Geostatistical Approaches for Geovisual Data Exploration, Analysis and 3D-Visualisation in Civil Security

Markus Wolff and Julia Gonschorek

Abstract— This contribution presents selected approaches, methods and tools to facilitate geovisual analytical data exploration for civil security purposes. To analyse large emergency service data of a major German city's fire department, different data mining techniques are applied. This allows identifying statistical significant clusters in space and time. To facilitate convenient methods for exploring such complex datasets, a GIS-based software is developed. For visualisation and interactive exploration these results are integrated into a three-dimensional geovirtual environment.

Index Terms—GIS, geospatial data analysis, geovirtual environments.

Introduction

With the spreading of geospatial data into daily life-routines this particular type of data becomes somewhat ubiquitous (e.g. mobile navigation, Google Earth, Google Maps). With this increase in data availability demands for complex analysis increase as well. For a wide range of purposes highly dimensional data needs to be analysed rapidly to discover relationships, clusters and trends. To allow for such analysis, complex data can be transformed into "visual displays which allow an analyst to look for patterns, trends, relationships and structures that describe the significant aspects and characteristics of the data" [1] cited in [2]. To efficiently discover and visualise such spatial data patterns methods from computer science and knowledge discovery should be combined with those from cartography and geovisualisation [3]. This leads to the disciplines of Visual Analytics, the science of "analytical reasoning facilitated by interactive visual interfaces" that should lead to "detect the expected and discover the unexpected" [4], cited in [5] and Geovisual Analytics (defined as "the science of analytical reasoning and decision-making with geospatial information, facilitated by interactive visual interfaces, computational methods, and knowledge construction, presentation, and management strategies [6] cited in [2]). According to these definitions this contribution develops methods, techniques and software tools to visually explore large databases by using methods of data mining and geostatistics.

1 Methodology

The underlying database consists of more than 100.000 geocoded features. Each entry represents one particular case of emergency. In total more than 35 different cases are listened in the database, ranging from fire emergencies to surgery- and internistic-related emergencies respectively. The earliest entry is dated on early July, 1st 2007, the latest on late June 30th 2008. All cases are related to the German city of Cologne. As Figure 1 shows, a simple dot-based visualisation of, e.g., internistic-related emergencies, communicates just a very rough pattern representing the overall urban framework.

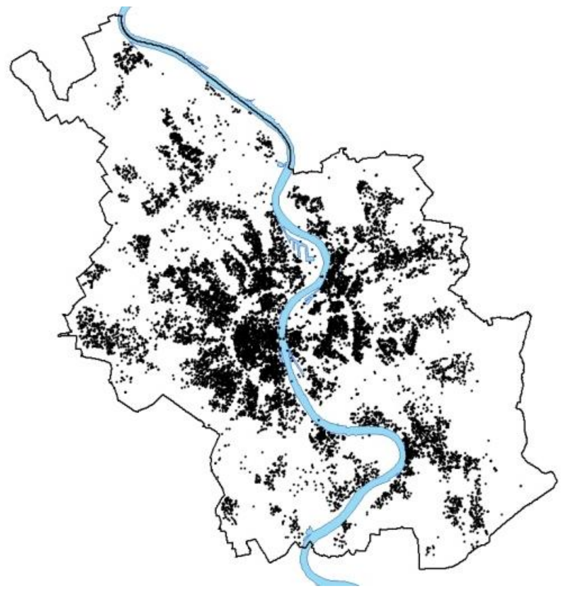

Fig. 1: Dot-based visualisation of internistic-related emergencies in the city of Cologne

To support decision makers with relevant information and to facilitate knowledge discovery, efficient methods and tools have to be designed that fit the needs of the potential user [7]. According to Keim's adjustment of Shneiderman's mantra "analyse first, show the important, zoom, filter and analyse further" [8], these methods should extract the most important spatio-temporal characteristics from a given emergency database and support the user with an easy-to-use environment. Therefore, the presented work targets on the development of techniques to analyse this massive amount of geodata by using efficient methods of knowledge extraction. A set of geostatistical analyses is conducted exemplarily for the cases of internistic-related emergencies. The results are visualised by using both traditional 2D techniques and complex 3D visualisation strategies. Eventually, a GIS-tool is programmed that allows even a GIS-untrained analyst to conduct geospatial knowledge discovery and visualisation with little effort.

To extract geospatial clusters of emergency services, kernel-density-estimations [9], are applied in a first instance. This method reveals first clusters. However, the results depend heavily on cell size / bandwidth parameterisation [10]. To verify these results, further methods are elaborated. Nearest neighbour analysis (NNA) is conducted to examine the distances between each emergency location and the closest emergency to it. On a global level, NNA compares the observed distributions of emergencies with a randomized distribution. This analysis confirms the overall cluster-

- *Markus Wolff is with University of Potsdam, Department of Geography, 3D Geoinformation Research Group, E-Mail: Markus.Wolff@uni.potsdam.de.*
- *Julia Gonschorek is with University of Potsdam, Department of Geography, E-Mail: julia.gonschorek@uni-potsdam.de.*

Manuscript received 01 November 2010;

tendency of this dataset. To proof whether this global observations apply on the local data level as well, Getis Ord Gi* statistics are calculated. The results of this analysis are shown in Figure 2.

