

ADDING VALUE OF CAR NAVIGATION- VISUAL APPROACHES TO KNOWLEDGE DISCOVERY

A.O. Dogru, H. Demirel, N. Ulugtekin

ITU, Faculty of Civil Engineering, Department of Geodesy and Photogrammetry, 34469 Maslak Istanbul, Turkey -
(dogruahm, hande.demirel, ulugtek)@itu.edu.tr

KEY WORDS: Cartography, navigation, knowledge base, data mining, user

ABSTRACT:

Car navigation systems are discussed in this study in terms of knowledge discovery, data mining and data quality. Since data, information and knowledge are inseparable levels of human cognition, integration of these methods aids to understand the real-world patterns for car navigation systems and add value. The car navigation system is multi-purpose, multi-scale, multi-temporal and varying users requires different levels of abstraction and visualization of spatial data and information. Hence, the data discovered, mined and the quality should be visualized on timely manner, without leading any confusing and for varying users and purposes. In this concept whole process was evaluated with in the context of research questions asked by Gahagen et.al (2001), where visualization problems and proposal for them were emphasized.

1. INTRODUCTION

Data mining and knowledge discovery are challenges in current GIS research, since the linkage of disparate datasets together across place, scale, time, theme and discipline is common. Consequently, uncovering and understanding real-world patterns, or real-world processes that can be found in the data stored in these databases presents a difficult challenge; they are vast, dense and may span many different areas of expertise (Gahegan et al., 2001). These methods are all considered in terms of knowledge discovery from the databases (KDD) and several newly introduced studies have been conducted (Buttenfield et al., 2000; Gahagen et al., 2001; Shekhar et al, 2002; Demsar et al, 2006). However, there are uncertainties in spatial data, and they may directly or indirectly affect the quality of spatial data mining (Wang et. al, 2002). Frequently, the spatial data are considered to be near perfect representations of the real world, and only rarely are they considered having an associate degree of uncertainty in either their location, topological and non-geographic attributes (Burrough, 2002). In order to prevent misleading decisions, spatial data quality issues should be addressed during data mining and knowledge discovery.

The car navigation systems, being one of the most mature and popular personal geo-services, involve complex design and modeling of spatial data including several issues still unsolved. The design process requires specific attention, since the usage is user dependant; used maps are multi-purpose and multi-scale. This study aims to investigate the possibilities of an integrated approach for handling knowledge discovery, data mining and spatial data quality for car navigation systems. Specifically car navigation maps are detailed, since varying user requirements are assessed. The relation between knowledge discovery and spatial data quality is examined and current status was evaluated by considering research questions of knowledge discovery asked by Gahegan et.al (2001). This effort of relating geographic visualization -cartography-, spatial data quality and knowledge discovery, aids to introduce new advances and will add value to the car navigation systems.

2. PROBLEM STATEMENT

Spatial knowledge and its acquisition in car navigation systems were divided into well-accepted levels, where these were defined as *landmark*, *route*, and *survey* (Stern and Leiser, 1988). Landmarks are the distinctive objects and/or views within the environment, which have great potential to support travelers in navigation tasks to discover the knowledge from the instructions transferred by the navigator. *Route* knowledge is considered as memory for procedural linking of landmarks, including order, inter-landmark distances and required actions. Finally, the *survey* knowledge is the memory in which landmark and route knowledge is integrated into a configurational, map-like whole (Brunett and Lee, 2005).

Car navigation entails route planning, the recognition of locations and landmarks on the route and negotiation of complex junctions. These activities require data at a wide range of scales and at different levels of abstraction. Hence, car navigation systems involve huge amount of spatially referenced data and its representations. Users need various types of maps in different situation and when usability of maps is concerned, one of the main issues is that the user has the right type of map, at a suitable scale and with the symbology adapted for the specific usage situation (Nivala and Sarjakoski, 2003). According to the map usage purpose, the content and symbolization technique of the map differentiate in terms of generalization. Additionally, mathematical calculations, such as finding the shortest or simplest path or optimal route are executed. Hence, accurate geometry in navigation is essential, where shapes are of much less importance. Besides, the topological consistency should be maintained while designing these maps.

