A workflow for improving the availability of routable data (OSM) for logistics in agriculture - using data from Telematics-systems and community-based quality management

JOHANNES LAUER & ALEXANDER ZIPP
CHAIR OF GEOINFORMATICS, DEPARTMENT OF GEOGRAPHY, UNIVERSITY OF HEIDELBERG, GERMANY

Summary
The increasing use of contractors in the agricultural sector and the increasing sizes of farms require the optimization in the field of agricultural logistics. An example for this is the supply chain for biogas plants. An improved transportation planning of the agricultural produced biomass to the plant and the deployment of the nutrient-rich solids is increasing the economic efficiency enormous. Such a system needs a much higher effort in logistics, as there is usually a larger number of service providers, farmers and transportation companies involved. Harvesting is a time-critical task depending from weather and other factors and fast response is a must. Planning a optimized transportation schedule for such a task is very complex. In order to optimize the logistics chain an efficient route planning is a fundamental need. One of the biggest problems to find the best transportation routes is the lack of accurate map data with all ways that can be used for agricultural machines. The commercial mapping companies are not mapping these ways because surveying the tracks cannot be accomplished in an economic manner.

We propose and develop a method for solving this problem. In this paper a way is presented how to get routable maps as additional benefit from surveyed data. This data can originate from e.g. a telematic system that tracks agricultural machines. The generated network data can be used to improve an open source base map (OpenStreetMap). Furthermore a method for the calculation of field gateways is shown and finally some examples of software and services are given that uses the data for solving routing problems.

Introduction
Looking at today road users, navigation systems are common standards in nearly every car or truck. In the field of car navigation, the most part of the relevant data source is collected. Ten percents of the road network changed every year. This data should be collected from time to time to keep the data source on present state. To navigate trucks there are some more attributes needed, such as width and clear high of streets, transit restrictions for hazardous good transportations. The main problem is getting those necessary attributes. There don’t exist any approaches to get this data in a low priced way. For mapping companies, like Teleatlas or Naveq it isn’t attractive to collect this data because the benefit is much lower than the produced costs. Further there is generally not enough expert knowledge about the needed attributes and how to collect them.

State of the art solutions
This subject isn’t new and there are a few, not really satisfying solutions on the market. In the field of agricultural logistics there are partially used truck navigations systems. The advantage is the more comprehensive map data compared to standard car navigation systems. There are some attributes like clear high and the width of roads. And further they’ve mapped avoid roads for hazardous goods or vehicle weight limitations. But this is only for trucks. The vehicle structure in agricultural business is more multifaceted than only some classes of trucks. Agricultural vehicles like harvesters and tractors with a multiplicity of possible accessory equipment have many differential requirements on the road tags. Still there are missed in greater parts the needed field paths that are mainly used by those
vehicles to reach the fields. So it is not possible to calculate a route directly from farm to field.

One system that solved some existing problems, is the FieldNav System of Lacos (http://www.lacos.de). This software is implemented on an ISO-BUS-terminal, how they get used at many agricultural vehicles. The software calculates routes on routable road data up to the end of the data (next point to the field). If this point is reached, the display toggles to a raster map where the actual position is marked. The driver can now visual navigate to the field (without routing).

So the main problem are the missing road attributes, the geometries of the relevant tracks and the needed connection points to drive from track to field (not blocked by vegetation, ditches or not passable because the slope is too high).

**OpenStreetMap**

In the last few years, OpenStreetmap ascends to an usable geographic map source. The data is in parts derived from free map data sets (like data from FRIDA project, US-TIGER data or AND data from the Netherlands. The main part is getting collected by the community. Users can upload own GPS tracks and sketches them with one of the popular OpenStreetMap editing tools. Then they can gives road attributes and load everything up to the OpenStreetMap Server. The advantage of community based systems is the actuality which cannot be reached within a commercial system. Depending on the size and the activity of the community, the hosted data will be very up to date. A current example is the mapping of Haiti after the earthquake in January 2010, that allowed to set up a first emergency routing service (http://opens.geog.uni-heidelberg.de/osm-haiti) within two days after the earthquake in spite of the sparse coverage of data just before the disaster occurred. The knowledge of persons that are familiar with the places they mapped can be a further advantage of community based mapping. Those people can react very closely on changed conditions. For example tagging a destroyed road by intense rain or marking an avoid road. GOODCHILD (2007) describes this phenomenon as “Citizens as Voluntary Sensors”.

**Claas Telematics**

The Claas Telematic system provides the tracking and tracing of geo referenced data from agricultural vehicles, mainly harvesters. These data is getting uploaded to a central server, get processed and showed on a user friendly GUI within a web portal. Beside data about fuel consumption, threshing drum rotation, engine rpm, crop moisture, the logger also saves the actual GPS position. Using these positions we can derive the way geometry that the tracked machinedroves. From this data we can derive map data for routing.

**The field record system**

„Eine Schlagkartei ist eine auf den einzelnen Feld- oder Wiesenschlag bezogene, systematische und geordnete Aufzeichnung über produktions-technische Maßnahmen und Ergebnisse in der Bodenproduktion.”(ALSING 1995, S.584) In short it can be regarded as a database for the field parcels.