Fig. 2: Hot- and cold spots according Getis Org Gi* statistics

Particular highly significant hotspots of internistic-related emergencies exist in the city centre of Cologne, in the north-west and in the south-east (Airport Cologne/Bonn). Furthermore, a significant cold spot (many emergencies with low incidents per location) exists southwesterly of the city centre. K-Means analysis verifies this spots as well. As Figure 3 shows, the city centre, the north-west and the Airport region as well as nine other areas are detected as clusters.

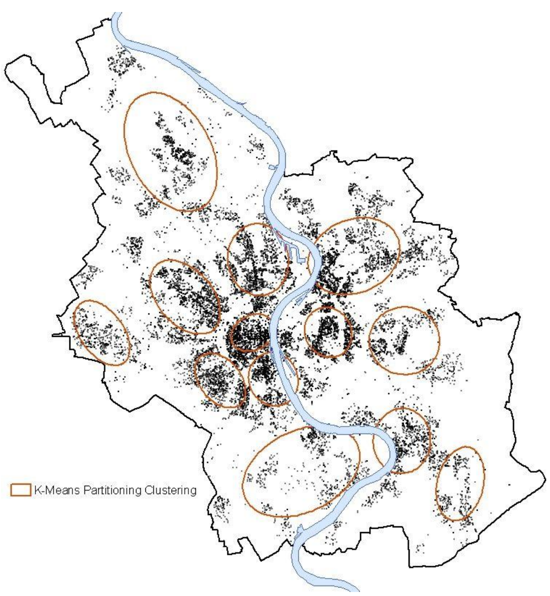

Fig. 3: Results of K-means clustering

At this point of investigation it is to conclude, that the data contains particular, highly significant clusters of internistic-related emergency services. This can be proved by different geostatistical methods. However, no temporal trends are revealed so far. The question is therefore, whether these spatial clusters are related to temporal trends and clusters as well. To detect such temporal trends and clusters, the chi-square goodness-of-fit test is applied. This reveals certain temporal trends. Concerning the significant level (at $p<0.001$ level) the hypothesis of equal distribution can be refused for certain monthly- and daily-based periods. Figure 4 shows, e.g., a maximum of internistic-related emergency services during December 2007. The daily minimum of these emergencies can be identified between 03 and 06 o'clock AM.

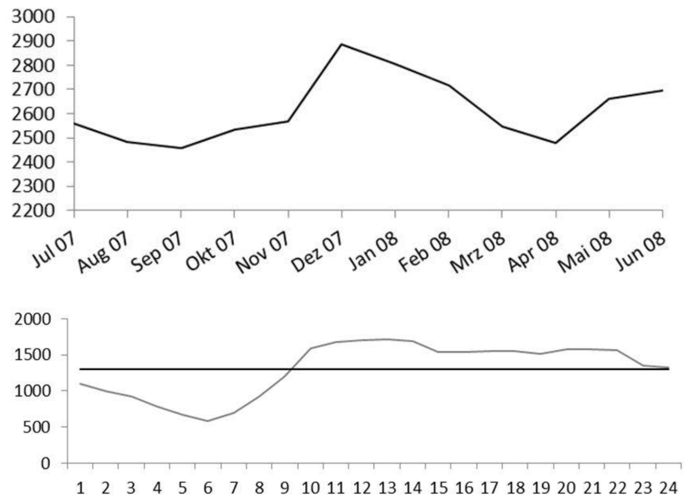

Fig. 4: Temporal trends of internistic-related emergency services: monthly (upper image) and daily (lower image), respectively.

2 Software development

To apply the used methods automatically and for different kinds of emergency services, a GIS-based graphical user interface (GUI) is developed. This GUI is programmed by using the .NET framework and ERSI's ArcObject / ArcEngine technology. By using the ArcEngine software development framework this application is developed as a standalone application (.exe). ArcGIS Engine is applied here because this SDK allows for integrating (Arc-) GIS functions into custom applications. However, since proprietary ArcGIS functionality is addressed, the end user needs the ESRI ArcGIS Engine runtime library or a licensed ArcGIS to be installed at the local machine. The tool allows at first for select a specific emergency service (e.g. fire incidents or surgery services) from the database. Subsequently, statistics are calculated and the results visualised by using different types of charts and maps, respectively. Applying this tool, analysts are able to analyse and visualise spatio-temporal trends and clusters of massive data without the need of professional GIS training or education.

