A Spatial Decision Support System for Property Valuation

Katerina Christopoulou, Muki Haklay

Department of Geomatic Engineering, University College London, Gower Street, London WC1E 6BT
Tel. +44 (0)20 76792740
Fax +44 (0)20 73800453
katerina@ge.ucl.ac.uk

1 Introduction

In recent years it is increasingly recognised that land management is a multi faceted challenge, in which social, economic and physical dimensions must be considered. In addition, the call for sustainable development (WCED, 1987) increases the importance of effective land management at all levels. Hence, there is need for new approaches that will be able to handle this complexity. Among various decisions at the operational level of Land management, property valuation is considered to be a challenging process. For this reason, several methods have been developed to assist in property valuation. However, in the majority of these methods several important factors that affect property value are under-represented. One such factor is the location of the property and its relationship to other spatial entities. Another is the large amount of data that must often be taken into account to extract useful knowledge that can potentially assist in valuation.
This paper discusses these issues and proposes the application of knowledge discovery methods in a computer assisted property valuation framework, currently in the first stages of development. The paper also examines the incorporation of the location of the property into the framework

2 The Property Valuation process

When considering the important economic value of the land, the process of property valuation stands out as a significant element in land management. Property Valuation is the estimation of the market value of a property. Property valuation for a specific purpose is a non-trivial process since it involves the consideration of a variety of underlying factors of the market and the way they affect the value of the property at a given time. Such factors may include governmental policies, geographical factors or even factors such as fashion; season etc. Property valuation also depends on the purpose (e.g. sale, taxation, financing etc.) and the type of the property (residential or commercial), for which, it is exercised.

It is widely recognise that there are five main standard valuation methods (Lawrance *et al*, 1971), of which the Comparative Method is considered as the most reliable but also heavily dependant on the quality of the selected comparables. Recently, a number of techniques for determining the value of property by trying to mimic the thought process of the actors of the market have been developed (Pagourtzi *et al*, 2003).

The successful application of a valuation method is heavily dependant on the quality and the variety of the data. Among the factors (legal, physical, economic) that influence the value, location is considered to be of outmost importance. Location in terms of proximities to infrastructure or amenities, neighbourhood quality, topology plays an important role in the formation of the value therefore can generate variations in price among similar properties. Although the importance of location in the value of a property is widely recognised (Kauko, 2003), current valuation processes consider location implicitly (Wyatt and Ralphs, 2003). This is mainly due to the modelling

difficulties that relate to the wide variety of spatial factors and their interactions that may or may not affect the property in question at given instances of time (Deddis *et al*, 2001). As a result, in the majority of the cases the incorporation of location is based on a valuer's knowledge and experience (Wyatt & Ralphs, 2003). Examples of research projects that face these challenges include the hedonic modelling with location (Orford, 1999); Artificial neural networks (Jenkins *et al*, 1998); Accessibility Index (Wyatt, 1995) and others. However, there is still a need for new and more efficient and accurate location based valuation models (Deddis *et al*, 2001).

3 Geographical Knowledge Discovery and Spatial Data Mining

Knowledge discovery in databases is a non-trivial process that involves the discovery of valid, novel, potentially useful and ultimately understandable patterns in data (Fayyad *et al*, 1996A). A number of methodologies have been proposed for the knowledge discovery processes that are mainly variations of the main scheme: data preparation-data mining-interpretation of the extracted knowledge. In this study the methodology proposed by Fayyad *et al* (1996B) will be adopted and involves five basic activities: selection, pre-processing, transformation, data mining and interpretation. These five activities are also relevant when the analysis is focusing on geographical information (Miller & Han, 2001).

In short, the actions that are being taken in each activity are as follows: Selection, Pre-processing and Transformation are preparatory stages which lead to the core knowledge discovery process – data mining. Data mining is commonly broken into three more sub-stages. The first involves the search and identification of the generic pattern type. The second step includes the identification of the specific data mining technique that is relevant for the problem at hand. The final step is the application of the selected technique for pattern search. The final stage of knowledge discovery is the interpretation/evaluation in which visualization techniques are being used and the discovered knowledge is either integrated into a knowledge-based database or used in a report. Figure 1 summarises the knowledge discovery process.

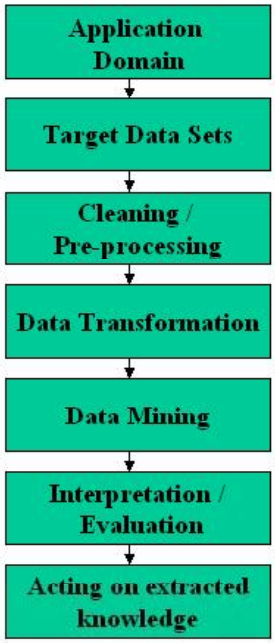

Figure 1: Knowledge Discovery Process
(after Fayyad *et al*, 1996B)

In many applications, these activities are not carried out in a sequential manner, but rather iteratively.

