The Prisma System

Intelligent Agents Working on Crime Pattern Analysis supported by Geographic Information Systems

Maribel SANTOS, Orlando BELO and Luís AMARAL

The process of extracting useful knowledge from large databases became one of the tasks of prior importance in today's organizations. The collection of excessive amount of information makes very difficult its treatment and analysis without appropriated means. Police Departments are real examples of organizations that currently debate themselves with situations involving large volumes of distributed information and requiring effective real time decision making. Some of these situations are critical in the normal Police Department's activities, namely the ones related to Crime Pattern Analysis. These are concerned with the recognition of spatial and temporal regularities in reported crime and the ability of predict future criminal activity. This is very important due the possibility to provide effective elements to increase patrol actions, improve priority investigations or even perform better public notification. Through the combination of Multi-Agent Systems and Geographic Information Systems technologies we design a computational system Intelligent Crime Pattern Analysis: the Prisma system. It considers a community of intelligent agents, divided essentially into two classes, that will be responsible respectively to populated specialized Data Marts and make Criminal Patterns Identification. With Prisma, Police Departments will be able to examine patterns related to notified incidents and analyze their movement in relation to police initiatives.

KEYWORDS: Crime Pattern Analysis, Geographic Information Systems, Intelligent Agents, Data Marts, Data Mining.

INTRODUCTION

It is not necessary to do an extensive work to verify that the complexity of today's organization problems and computer-based systems grows day after day, in some cases almost exponentially. The demand of new working methods, the necessity to deal with large volumes of information, the development of adequate procedures to the analysis and generation of specialized knowledge are only some determinant factors that affect significantly the organization's strategies definition and the behavior of their members.

As greater is the organization's dynamics and its needs of information gathering, more frequently are the mutations on its cognitive patrimony, more critical are the decision making tasks, and more necessary is to have efficient and flexible mechanisms of control. The most critical situations occur normally due to a deficient combination of expertise and knowledge by organization's members that are preponderant on decision making processes. Frequently, the lack of political will, the weak aptness to the development of coordinate tasks, the distribution of resources, or the unavailability of the organization's experts are serious cases of contention in the efficacy of the organization. In order to reduce the influence of those factors, the integration of high specialized teams in the most critical organization's application domains combined with the availability of large volumes of information organized accordingly the organization's needs, constitutes an attractive and positive approach.

Nowadays, *Police Departments* (PD) are real examples of organizations that debate themselves with situations involving large volumes of distributed information and requiring effective real time decision making. Some of these situations are critical in the police personnel performance and PD success, namely the ones related to *Crime Analysis* (CA). CA is defined as the systematic process of collecting, categorizing, analyzing, and disseminating timely, accurate, and useful information that describes crime patterns, crime trends and potential suspects [8].

This paper presents a general model for criminal pattern information gathering and analysis and describes the design of an intelligent crime pattern analysis system, which model was study and develops in order to support and improve PD's activities on crime prevention and investigation. Through the combination of *Multi-Agent Systems* (MAS) and *Geographic Information Systems* (GIS) technologies we develop a computational system, Prisma, that considers a community of intelligent knowledge based agents that are able to act together in order to reach criminal patterns.

The work presented in this paper has been supported by the District Police Department of Braga, in the north of

Portugal, and developed based on its resources (data and know how). The Prisma system will be very useful in the referred PD on day by day activities related to the efficient management of data resources that it must maintain, analyzed and keep up-to-date.

CRIMINAL RESEARCH

CA is a key step in the sequence activities aimed at conceiving, implementing and evaluating measures to prevent crime. This is an exploratory process that goes through several stages (Figure 1).

The process includes the collection of large amounts of detailed information, the analysis of the collected information to identify crime patterns, the formulation of preventive strategies, the implementation of the strategies and the evaluation of their impact on crime.

