Geospatial Data Confidentiality: Collection, Sharing, and Use
Special Session Track, Tuesday, April 9, 2013
Annual Meeting of the Association of American Geographers (AAG)
Los Angeles, California

In recent years, advances in GIS, GPS and related geospatial technologies have led to the collection and use of enormous amounts of more precise and accurate georeferenced data that are increasingly rich in attributes and available, from street network to land parcel data, from crime waves to disease outbreaks, and more. In particular, the collection of continuous space-time coordinates of people's daily lives using GPS has become more common in recent years, and the growing use of GPS-enabled mobile phones has transformed locationally-specific data collection. Locational data about people's activities and trips are also collected in many social surveys and high-resolution, remotely-sensed imagery has become readily available (imagery can be combined with population data, along with locational information related to land use, ownership and household or individual characteristics (Liverman et al. 1998)). While spatial analysis and mapping of georeferenced individual-level data can help identify important geographic patterns or lead to significant knowledge for dealing with specific social issues (Thomas et al. 2008, Kwan 2004, Kwan 2000), such locational data can be sensitive and carry considerable disclosure risk.

The five sessions below have been organized into a special track at the 2013 Annual Meeting of the Association of American Geographers in relation to an NSF-funded research project on geospatial data confidentiality and data sharing (see Session 5). The purposes of the track are to: 1) explore new developments in the use of GPS-enabled mobile phones applied to geographic research, 2) present new approaches and methods for confidentiality protection in geospatial studies, and 3) discuss the unique confidentiality characteristics of geospatial data and their visualizations and implicates for research on disclosure risks and on the potential for sharing geospatial data.

Kwan, M.-P. 2000. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transportation Research* C 8:185-203.

Kwan, M.-P. 2004. GIS methods in time-geographic research: Geocomputation and geovisualization of human activity patterns. *Geografiska Annaler* B 86(4):267-280.

Liverman, D., Moran, E.F., Rindfuss, R.R., and P.C. Stern, eds. 1998. *People and Pixels: Linking Remote Sensing and Social Science.* Washington, DC: National Academy Press.

Thomas, Y, Richardson, D.B., and I. Cheung (eds.). 2008. *Geography and Drug Addiction.* Dordrecht, The Netherlands: Springer.

SESSIONS

1. Geography and Mobile Phone Data: from theory to empirics I (paper session)

8:00 AM - 9:40 AM in Santa Monica C, Westin, Level 3

Organizer(s): Emmanouil Tranos, Vrije University and Rein Ahas, Department of Geography, University of Tartu

Presenter: Rein Ahas, Department of Geography, University of Tartu
Measuring movement patterns of incoming tourists in destination with passive mobile positioning data

Presenter: Yihong Yuan, Institute of Cartography and Geoinformation, ETH Zurich
A framework for characterizing human mobility from georeferenced mobile phone data

Presenter: Siiri Silm, University of Tartu
Measuring ethnic composition by passive mobile positioning data in case of Estonians and Russians in Tallinn

Presenter: Louis Gutierrez, Rensselaer Polytechnic Institute
Solazo Weather App (SWAP): Approximating the Weather

2. Geography and Mobile Phone Data: from theory to empirics II (paper session)

10:00 AM - 11:40 AM in Santa Monica C, Westin, Level 3

Organizer(s): Emmanouil Tranos, Vrije University and Rein Ahas, Department of Geography, University of Tartu

Presenter: Emmanouil Tranos, Department of Spatial Economics, Vrije University
Mobile phone usage and motorway traffic: a simultaneous equation approach

Presenter: Benjamin Hopkins, California State University, Chico
Are we there yet? How ubiquitous use of mobile devices alters our ability to navigate

Presenter: Kati Nilbe, University of Tartu
Factors affecting visitors travel distance: a comparison of foreign event visitors´ and regular visitors´ in Estonia

Presenter: Ruth Lane, Monash University
The Second Life of Mobile Phones

Discussant: Harvey Miller, University of Utah

Research based on mobile phone data is becoming more and more common in geography, planning and social sciences in general as data collection from mobile networks is cost-effective compared to traditional methods due to the immense penetration rate. Numerous applications can be found in the literature: from signature identification in human mobility in the frame of complexity science to transport management applications. Geography and spatial sciences seem to be highly benefit from this development as data from mobile phone operators, despite the privacy issues, can be seen as an opportunity to model the spatio-temporal dynamics of human behavior.