When we look at the specifics of spatial data mining, two are the high-level aims that have been identified by Fayyad *et al*. (1996B): prediction and description of the datasets. To accomplish these aims the appropriate data mining task must be selected and applied to the dataset. Data mining tasks include (Miller & Han, 2001): Segmentation which can further analysed to Clustering and Classification, Dependency Analysis, Deviation and Outlier Analysis, Trend Detection and finally Generalisation and Characterisation.

Although knowledge discovery is a quite well established area in conventional databases its application in spatial databases is a new but very promising area for research (Ester *et al*, 2001). The complexity of geographical phenomena (Gahegan, 2001) along with the large size of spatial datasets not only justifies the application of knowledge discovery to spatial datasets but they also make it highly attractive.

Adoption of such an approach into geographical problem solving presents a number of challenges. Among them are the modification of existing or the development of new algorithms that can handle spatial data, the representation and storage of the extracted Knowledge into spatial databases and further incorporation into the model, the role of visualisation in such a methodology and also the mining of disparate and different in format, data (Koperski *et al*, 1998; Gahegan, 2001; Miller, 2004)

4 Computer assisted property valuation framework

As the previous sections demonstrated, knowledge discovery is applied in complex problems to reveal previously unknown information. To develop explicit and accurate location based valuation models is difficult since the valuation process involves the valuer's particular knowledge and experience. The application of knowledge discovery techniques to property valuation, where location is one of the main factors, is a promising approach to accurately automate the valuation process and this is the core of this particular research.

The impact of location on property valuation is suitable for geographical knowledge discovery as it fulfils two core conditions. Firstly, it is a 'non-trivial' process. Today valuers take location into account by employing the background knowledge of a limited geographical area and experience and primarily compare the known value of other similar properties within this area. Additionally, in most computer-based valuation models the incorporation of the location usually follows oversimplifying approaches. This data-driven approach is not tied to theories (e.g. hedonic modelling) that attempt to explain the role of location in the property value therefore it doesn't require *a priori* assumptions about the variables. Knowledge discovery will help in revealing this missing information and by using the appropriate representation it can be incorporated into the valuation model. Secondly, valuation-related datasets offer the desired diversity and volume. Valuation modelling may include data that involve the property in question, locational characteristics, environmental characteristics, transactional data and so on. These potential datasets apart from the fact that vary in type may be also distributed. As mentioned above some of the research needs in spatial data mining involve the handling of several and diverse data types.

4.1 Conceptual Analysis

A conceptual analysis of a property valuation decision support system that uses Knowledge Discovery for residential property valuation is shown below. Figure 2 illustrates the conceptual architecture of this system.

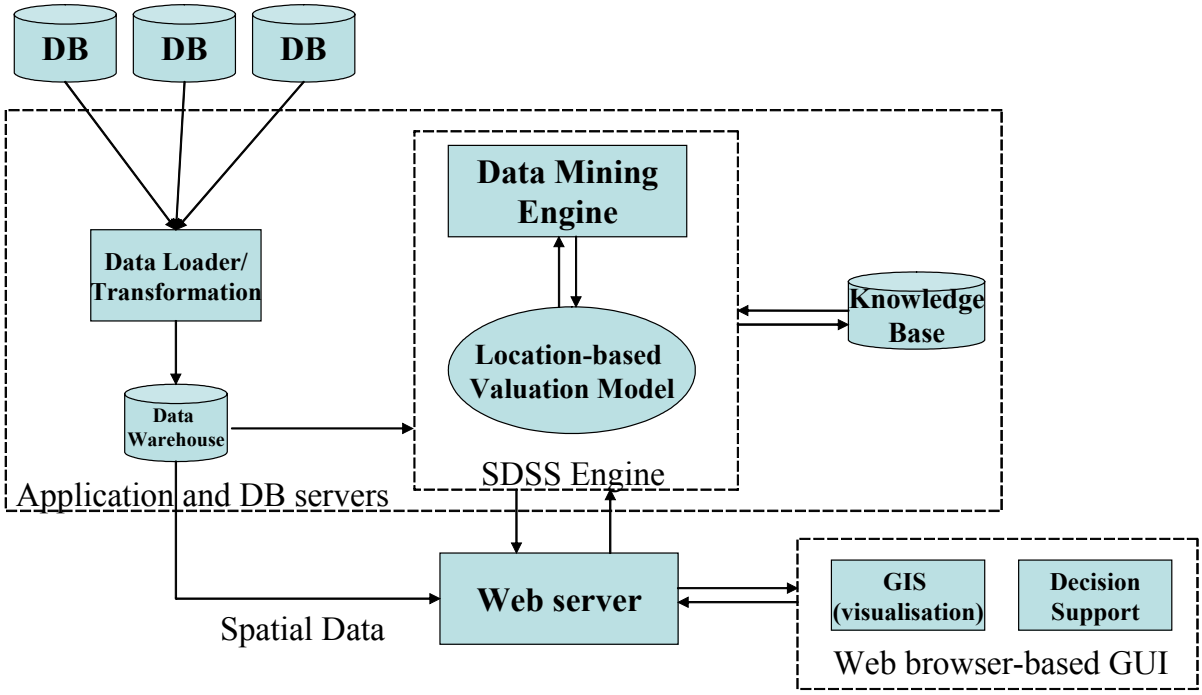

Figure 2: Conceptual Architecture

This system is organised in three main components: Data Loader, SDSS Engine and finally the visualisation component.