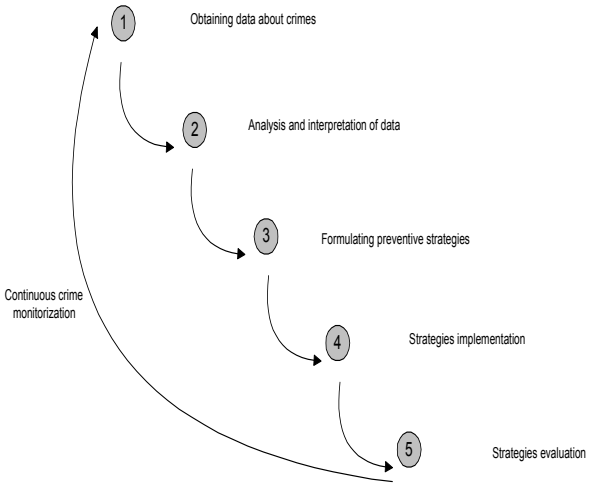


Figure 1 – The Stages of a Preventive Process

Essentially, CA activities may be organized in three main classes:

- Tactical, that provides information to assist operational personnel (patrol and investigate officers) in the identification of specific and immediate crime problems and the arrest of criminal offenders. Data analysis is used to promote a quick response to field situations.
- Strategic, which is concerned with long-range problems and projections of long-term increases or decreases in crime (crime trends). Strategic analysis also includes the preparation of crime statistical summaries, resource acquisition and allocation studies.
- Administrative, that focuses on provision of economic, geographic or social information to administration.

There are several forms of CA, which are relevant to the treatment of high volume crime: crime patterns, crime trends, crime clusters and crime series [17]. Crime patterns look for the nature and distribution of crime within a specific area; crime trends are related to significant changes in an area crime pattern over time; crime clusters find groups of crimes linked through similar characteristics; and crime series looks for crimes with common offenders.

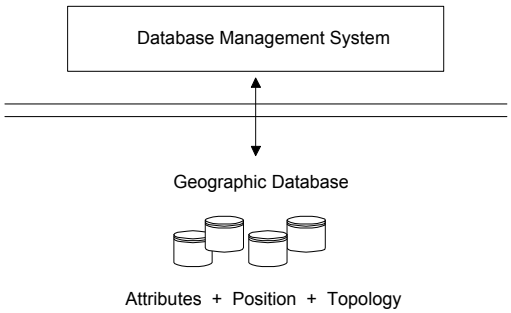
Crime Pattern Analysis (CPA) is concerned with the recognition of spatial and temporal regularities in reported crime, and the ability of predict future criminal activity, given past reality [14].

Traditionally, officers collect and catalogue information about crimes, their location and time. Additional information is collected as the *modus operandi* of the crime and a description of its general characteristics. CPA is based on the assumption that individual crimes and incidents are not unique events but share a number of common characteristics or features [11].

Pattern detection is a process of instantiating and evaluating patterns to fit particular set of data. Patterns extraction methods can be selectively invoked based on application tests. An application test is a filter which determines whether a given type of pattern can be detected from given data [12]. A pattern describes any or two more crimes that exhibit similar characteristics such as location, time of day, day of week, weapon used, suspect description, victim type, property taken and so on. A pattern may suggest that a single suspect or group of suspects is committing a number of crimes, or it may simply indicate a general trend. Crime patterns are identifiable in the spatial distribution of offenses, in the spatial and social distribution of victims of crime and in the demographic characteristics of perpetrators or offenders. The identification of a pattern may result in increased patrol, priority investigation, public notification, or more than one of these.

INTEGRATING GEOGRAPHIC INFORMATION SYSTEMS TECHNOLOGY

The information collected must be organized and stored



accordingly their nature and remains accessible to PD's purposes on CPA. Given the strong dependency that criminal data has with the places where the crimes occur, it is obvious that must exist adequate and specialized means that ensure such geographic data relationship and provide information retrieval facilities. Geographic data have essentially four major components: the geographic position, the attributes, the spatial relationship and time [1]. Geographic information allows describing a phenomenon at a location it existed at a specific point in time.

Figure 2 – The Basic Components of a Geographic Database

Usually, GIS are able to deal with such kind of geographic data. In their structure they includes tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world [6]. Functional GIS

comprise an integrated collection of computer hardware, computer software and geographical data [13]. The entities of a GIS are represented by their attributes, and most important, by their geographic position and spatial configuration (Figure 2). The spatial configuration describes the entities spatial interrelations with each other (topological relations).

GIS have enormous potentialities in criminal research because of its three main abilities: database management, spatial analysis and geographic visualization. With these three components, GIS provide the basic capabilities of linking crimes and their characteristics with their geographic locations. These systems also provide overlay functions that allows the association of the locations of crimes with the demographic characteristics of the areas in which they occur.