In the car navigation systems, the expectation of the user is high, being easily, rapidly and accurately navigated by the system; however the spatial knowledge is inherently incomplete, or partial. Since data, information and knowledge are the three inseparable levels, the transmitted information and the comprehended knowledge is incomplete, involves uncertainties. An example of a junction uncertainty can be

given from a conducted study (Ulugtekin et al., 2004). Two different representations of a junction for varying users were selected for the same motorway, a part of Transit European Motorway –TEM. These representation levels are presented in Figure 1: (a) represents the two-lane motorway, represented with two lines, where the (b) part of the same figure represents the same world reality as a single line. For the same junction, two different shortest paths, which are displayed as bold lines, were calculated for the same origin and destination. The shortest path and the route time for the (a) option is obtained as 3,016 km and 1 min 26 sec, where for the (b) option route length was 2,605 km, and route time was 1 min 44 sec. The driver in this case should choose the (b) option as it is the shortest path, however as it is understood from the marked arrows presenting the directions of the roads at Figure 1.b, in reality some roads are one-directional and not appropriate for the entrance of the cars on this road network.

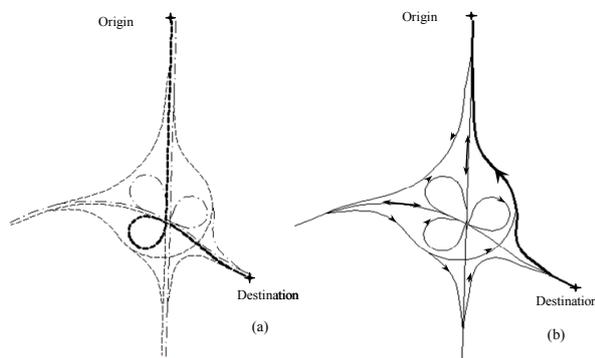


Figure 1: Calculated Shortest Paths for Different Representations of the Same Road Network

The results obtained from this example, are originated from data acquisition, data modeling, cartographic generalization, visualization and presentation of data/map quality. Quality data may be attached to elements at different generalization levels on geometric dimension, such as geometric primitives (ex: line segments composing road), complete geometric shape of an object (ex: complete network of Transit European Network), all geometric shapes of a same object class (ex: all polygons representing roads), and dataset or all geometric shapes of all object classes (ex: car navigation map of Istanbul). Quality may be attached to elements at different levels of detail on the semantic dimension, where levels are value, domain, attribute, object and semantic. For the temporal dimension, different levels of dimension is available being; temporal primitives, complete temporality of an attribute (evolution) or an object (existence); complete temporality of an object, all temporalities of a same object class or of same attribute type of a same object class and dataset or all temporalities of all object classes.

3. KNOWLEDGE DISCOVERY FOR CAR NAVIGATION SYSTEMS

Within the car navigation system, generally driver or another user submits queries for destination. The driver may either provide the location of the vehicle and destination, or a specific route was required regarding the landmarks or points of interest (POI). Within the navigation system a geographic database is available for semantic and spatial queries and network analysis. Network analyses use generalization, where results of queries also benefits from cartographic visualization and symbolization.

The process is a feed-back cycle, since the driver is prompted to take the next action based upon receiving a particular view of the environment and understanding. A simple flow chart, which starts with the acquisition of the positional data by the entry of the driver or by the use of spatial or semantic queries executed through geographic database and ends with the vehicle control, for the system operation is provided in Figure 2. In this Figure, dashed arrow lines represent the human interaction of the system while solid arrow lines represent the common process of the designed system.