- **Data Loader component**: is responsible for getting the data from various disparate sources, and places it in the Data Warehouse.
- **SDSS engine:** is an implementation of the location-based valuation model and the data-mining algorithm. The data-mining algorithm extracts knowledge from the data in the data warehouse and performs property value estimates.
- **Visualisation component**: in this particular diagram a web-based implementation is shown.

The user interacts with the system through a graphical user interface. This is either implemented as a traditional desktop application or as a web based interface utilising a standard web browser. Figure 2 shows the web-based implementation, which is realised by the use of a standard web server. There can be two main function groups available to the user. The first relates to the functions for property valuation decisions. The second relates to visualisation.

The diagram also shows direct queries to the data warehouse, which relate to accessing complementary information to the property value estimations delivered by the SDSS engine and also to accessing spatial data.

Acknowledgements

This research project is supported by the Greek State Scholarships Foundation (IKY). Their support is greatly appreciated.

References

DEDDIS, W.G., MCCLUSKEY, W.J., LAMONT, I., 2001, The Application of Spatially Derived Location Factors Within a GIS Environment, World Valuation Congress, Singapore.

ESTER, M., KRIEGEL, H-P., SANDER, J., 2001, Algorithms and Applications for Spatial Data Mining. In Miller, H.J., Han, J.(eds), Geographic Data Mining and Knowledge Discovery, Taylor and Francis, p.131-159.

FAYYAD, U., PIATETSKY-SHAPIRO, G., SMYTH, P., 1996A, From Data Mining to Knowledge Discovery: An Overview. In Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (eds), 1996, Advances in Knowledge Discovery and Data Mining, AAAI Press, p1-30

FAYYAD, U.M, PIATETSKY-SHAPIRO, G., SMYTH, P., 1996B, From Data Mining to Knowledge Discovery in Databases. AI Magazine, p37-54

GAHEGAN, M., 2001, Data Mining and Knowledge Discovery in the Geographical Domain. A National Academies white paper, Intersection of Geospatial Information and Information Technology

JENKINS, D.H., LEWIS, O.M., ALMOND, N., GRONOW, S.A., WARE, J.A., 1998, Towards an Intelligent Residential Appraisal Model. Journal of Property Research, vol. 16, no.1, p.67-90.

KAUKO, T., 2003, On Current Neural Network Applications involving Spatial Modelling of Property Prices, Journal of Housing and the Built Environment, Vol. 18, p. 159-181.

KOPERSKI, K., HAN, J., ADHIKARY, J., 1998, Mining Knowledge in Geographical Data, Communications of ACM

LAWRANCE, D.M., REES, W.H., BRITTON, W., 1971, Modern Methods of Valuation of Land, Houses and Buildings, The Estates Gazette Limited.

MILLER, H.J, 2004, Geographic Data Mining and Knowledge Discovery. In Wilson, J.P., Fotheringham, A.S.(eds), in press, Handbook of Geographic Information Science, Blackwell.

MILLER, H.J., HAN, J., 2001, Geographic Data Mining and Knowledge Discovery-An Overview. In Miller, H.J., Han, J.(eds), 2001, Geographic Data Mining and Knowledge Discovery, Taylor and Francis, p3-32.

ORFORD, S., 1999, Valuing the Build Environment: GIS and House Price Analysis, Ashgate Publishing Ltd.

PAGOURTZI, E AND ASSIMAKOPOULOS, V., HATZICHRISTOS, T., FRENCH, N., 2003, Real Estate Appraisal: a Review of Valuation Methods, Journal of Property Investment and Finance, vol. 21, no. 4, p.383-401

WCED (WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT), 1987, 'Tokyo Declaration.' Our Common Future. Oxford University Press.

WYATT, P., RALPHS, M., 2003, GIS in Land and Property Management, Spon Press, London.

WYATT, P., 1995, Using a Geographical Information System for Property Valuation, Journal of Property Valuation and Investment, Vol. 14, No.1, p. 67-79.

Biography

Katerina Christopoulou is a PhD student in the department of Geomatic Engineering at UCL, where she researches spatial data mining and in particular its application to property valuation. She holds an MSc in GIS from the University of Leeds and BSc in Topographic Science from Technological Educational Institute of Athens (T.E.I). She has worked as a GIS and Digital Cartography Engineer in an engineering firm in Greece. Since 2000 she is working at the Hellenic Military Geographic Service in the Remote Sensing department.