Current GIS allow users to search for patterns in space and time, using graphic and/or database tools. In such systems the user is responsible for specify the criteria of that search. Therefore, GIS are able to be hypothesis-testing devices but are unable to generate suitable hypothesis themselves [14]. Despite widespread recognition that the analysis of patterns and relationship in geographical data should be a central function of GIS, the sophistication of certain areas of analytical functionality in many existing GIS continues to be not enough. One difficulty experienced in any discussion of links between GIS and spatial analysis is the clarification of exactly what is to be considered as spatial analysis. In the context of this work, is assumed that spatial analysis is a general ability to manipulate spatial data into different forms and extract additional meaning as a result [2].

APPLYING MULTI-AGENT SYSTEMS TECHNOLOGY

Over the past few years, *Distributed Artificial Intelligence* (DAI) [5] has reinforced significantly its impact in terms of new problem solving concepts and methodologies in a wide group of application areas ranging from Robotics, Knowledge Based Systems, Artificial Life, Virtual Reality, Medical Diagnostic Systems, to Geographic Information Systems. However, in nowadays, an area has lead the main research and applicative efforts of DAI to real world problems: the MAS [20] [19] one.

The MAS area is concerning with the development of new theories, languages, architectures, and tools [15] involving artificial entities designed as agents [10] [7]. Agents are frequently referred as computational entities with specific and well-defined application domains, that are able to act as kind of autonomous artificial assistants, independent, and with capacities of decision over their environment. The application of MAS technology in the implementation of CPA systems can be very helpful and advantageous in many aspects concerning with intelligent information retrieval, inter-disciplinary approaches to criminal pattern identification and selection, and distributed cooperation on CPA.

The application specter of MAS technology has been very wide, contributing significantly to the design and development of distributed computational environments oriented to the integration of cooperative Knowledge Based System [18] and to support cooperation among intelligent agents [3] [16]. In MAS, agents are organized by competence areas and distributed according the applications' needs. The execution of a task or a problem solving process is normally a cooperative process involving agents that have the expertise and knowledge required for do it.

In CA we may adopt all these ideas in the design of a MAS for CPA. Thus, we can think in a system that integrates a community of intelligent knowledge based agents that are able to act together in order to reach criminal patterns through the combination of their expertise and criminal knowledge. If we connect all the agents through a dedicated and robust communication medium, we provide the means to ensure inter-agent communication and cooperation. We can allocate some of the agents doing information gathering on geographic information databases in order to create and maintain specialized Data Marts, organized according to the needs of each PD. Another group of agents can be allocated in criminal pattern identification working over their own knowledge bases and the information provided by the former agents through their Data Marts. All together can act as a real "task-force"' on crime prevention and investigation.

INTELLIGENT CRIME PATTERN ANALYSIS

The analysis of a geographic database by a GIS requires the interaction of one or more users, which are responsible for the specification of the search criteria. Techniques of knowledge discovery in databases [9] and the use of intelligent agents technology are very important aspects in the success of such task.

The implementation of the Prisma system has as first aim the construction of a system for knowledge discovery in geographic databases, where the participation required from users be it extremely reduced. This system is supported by GIS, which allow efficient mechanism for the manipulation and visualization of geographic information. The intelligent agents manage all the process, using for that the knowledge about the application domain transmitted by the users.

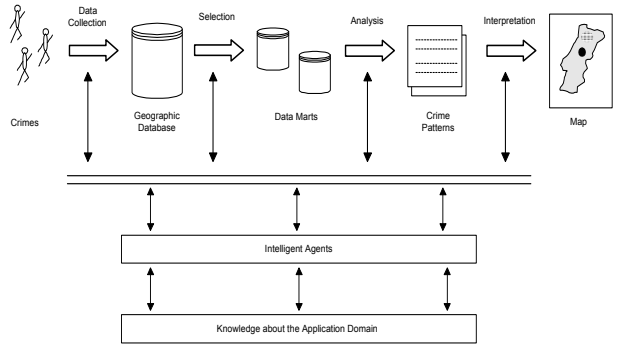


Figure 3 – General Model for Criminal Pattern Information Gathering and Analysis

The model developed for CPA based on intelligent agents (Figure 3), follows the next steps:

- Data collection, the data regarding the locations and characteristics of all crimes must be collected by the officers and made available for further investigations. The information that constitutes the core of the analysis can be summarized in the following questions: what kind of offenses occurred; where, when, under that circumstance and by what method was they committed; and who or what was the victim or target? The answer to those questions can contribute to understand better the crime pattern, the criminal opportunities that underlie it, and the scope for prevention. This step involves the transformation of the original information, which is usually in the form of cases (a set of details about an individual criminal activity), into a number of entities and their attributes that constitutes the geographic database facts.
- Data Selection, depending on the objectives of analysis wanted, the data must be selected and prepared for the next step. One or more intelligent agents support this task and are responsible for identify, select and, if needed, transform the data for pattern identification. This step is very important because the selected data should contain much relevant information as available, and if possible, contain no irrelevant information. To a correct Data Marts construction, the data selected must be checked to ensure that missing values and eventual corrupted data are treated. In this step are implemented strategies for removal of noisy information and for handling missing data fields.
- Crime Patterns Identification, based on the specification of what we need to know, agents must be able to apply the search algorithms needed to satisfy the user's objectives. Sometimes we do not know exactly what we look for. In this case, the job of the intelligent agents acquires one more obstacle, not only in the search of patterns, but also in the selection of data being analyzed (previous step).
- Patterns Interpretation and Visualization, the patterns discovered must be transform in order to be understood by the users and whenever wanted integrated in the geographic database. This integration allows the visualization of the results in a map.

This is an iterative process in that is always possible going back to a previous step and there recover the course of CPA.

For the identification of spatial and temporal crime patterns, agents must have tools, algorithms, which allow them to develop efficient pattern searching. These algorithms can be easily specify, but we must have a hypothesis to the solution we are looking for (this hypothesis guides our analysis). If we only want to specify what we need to know, then our agents must

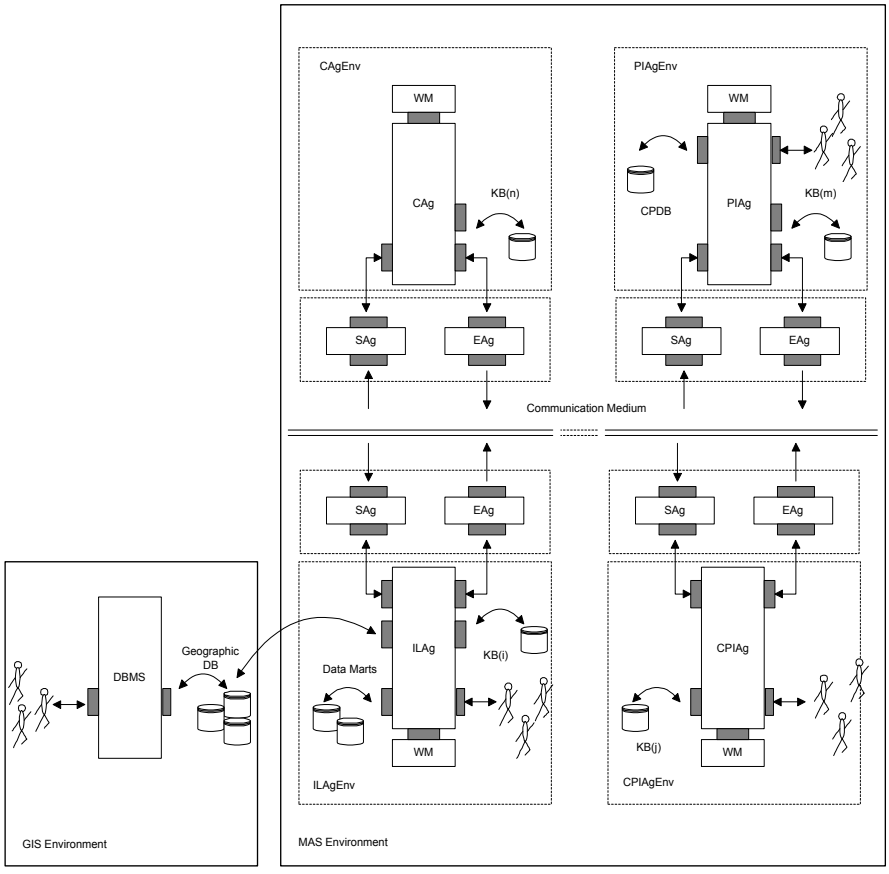


Figure 4 – The Prisma System

implement some of the principles of the Data Mining tools. These algorithms will allow a generic implementation for searching patterns in geographic databases.