System is based on the interaction of human (user/driver) and machine (system). The system uses an input data of location determined by the user or obtained by semantic or spatial queries executed by the user. The knowledge discovery from database in car navigation systems starts in this stage while obtaining the positional data from the geographic data base and it continues through the database generalization to derive the route data and instruction generation to navigate the driver. Whole process is interactive; especially vehicle control depending on the instructions displayed to the driver is the last and the most important part of this interaction since the confidence of the travel depends on this usable knowledge derived from the final instructions. Hence, these systems require both highly accurate geometric and semantic data of which relations should be carefully established. The integrated approach of KDD, spatial data mining and spatial data quality methods were applied in car navigation systems. Since the conceptual knowledge processing aims to provide methods and tools for acquiring, reasoning with, and representing knowledge, and for making it available to human communication (Stumme, 2003), it should be considered as a subject associated with cartography, which is the science of modeling and producing maps as a communication tool of spatial information by manipulating spatial databases. Several research questions on knowledge discovery asked by Gahegan et al. (2001) are adapted to car navigation systems for covering varying requirements in this study.

“From the perspective of the scientific method, what are the different tasks involved in a knowledge construction process applied to a geographical database and which visualization techniques are the most suitable for them?”

Visualization techniques used in knowledge discovery process are geometric projection, map-based, chart-based, iconographic, pixel-based, hierarchical, and network techniques (Hinneburg et al. 1999; Gahegan et al. 2001). Visually-based approaches in knowledge discovery can offer an opportunity for humans and machines to work together in constructing and evaluating the objects and relationships required by the analysis phase, making the best use of the abilities of each car navigation systems basically use the map based techniques for the visualization (Sui and Meng, 2002). Maps are used as communication tools in these systems to present the world reality to the user, driver.

“To what extent do the different visualization techniques affect logical inference?”

Since there is a clear a human-machine interaction in navigation systems, different visualization techniques are used in these systems to improve the logical inference of the user. As these techniques affect the perception of the user, they should be applied carefully for car navigation systems. However, as the level of the user perception increases, the success of the process

increases simultaneously. For instance, simple “Line graphs” can be used to present the traffic density data or to show the

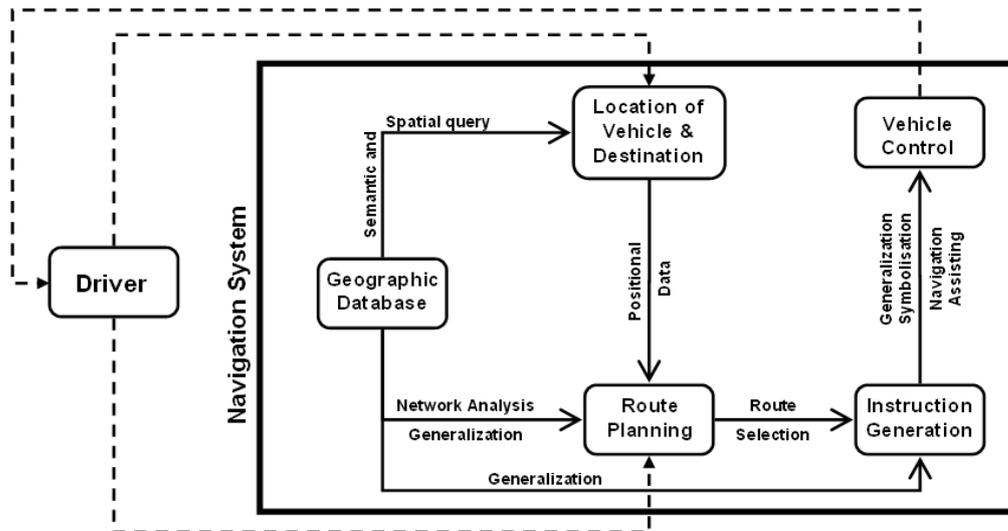


Figure 2. The Car Navigation System Operation

relation between velocity and distance to the destination. Since this technique is not sufficient to easily understand, it will not be consistent with one of the basic aims of the navigation systems that is presenting the world reality visually as simple and as understandable as possible. Hence, it should only be used while querying the data not to reduce the logical inference of the driver while traveling. On the other hand, visualization techniques just as “box plots”, “parallel coordinate plots”, and “chernoff faces” are too complex for transferring the knowledge to the user in a car navigator. If they are directly used in navigation systems, they will decrease the logical inference so this will increase the number of user errors of navigation process. Pictorial symbolization of the POI or use of orientation arrows for explaining the direction in addition to the audible warning services could be given as an example to improve the cognition for a confident navigation.

