by usage-efficiency of the container capacity whereas inside the grid reallocation is proportional to the container capacity.

5 Utilising the modelled data

The refined data, in combination with raw data, can be used to support the planning of waste management. General techniques used in the Geographical Information System (GIS) can be used to visualise and analyse the different modelling results and provide new information based on the combinations of modelled data [8, 9, 10]. Multiple views are designed to give a broad perspective over the problem at hand. Waste management companies and haulers can use this new information in order the preventively solve the future challenges and changes in strategic, tactical and operational level.

Example 1:

The information about the waste formations can be utilised in numerous ways and for different purposes. For statistical use the information about regional waste formation (e.g. the amount and the variation over time) can be calculated and predicted (Figure 3). This information can then be used to analyse pricing agreements or distribute the shared waste taxes fairly.

Table 1: Variation of collected household waste (ton).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>297</td>
<td>163.31</td>
<td>216.98</td>
<td>209.36</td>
<td>146.92</td>
<td>736.57</td>
</tr>
<tr>
<td>749</td>
<td>42.02</td>
<td>44.35</td>
<td>42.24</td>
<td>29.11</td>
<td>157.72</td>
</tr>
<tr>
<td>593</td>
<td>44.71</td>
<td>3.55</td>
<td>5.00</td>
<td>3.45</td>
<td>56.70</td>
</tr>
<tr>
<td>778</td>
<td>14.43</td>
<td>15.87</td>
<td>14.83</td>
<td>8.96</td>
<td>54.09</td>
</tr>
<tr>
<td>534</td>
<td>8.61</td>
<td>8.26</td>
<td>7.06</td>
<td>5.09</td>
<td>29.03</td>
</tr>
<tr>
<td>174</td>
<td>8.63</td>
<td>7.96</td>
<td>6.06</td>
<td>3.47</td>
<td>26.12</td>
</tr>
<tr>
<td>227</td>
<td>7.75</td>
<td>6.12</td>
<td>5.54</td>
<td>3.59</td>
<td>23.00</td>
</tr>
<tr>
<td>476</td>
<td>6.66</td>
<td>6.66</td>
<td>5.31</td>
<td>3.29</td>
<td>21.92</td>
</tr>
<tr>
<td>686</td>
<td>5.91</td>
<td>5.63</td>
<td>5.37</td>
<td>3.40</td>
<td>20.31</td>
</tr>
<tr>
<td>Other</td>
<td>38.93</td>
<td>28.27</td>
<td>28.81</td>
<td>18.96</td>
<td>114.97</td>
</tr>
<tr>
<td>Total</td>
<td>340.96</td>
<td>343.64</td>
<td>329.56</td>
<td>226.24</td>
<td>1,240.41</td>
</tr>
</tbody>
</table>
Figure 3: Overall waste production in different communes in Jätekukko’s service area.

Example 2:

The waste collection system has highly a stochastic nature and for example the waste formation and collection times may differ in time and place. Before applying any vehicle routing models or algorithms the system and the stochastic components have to be modelled properly and reliably; within the given limitations and restrictions (Figure 4a). During the optimisation of waste collection and transport the vehicle capacity, optimal pick-up periods and optimal container size and type are highly dependent on the predicted waste formations and the travelling times (Figure 4b). Only after analysing waste formations in a certain area the planner is capable of planning the collection reliably. In the more dynamic planning of routing and scheduling the predictions of waste formations can be used together with container filling rate measurements in order to minimize the collection costs.

Example 3:

Due to changes in the population or commercial activities in a region the operations of waste management companies may vary over time, therefore it is important that the information in the operational database is monitored and registered. One of the uses of the modelled data might be the monitoring of complaints, good insight in complaints will allow planners to alter the operations...
in a region in order to restore the agreed service levels or fee additional cost if the service agreed is not matching the client requirements.

Another use of this type of information could be the monitoring of individual trucks and labour. Due to the national and communal regulations usually the actual collection and disposal should be done in a certain limits. Because the collection and disposal time in this case are used as input parameter it possible to control the collection. The same data can be used to support the planning of collection.

Figure 4: A. Collection periods in postal areas; B. Waste formation per container in the City of Kuopio (collection time & predicted mass (kg)).

6 Conclusions and discussion

The study showed that the data stored in the different data management systems of waste management companies can easily be refined to useful information. Additionally, also public and private domains can provide more accurate background information which could be utilised in the waste management information system.

The multi-layer perceptron (MLP), which was used in this study, can be applied to determining and predicting stochastic components like regional waste formation. Exact waste formation in the containers can only be known by measuring it e.g. with the collection vehicle with quite expensive weighting system but by using existing data regional waste formation can be estimated.
quite reliably. Neural networks, like MLP, can also be used to evaluate the needs and profitability of measurements.

The refined data, in combination with raw data, can be used to support the planning of waste management. General GIS techniques and Geo-statistical methods can be applied to visualise and analyse the different modelling results and provide new information based on the combinations of modelled data. The information about the waste formations can be utilised for statistical purposes or for planning and monitoring of waste collection and transport.

Utilising the available data more effectively could enhance the benefits of the necessary investments for the new software systems even more. New software technologies have made the constructing a coupled system between different data management systems and data refining models possible. In time the decisions in the process of planning and monitoring waste management operations could be supported with advantageous and valuable new information.

Acknowledgements

We would like to thank especially Jätekukko Ltd, Ecomond Ltd, Pivotex Ltd, Tietomitta Ltd and Statistics Finland for providing the data and other information needed in this study. Also we would like to thank the other project financers and partners: National Technology Agency of Finland (Tekes), Isalmen Keräysöljy Ltd, Finnish Solid Waste Assosiation, Pirkanmaan Jätehuolto Ltd, Syncrontech Ltd, Kymenlaakson Jäte Ltd, ENW Management Ltd, West Finland Environmental Centre, Helsinki University of Technology Ympäristökonsultointi Hyötyvisio Ltd, Wavin-Labko Ltd, Molok Ltd, Elcard Ltd and Ylä-Savon Jätehuolto Ltd as well as other members contributed to the results namely professor Juhani Ruuskanen, researchers Harri Niska, Toni Patama, Jarkko Tiirikainen, Kari Pasanen, Mikko Heikkinen and Janne Kankkunen and Simo Isoaho, Milla Peltoniemi and Timo Hämäläinen from Tampere University of Technology. Finally we would like to thank our international partners IRADO B.V. & ONS Group for their cooperation and information.

References