3 3D-Geovisualisation

To facilitate a graphic visualisation, some of the results are further processed and pipelined from the GIS analysis back-end to a sophisticated 3D-visualisation front-end. Integrated into a geovirtual environment, these visualisations are suitable for an intuitive communication of complex geospatial phenomena [11]. Figure 5 shows, for instance, the integration of KDE and NNH ellipses into an interactive 3D environment.

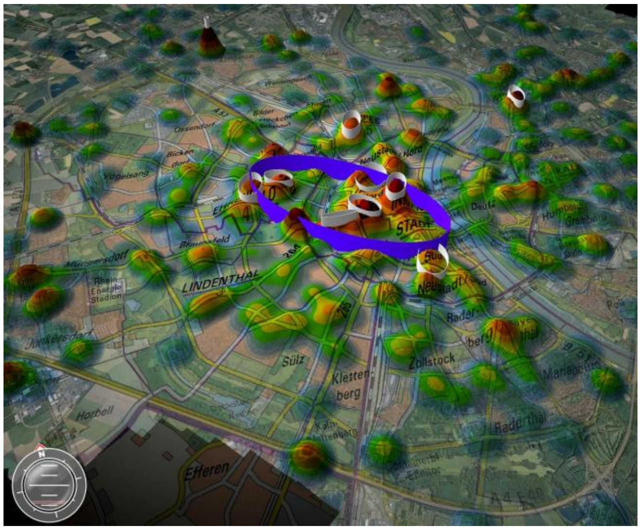

Fig. 5: Integration of KDE and NNH ellipses into an interactive geovirtual 3D-environment

4 Conclusion

This contribution presents geostatistical methods for exploring spatio-temporal explore large databases. It is shown how specific emergency services cluster in space and time. For an intuitive communication of such complex analysis, results are passed to a 3D-visualisation system. To support geovisual data exploration and analysis a GIS-based application is programmed that allows an analyst to conduct major analyses without the being in need of a particular GIS or VIS.

Acknowledgments

Funding of this study by the German Federal Ministry of Education and Research (BMBF) within the framework of the InnoProfile research group ‘3D Geoinformation’ (www.3dgi.de) is gratefully acknowledged. The author also likes to thank the fire department of the city of Cologne for providing extensive emergency datasets. Furthermore the author gives a special thank to the municipality of the city of Cologne for providing the 3D city model and other geospatial base data. Finally, the author thanks Autodesk, Inc. for supplying the LandXplorer system.

References

[1] Keim, D.A. and Ward, M. (2003). Visualisation. In: Intelligent Data Analysis. Berthold, M. and Hand, D. J. (eds.), 2nd edition, Springer Verlag, Berlin–Heidelberg, pp. 403–428

[2] Demsar, U. and Fotheringham, S.A. and Charlton, M. (2008). Combining geovisual analytics with spatial statistics: the example of geographically weighted regression. In: The Cartographic Journal, 45 (3), Maney Publishing, pp. 182-192

[3] Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M.J., MacEachren, A. and Wrobel, S. (2007). Geovisual Analytics for Spatial Decision Support: Setting the Research Agenda. In: International Journal of Geographical Information Science, 21 (8), Taylor & Francis, pp. 839-857

[4] Thomas, J.J. and A.C. Kristin (2005). Illuminating the Path: The Research and Development Agenda for Visual Analytics”. National Visualization and Analytics Center, Piscataway, NJ: IEEE Computer Society Press

[5] Kraak, M.J. (2007). Geovisualization and Visual Analytics. In: Cartographica: The International Journal for Geographic Information and Geovisualization. 42 (2), UT Press, pp. 115-116

[6] MacEachren, A. M. (2008). Invited paper in preparation on Geovisual Analytics (title to be decided), Information Visualisation

[7] Roth, R. E., Ross, K. S., Finch, B. G., Luo, W. A. and M. MacEachren A user-centered approach for designing and developing spatiotemporal crime analysis tools

[8] Keim, D., Andrienko, G., Fekete, J.D., Görg, C., Kohlhammer, J. and G. Melancon. (2008). Visual analytics: Definition, process, and challenges. In: Information Visualization, Springer, pp. 154-175

[9] McCullagh, M. J. (2006). Detecting Hotspots in Time and Space. ISG06

[10] Smith, de M.J., Goodchild, M. F. and P.A. Longley. (2006). Geospatial Analysis, Troubador Publishing

[11] Wolff, M. and H. Asche. (2010). Towards 3D Tactical Intelligence Assessments for Crime Scene Analysis. In Taniar, D., Gervasi, O., Murgante, B., Pardede, E., Apduhan, B.O. (eds.), Computational Science and Its Applications - ICCSA 2010. Fukuoka/Japan, March 23-26 2010. Springer Lecture Notes in Computer Science (LNCS), Volume 6016, Proceedings, Part I. Berlin. pp. 346-360.