“To what degree should we rely on automated reasoning in producing the visual scenes used for data exploration (automated scene construction) and what implications does reliance bring?”

Users need various types of maps in different situation and when usability of maps is concerned, one of the main issues is that the user has the right type of map, at a suitable scale and with the symbology adapted for the specific usage situation (Nivala and Sarjakoski, 2003). The user needs may be totally different in various map usage situation (e.g. city or road navigation in national or international scale). Therefore, all maps are designed for different purposes. Aim of the map mainly affects its content and symbolization techniques that will be used in terms of generalization. In addition to the user requirements to easily reconstruct the world reality, navigation systems itself requires different levels of representation especially because they uses small display devices in application. An automated system to produce various representations with different levels of details of world reality is needed in this concept. This system should also enable automated update and generalization implementations. Since the miscognition of the presented map results in the deficiencies in confidence, it should be tested to obtain most reliable systems. The reliability of the system could be achieved through visualization of the uncertainties.

Studies still continue to obtain the most reliable network dataset. The data set will be used for more reliable network analyzing and comprehensible user interfaces. (Dogru and Ulugtekin, 2004; Ulugtekin and Dogru, 2005).

“To what extent, and for what tasks, must a data mining process consider the spatial and temporal properties of the data along with other attributes?”

Time is an important factor especially in the car navigation systems. Changes in navigation conditions should be updated regularly and, if possible, in real time. For example, in case of traffic jams, the car navigation system could provide alternative solutions for the driver via updating the topology information. Hence, the designed data model for car navigation system should permit such updates. Additionally, day and night-drive visualization issues and also seasonal view of the environment should be considered.

“How does the presence or absence of different data sets affect the success of visual data mining and what is the optimum arrangement of data to locate a given object or relationship?”

Navigation systems aim to present the world reality with optimum amount of data because of its small display medium in practice. Therefore, both the presence and the absence of the data will effects the efficient use of the system. The presence of another linear object, such as pipe-line, is confusing for the driver, while navigating. Parallel, the absence of the railways will be risky. Additionally, too much detailed data will not be understandable and readable. If the level of details is not appropriate the user could not reconstruct the mental map. Hence, the navigation maps should be well organized and generalized to eliminate such effects. The mapping data should be arranged in terms of multiple representations that will be used for different scales and aims. Additionally the relations of the objects represented in these levels should be formalized and maintained for automated data derivation processes. For example, Ulugtekin and Dogru (2005) organize the navigation data in five different representation levels for small display car navigation, which is presented in Table 1.

Table 1. Map details for car navigation

Levels	Scale	Detail
International	Depends on coverage	National boundaries, international motorway, related text.
National	1 : 8 000 0000 (Turkey)	National boundaries, international motorways, state roads, city boundaries, related text.
City	1 : 250 000 (Istanbul)	Administrative boundaries, motorways and the state roads, railways, airports, related text.
County	1 : 25 000 – 1 : 50 000	POI (with symbols), motorways, state roads, arterial highways, land use (settlement or green area), related text.
District	1 : 5000 – 1 : 10 000	Land use (buildings, green area) POI, district and building block boundaries, all roads (all roads are represented as aerial symbols), related text.

4. CONCLUSION

The spatial data visualization techniques aid to integrate knowledge discovery from database, data mining and data quality. The association between these subjects can be formulated for the car navigation systems as follows; via generalization of multi-scale maps and databases, various interfaces for varying users, different visualization techniques to improve the logical inference of the user and optimum amount of information via generalization. For a successfully operating car navigation system, the designed structure should also consider the subjects covering aim of the use, various users profiles, confidence of the driver and driving conditions such as day or night view, selected road data, weather, and etc. Additionally, consistency and up to datedness of the geometric data of the road network is also determining factors for the success of the navigation process. All this requirements can directly or indirectly be covered by the use of a visualization technique of which outputs can easily represent and communicate the world reality to the user. Use of map based techniques that can be totally considered as a knowledge discovery process with its generalization, symbolization, and presentation are indispensable opportunities in this manner. Since knowledge discovery techniques and related fields empower the exploration and facilitate geographic visualization, this research will aid to establish the linkage between these fields on a car navigation system application.