*These three sessions are also included in the AAG Annual Meeting's
Symposium on Geography, GIScience, and Health: Spatial Frontiers of Health Research and Practice
(http://www.aag.org/AM2013/GIS-Frontiers)

3. Encryption for confidentiality protection in geospatial studies with human subjects (paper session)
12:40 PM - 2:20 PM in Santa Monica C, Westin, Level 3
Organizer(s): Geoffrey Jacquez, State University of New York at Buffalo and Khaled El Emam, Electronic Health Information Laboratory, CHEO Research Institute

Presenter: Geoffrey Jacquez, State University of New York at Buffalo
Exploratory Evaluation of Homomorphic Cryptography for Confidentiality Protection: Study Motivation, Objectives and Design
Presenter: Daniel Goldberg, Texas A&M University
Overview of Approaches to Confidentiality Protection in Geospatial Studies with Human Subjects
Presenter: Khaled El Emam, Electronic Health Information Laboratory, CHEO Research Institute
An overview of secure computation methods for sharing and analyzing personal information and data on small areas
Presenter: Aleksander Essex, CHEO Research Institute
Invisible Pushpins: Techniques for Computing Geospacial Statistics on Encrypted Data
Presenter: Andrew J Curtis, Kent State University
Can I see your house from here? Reengineering risk in health maps

Confidentiality protection is an important priority in studies involving human subjects, but can restrict data access thereby limiting result validation and slowing the pace of basic, applied and translational research. In geospatial studies, techniques for protecting privacy include masking of results, aggregation and "jittering" patient locations through the addition of spatial errors to the coordinates of place of residence. This session reports the status of a research project, in progress, funded by the National Library of Medicine "Exploratory evaluation of homomorphic cryptography for confidentiality protection".

4. Geography and Mobile Phone Data: is there a privacy caveat? (panel session)
2:40 PM - 4:20 PM in Santa Monica C, Westin, Level 3
Organizer(s): Rein Ahas, Department of Geography, University of Tartu

Panelist: Rein Ahas, Department of Geography, University of Tartu
Panelist: Michael Batty, University College London, Centre for Advanced Spatial Analysis
Panelist: Erki Saluveer, University of Tartu

The use of mobile phone data for scientific research has become increasingly popular among geographers in recent years. Geo-referenced, individual and aggregated data from mobile phone operators enables analysts to uncover fine-grained spatiotemporal dynamics of socio-spatial phenomena that would have not been identified if traditional spatial statistics had been employed. Despite the undoubted potential of mobile phone data for deepening our understanding of critical social issues, its use also raises questions about privacy protection and confidentiality, as it may disclose the home, workplace, activities and trips of individuals in a population under scrutiny.

*These three sessions are also included in the AAG Annual Meeting's
Symposium on Geography, GIScience, and Health: Spatial Frontiers of Health Research and Practice
(http://www.aag.org/AM2013/GIS-Frontiers)

This panel session seeks to stimulate discussion and reflection about privacy issues related to the use of mobile phone data in social science. The panel will convene experts from different disciplines and backgrounds who have been concerned with the trade-off between privacy protection and accuracy of analytical results. Each panelist will express her/his views on the progress and prospects in this area, followed by open discussions. The objective is to provide an up-to-date picture on what is happening, outline the challenges and facilitate the dialogue on a wide range of aspects related to (geo)privacy including legislation, policy, anonymity and obfuscation strategies. Our plan is to continue the debate initiated by this panel session in a form of a special issue in a high impact journal.

5. Addressing Challenges For Geospatial Data-Intensive Research Communities: Research on Unique Confidentiality Risks & Geospatial Data Sharing Issues (panel session)

4:40 PM - 6:20 PM in Santa Monica C, Westin, Level 3

Organizer(s): Douglas Richardson, Association of American Geographers

Panelist: Douglas Richardson, Association of American Geographers
Panelist: Kristine Witkowski, Inter-university Consortium for Political and Social Research
Panelist: Mei-Po Kwan, University of Illinois at Urbana-Champaign and Utrecht University
Panelist: Jean McKendry, Association of American Geographers

Research combining a variety of geographically-referenced data streams is spreading across many scientific domains, ranging from environmental science to transportation to epidemiology, and opportunities to create new multi-disciplinary and data-intensive scientific collaborations are expanding. Yet, the unique characteristics of georeferenced data present special challenges to such collaborations. These data are often highly identifiable when presented in maps and other visualizations. The potential opportunities and benefits of collaboration are constrained by the need to protect the locational privacy and confidentiality of subjects in research using georeferenced data.

The Association of American Geographers (AAG), together with the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan, is leading an NSF-funded program of research to address these challenges facing geospatial data-intensive research communities, by focusing on the unique confidentiality characteristics of geospatial data and their visualizations, on disclosure risks, and on the potential for sharing geospatial data within a Virtual Data Enclave (VDE). The purposes of this panel are to: 1) provide a capstone to the previous sessions in this track, 2) share information about this research project, and 3) to engage the broader geography and geospatial research community in a discussion on the challenges and opportunities associated with geospatial data-intensive research and data sharing. This research project is funded under National Science Foundation Award BCS-1244691.