To calculate the field connection points between the new street segments and the fields we need the geometries of each field. Those can be exported from existing farm management systems (like AGRO-NET), that used within many farms. This software manages all fields of a farm in a field database, including the geometries. The needed field borders and the data used to identify the field (farm no, field no, field part no, field name) can be exported by the AGRO-NET software to the widespread ESRI –Shape file (ESRI 1998) and to KML (WILSON 2008).
The field borders are also useful to differentiate the parts of the GPS tracks that are within a field from the parts beyond field. By intersecting the field borders with the GPS tracks we can cut off the GPS points within the fields that are useless to derive street data.

**Processing the data**

GPS-Track-logs from a telematic system like Claas Telematics consists of several point measurements. These measurements are normally defective (e.g. GPS variations or missing GPS Signal). There also usually exists more than one GPS track per street (e.g. journey and return, different vehicles). The multiple logged geometries have to be averaged to derive a street geometry. The averaging also causes a spatial correction of the geometries.

Averaging of GPS Data is part of some scientific studies that we want shortly describe. 

Roth 2008 uses a theoretical probabilistic model to calculate the track averaging. Compare the histograms he finds similar track parts. After that the similar parts averaged by using a probabilistic model, that includes the variances of the GPS measurements.

To average GPS tracks from pedestrians, Krämpe (2006) uses clustering. He clusters the measurement points and average them using the centroid of the cluster. After that he connects the centroid points and gets the averaged line. The problem of this approach is the missing knowledge of coherence of several points.

An interesting approach is the one of Morris et al. 2004. They use algorithms from graph theory. From all GPS tracks they build a street graph. The nodes represent the intersection points of the tracks. The connection lines between two nodes represent the graph edges. Those connection lines (normally there is more than one that connects two nodes) will be reduced by graph algorithms. Morris et al. 2004 describe four algorithms for graph reduction.

Parallel reduction reduces two parallel edges (each start- and endpoint must be the identical) to one edge. The edges will be reduced, if the geometry between the two nodes is nearly similar. The similarity is calculated with the Hausdorff distance between the two edges. If the edges satisfied the similarity condition, the algorithm calculates pairs of the edge points and determines an average coordinate. The geometry starting at start node, ending at end node using the averaged coordinates will be the new graph edge. The general case of parallel reduction is face reduction. Face reduction enlarges the used edges to more than two. This case appears if more GPS-tracks intersect. Depending on variances of the several tracks there will be more intersection points (Figure 1).

![Figure 1: Face with grade four; the white squares represent the nodes (Source: MORRIS ET. AL. 2004)](image)

At first we look for the nodes with the highest distance to each other. These nodes are connected by several tracks. The tracks are getting concatenated (assembled at the nodes between). The result is the input for the parallel reduction. Using this algorithm, we calculate the connection R between nodes a and b. After that, the rest of the nodes connected to the spatial nearest node of R. the third algorithm is to delete dead ends (e.g. tracks that are...
not belong to the road system, like parking in garage doorways). The fourth algorithm, serial reduction, erases nodes with degree \( \leq 2 \). Nodes with a degree less or equal two are no crossing points or connection points to other tracks. If the node is member of two different polylines, the lines will be assembled and the node will be erased from graph structure. This algorithm eliminates the nodes, which are produced by the other three algorithms. An implementation of these algorithms is within the TopoFusion software of Alan and Scott Morris.

**Implementation of the workflow**

From raw data to routable road data, there was developed a workflow (Laue 2008). This workflow was largely implemented prototypically as an independent software system.

![System architecture of the realized software solution](image)

Figure 2: System architecture of the realized software solution

From raw data to routable road data, there was developed a workflow (Laue 2008). This workflow was largely implemented prototypically as an independent software system. From the external databases the data will be imported via standardized interfaces into a local database. The Claas Telemetrics data will be converted from CSV-type (Shapranovich 2005) (export of the Telematics database) to the GPS XML data format GPX. The GPX format is the interface to the track database. That brings the benefit to import other GPS-data with only sparse work into the database.

The field borders will be exported from the AGRO-NET software in the kml format and can be imported via the implemented interface into the local PostGIS database.

To import the OpenStreetMap Data into the database we used the free osmosis OpenStreetMap tool.

The third extern component is the Java OpenStreetMap editor JOSM. This software provides the ability to make last manual changes and corrections on the derived data. In the last step, the data is uploaded to the OpenStreetMap Server via the OpenStreetMap API.
Figure 3: Workflow from raw data to routable OpenStreetMap Data

Figure 3 shows the whole procedure from raw data to the routable map. In a first step, the raw data will be exported from the miscellaneous databases. The Claas Telematics tracks are preprocessed and imported into the PostGIS database. Further the field borders will be exported from the AGRO-NET software and imported via the KML-Converter into the database. The field specific Tags (field no, field name, crop year, ...) are parsed and saved in the database. The OpenStreetMap Data within the bounding box of the GPS-track are downloaded. This part and the importing of the OpenStreetMap Data into the database is solved by the osmosis software. A SQL script eliminates the not needed data (POIs, land cover, ...).