REFERENCES

Burnett, G.E., Lee, K., 2005. The effect of vehicle navigation systems on the formation of cognitive maps, Final draft of paper for G. Underwood (Ed.) *Traffic and Transport Psychology: Theory and Application*. Elsevier

Burrough, P.A., 2002, GIS and Geostatistics for Environmental Modelling, edited by Shi, W., Fisher, P., Goodchild, M.F., Spatial Data Quality, Taylor&Francis Inc., pp. 18-35.

Buttenfield, B., Gahegan, M., Miller, H., Yuan, M., 2000. Geospatial Data Mining and Knowledge Discovery , A UCGIS White Paper on Emergent Research Themes. http://www.ucgis.org/priorities/research/research_white/2000whitepapersindex.htm

Demsar, U., Krisp, J.M., Kremenova, 2006. Exploring Geographical Data with Spatio-Visual Data Mining, edited by Riedl, A., Kainz, W., Elmes, G., Progress in Spatial Data Handling, 12th International Symposium on Spatial Data Handling, pp.149-167.

Dogru, A.O., Ulugtekin, N., 2004. Junction Modeling in Vehicle Navigation Maps and Multiple Representations. Proceedings of the XXth Congress of ISPRS, Vol:35, Part B4, Com.IV, Istanbul, Turkey, pp: 244-248.

Gahegan, M., Wachowicz, M., Harrower, M., Rhyne, T. M., 2001, The integration of geographic visualization with knowledge discovery in databases and geocomputation. *Cartography and Geographic Information Systems*, special issue on the ICA research agenda.

Hinneburg, A., Wawryniuk, M., Keim, D.A., 1999. HD-Eye: Visual Mining of High-Dimensional Data, IEEE Computer Graphics & Applications Journal, Vol. 19, No. 5, pp. 22-31.

Nivala, A. M., Sarjakoski, L.T. 2003. An Approach to Intelligent Maps: Context Awareness. Workshop "HCI in Mobile Guides", Udine, Italy.

Shekhar, S., Lu, C.T., Liu, R., Zhou, C. 2002, CubeView: A System for Traffic Data Visualization, The Proceedings of the Fifth IEEE International Conference on Intelligent Transportation Systems, <http://www-users.cs.umn.edu/%7Ectlu/PaperTalkFile/its02.pdf> (accessed 14 Feb. 2007)

Stumme, G., 2003. Off to newshores: conceptual knowledge discovery and processing, Int. J. Human-Computer Studies, Elsevier, pp. 287-325

Stern, E., Leiser, D., 1988. Levels of spatial knowledge and urban travel modelling. *Geographical Analysis*, 20(2), pp. 140-155.

Sui, L., Meng L., 2002. Exploration of the road database with data mining methods. IAPRS, Vol XXXIV, Part 2, Commission II, Xi'an.

Ulugtekin N., Dogru, A.O., 2005. Map Design for Navigation Purposes. XXIInd International Cartographic Conference, CD. A Coruna, Spain.

Ulugtekin N., Dogru A.O., Thomson, R., 2004. Modeling Urban Road Networks Integrating Multiple Representations of Complex Road and Junction Structures, Proceedings of the 12th International Conferences on Geoinformatics, Gavle, Sweden, pp. 757-764

Wang S.L., Wang X.Z., Shi, W.Z., 2002. Spatial Data Cleaning. *Proceedings of the First International Workshop on Data Cleaning and Preprocessing*, edited by Zhang Shichao, Yang

Qiang, Zhang Chengqi, Maebashi TERRSA, Maebashi City,
Japan, pp.88-98