

Geographic Privacy-aware Knowledge Discovery and Delivery

Fosca Giannotti
CNR Pisa
Pisa, Italy
fosca.giannotti@isti.cnr.it

Dino Pedreschi
University of Pisa
Pisa, Italy
pedre@di.unipi.it

Yannis Theodoridis
University of Piraeus
Piraeus, Greece
ytheod@unipi.gr

ABSTRACT

A flood of data pertinent to moving objects is available today, and will be more in the near future, particularly due to the automated collection of privacy-sensitive telecom data from mobile phones and other location-aware devices. Such wealth of data, referenced both in space and time, may enable novel classes of applications of high societal and economic impact, provided that the discovery of consumable and concise knowledge out of these raw data is made possible. Recent research activities have developed theory, techniques and systems for geographic knowledge discovery and delivery, some of them based on privacy-preserving methods for extracting knowledge from large amounts of raw data referenced in space and time. All these efforts aim at devising knowledge discovery and analysis methods for trajectories of moving objects.

The fundamental hypothesis is that it is possible, in principle, to aid citizens in their mobile activities by analysing the traces of their past activities by means of data mining techniques. For instance, behavioural patterns derived

from mobile trajectories may allow inducing traffic flow information, capable to help people travel efficiently, to help public administrations in traffic-related decision making for sustainable mobility and security management, as well as to help mobile operators in optimising bandwidth and power allocation on the network. On the other hand, it is clear that the use of personal sensitive data arouses concerns about citizen's privacy rights.

In this tutorial, we establish a framework for the challenges and the mining solutions for the geographic information collected by Moving Object Database (MOD) engines. We first discuss the challenges of collecting mobility data, and elaborate on the impact of trajectory data analysis in several modern applications. We then discuss methodologies and techniques to collect raw data, reconstruct trajectory information, and efficiently store it in MODs. We continue with an overview of knowledge discovery approaches for movement data. Finally, we propose a research agenda and identify areas where interdisciplinary studies are needed.

ABOUT THE AUTHORS

Fosca Giannotti (CNR Pisa, Italy).

Fosca Giannotti is a senior researcher at the Information Science and Technology Institute of the National Research Council at Pisa, Italy. Her research interests include data mining query languages, knowledge discovery support environment, web-mining, spatio-temporal reasoning, spatio-temporal data mining, and privacy preserving data mining. She has been the coordinator of various European and national research projects and is currently the co-ordinator of the FP6-IST project GeoPKDD: Geographic Privacy-aware Knowledge Discovery and Delivery. She is the author of more than one hundred publications and served in



the scientific committee of various conferences in the area of Logic Programming, Databases, and Data Mining. In 2004 she co-chaired the European conference on Machine Learning and Knowledge Discovery in Data Bases ECML/PKDD 2004. She is the program co-chair of ICDM 2008, the IEEE Int. Conf. on Data Mining.

Dino Pedreschi (University of Pisa, Italy).

Dino Pedreschi is a full professor at the Dipartimento di Informatica of the University of Pisa. He has been a visiting scientist and professor at the University of Texas at Austin (1989/90), at CWI Amsterdam (1993) and at UCLA (1995). His research interests are in data mining and logic in databases, and particularly in data analysis, in spatio-temporal data mining, and in privacy-preserving data mining. He is a member of the program committee of the main international conferences on data mining and knowledge discovery, and an associate editor of KAIS.



He has been a co-chair of ECML/PKDD 2004, the international conference on Machine Learning and Knowl-

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the ACM. To copy otherwise, or to republish, to post on servers or to redistribute to lists, requires a fee and/or special permissions from the publisher, ACM.
EDBT 2009, March 24–26, 2009, Saint Petersburg, Russia.
Copyright 2009 ACM 978-1-60558-422-5/09/0003 ...\$5.00

edge Discovery in Databases. He served as the coordinator of the undergraduate studies in Computer Science at the University of Pisa, and as a vice-rector of the same university, with responsibility in teaching affairs. He is the national coordinator of the project “GeoPKDD.it”, Geographic Privacy-aware Knowledge Discovery and Delivery.

Yannis Theodoridis (University of Piraeus, Greece). Yannis Theodoridis is Assoc. Professor with the Department of Informatics, University of Piraeus in Greece. He received his Diploma and PhD degrees in computer engineering from the National Technical University of Athens in 1990 and 1996, respectively. Since 2002, he is in the fac-



ulty of Informatics, University of Piraeus in Greece, currently directing the Information Systems Laboratory (InfoLab). His research interests include knowledge discovery and data mining (KDD), spatial/spatiotemporal databases and geographical information systems (GIS). He was scientist in charge and coordinator of two European FP6 projects, namely PANDA (IST-2001-33058 FET Working Group; 2001–04) and CODMINE (IST-2001-39151 FET Assessment Project; 2002–03), on pattern-based management and privacy-preserving data mining, whereas he is currently involved in GeoPKDD Project (IST STREP Project, 2005–05) on moving objects trajectory warehousing and mining aspects. Yannis Theodoridis has presented tutorials on spatial databases (ADBIS’99) and mobile location solutions (ADBIS’03).