Via geometrical intersection the field connection points are calculated and stored in the database. In the next step, we differentiate between GPS-points in field and on the road. For this, an intersection of the GPS points with the field polygons and marked the points within the polygons is applied. Those points will no longer be used to derive road data.

After that we the TopoFusion Software to calculate a graph from the GPS-tracks and the OSM-street data. For this, the OpenStreetMap data will be exported to the GPX-format and
imported into the TopoFusion software. With the graph reduction algorithms the tracks are averaged and so road network is getting reduced. The road network is stored in GPX-format and converted into the JOSM OpenStreetMap data format. This data type allows negative IDs for each object and some tags that shows the status of an object (e.g. "changed", "deleted"). So it is possible to update the server side OSM-data. Inside the converting algorithm we reduced the point density by the Douglas-Peucker Algorithm (DOUGLAS, PEUCKER 1973), because the density of the GPS measure points (each five seconds one point) is too high. Furthermore with this, the manual correction will be easier. Too short way parts and single points which are not connected to the road network will be deleted.

As way attributes we use the knowledge about the vehicle that provides the GPS-data. By knowing the machine type we added the attributes way width, clear high and maximum weight. The attributes derived from the data sheets of the vehicles (Claas Lexion 570, 580 and 600) (CLAAS 2008). We set the track type to the hierarchic lowest type (Tracktype – Grade1) that can be used by this vehicle. Those attributes gives minimal constraints for the ways. That means, for further route planning: All roads that was driven by a Claas Lexion 600 harvester can be driven by this vehicle or a smaller machine.

The generated data are imported into the JOSM-Editor and manually corrected. In this step, the new roads are connected to the existing OSM-data. Furthermore it is possible to close apparent small gaps between track parts and complete the road network.

In the last step the data is uploaded via the OSM-API to the OpenStreetMap Server.

**Using the generated data for routing**

The data generated through the woflow can now be used for routing. Based on the well known OpenRouteService (Neis & Zipf 2008) – an implementation of the OpenGIS Location Service Specifications (MABROUK ET. AL. 2005) - a specialized version and user interface was realized that uses the relevant data for routing the agricultural machines. The calculated field access points are available in the database and can be used similar to adresses. The user interface and map shows the relevant information about Ids of field parcels etc, that can be used a input for routing (figure 4).

![Figure 4](image)

*Figure 4:* Route-planning from parcel to parcel based on a special version of OpenRouteService
5 Conclusion and further work

This paper demonstrates an approach that automates to generation of street networks which are appropriate for routing to a large degree. Only the integration into the existing OpenStreetMap data set needs manual editing in order to solve conflicts with existing street segments. A major factor for the automated generation of the street network is the availability of a large set of raw data such as GPS-tracks. It has been shown that the data of telematics-systems such as Claas Telematics can be used to generate a street network for routing purposes in a cost-efficient way. Attributes of the street segments can be generated from data provided by the different agricultural machines that are being tracked. The different types of machines are known and available in a database. The measurements and weights of the machines allow to infer basic attributes with respect to routing for each street segment.

If a street has been used by a distinct type of machine it can also be used by other machines of the same type or of smaller measurements or less weight. A similar classification of the trafficability of roads is being used in forestry logistics (HAUCK 2003) and can be adapted.

OpenStreetMap offers a good data base for the integration of new ways and streets. The attributes can be chosen freely by the user and are not restricted as in commercial datasets such as TeleAtlas oder NAVTEQ. The workflow introduced offers a fundament for generating street data suitable for routing applications from GPS-tracks. The individual components such as weighting og tracks or the calculation of the field parcel access points are encapsulated in component that can improved or exchanged individually.

In general we can state that a geodata set that is kept up date by a user community can have very recent and actual data. In our case the new community of farmers and companies using the system can check and improve the data set on their own. The realized method and workflow to generate street networks for routing can be adapted to other types of routing applications if they also do have access to some kind of sensor information. For example there exists a first version of a routing prototype for people with wheelchairs or similar restrictions regarding mobility (http://www.rollstuhlrouting.de). They do not have telematics systems, but more and more smartphones with velocity sensors or similar are being used.

References

DOUGLAS D., PEUCKER T. (1973), Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, Cartographica: The International Journal for Geographic Information and Geovisualization 10, 112 (01 October 1973).
HAUCK B. (2003), Konzept zur Entwicklung von bundeseinheitlichen Geographischen Standards für die Holzlogistik, kwf Bericht Nr. 4/2003


LAUER (2008), Verbesserung der Datengrundlage für die Routenplanung im Bereich landwirtschaftlicher Logistik auf Basis offener Geodaten, Masterarbeit, Uni Bonn - Geographisches Institut


ROTH J. (2008), Extracting line string features from GPS logs, 5. GI/ITG KuVS Fachgespräch "Ortsbezogene Anwendungen und Dienste", Nürnberg, Sonderdruck Schriftenreihe der Georg-Simon-Ohm-Hochschule Nürnberg Nr. 42