THE COMPUTATIONAL MODEL FOR CRIME PATTERN ANALYSIS

The computational model, the Prisma system, to support crime pattern analysis activities (Figure 4) considers basically four classes of agents:

- Information Loaders (ILAg), will collect from the geographic database the information that they need to build the Data Marts, according the directives stored in their own knowledge bases (KB(i)). The Data Marts will be updated periodically being its information available directly to users or to other agents located in the system's environment.
- Criminal Pattern Identifiers (CPIAg), these agents will analyze users' needs, presented through a specialized interface, and try to reach some crime patterns according to: 1) the information presented by the users; 2) the reasoning strategies stored inside their knowledge bases (KB(j)); and 3) the information stored in the Data Marts maintained by the ILAg agents.
- Patterns Interpretation Agents (PIAg), who are responsible for the patterns format detection and the application of the correspondent interpretation mechanism. All the information needed to do it would be stored in their knowledge bases (KB(m)). The pattern interpretation is a very important task since some data mining techniques shows its results in a less understandable form. Additionally, they will maintain the patterns discovered in past processes in a specialized Criminal Patterns Database (CPDB), which keeps accessible for further analysis.
- Communication Agents (CAg), will be responsible to ensure communication among the ILAg, the CPIAg and the PIAg agents. The *Sensors* (SAg) analyze de information that circulates in the *Communication Medium* (CM) according to the instructions of the ILAg, CPIAg or PIAg to which they are associated. The *Effectors* (EAg) post in the CM the knowledge structures sent by their associated agents. They will be independent computational processes in order to avoid contention or overload situations in the internal activities of the ILAg, CPIAg or PIAg agents. Put them as autonomous agents makes possible to manage and schedule effectively the messages received or to be send in the local platforms where the ILAg, CPIAg or PIAg agents are located.

The CM acts as a direct communication channel between agents ensuring message passing among agents. The SAg and EAg will have local mechanisms to interpret the messages ensuring mutual understanding. The possibility for agents having their own language is not put aside. In these cases, to ensure message interpretation and understanding agents must have adequate translation mechanisms. The system will provide the means to ensure agent distribution through PD installations in order to follow the locations of data resources, criminal analysis experts and information retrievers.

Assisted by intelligent agents, PD will be able to examine patterns related to reported incidents and analyze the movement of these incidents in relation to policing initiatives. As a first instance project results, PD will be able to get information related to:

- Crime Trends based on computer analyses of existing and past criminal activity, the system will issue a projection of future crime activity for management decision making.
- Situational Analysis, will provide demographic data on victims and areas experiencing crime activity for proactive assistance through crime prevention efforts and for enforcement planning.
- Crime Patterns/Series, will allow identifying statistically unique events that have a recurring pattern. The analysis will focus on the criminal's *modus operandi*, the crime's day/time/location patterns and clusters, and past similar crimes to perform predictive analyses or crime prevention methods.

CONCLUSIONS AND FUTURE WORK

CA is based on the assumption that crimes are not totally random, isolated or unique events, but can be combined into sets sharing common features and showing distinct patterns. The emergence of GIS has provided PD with the ability to map and manipulate data sets faster and to link crime data with other data in order to understand the crimes occurred.

Despite GIS have great potentialities in criminal research, the enormous volumes of data being analyzed make very difficult the search of patterns without a system designed with this aim. Developing appropriate methods for spatial analysis, and its aggregation with GIS, is extremely important at a time when there is a large growth both in geographic information and in the number of users who are potentially interested in those results.

In this paper was presented a computational model for CPA based on intelligent agents technology and supported by GIS. In this model, agents are distributed and are cooperative entities responsible for the selection, verification and analysis of data, and for the transformation of founded patterns to a format understandable by the users. They must guarantee the correct Data Marts construction and the efficient implementation of the Data Mining algorithms. The GIS will provide the means to manipulate the data to be analyzed and to display the results (patterns) in an appropriate format.

The Prisma system is now under construction. At this moment, the cooperation and coordination strategy of the agents was defined and the data collection and construction of the geographic database takes place. In terms of hardware and software requirements, the system will be developed using SICStus Prolog, Oracle and Java, and implemented in a TCP/IP based network of Unix workstations. The Prisma system will be implemented under the principles established on the BEABLE system [4], a distributed computational environment for the implementation of MAS, developed in the University of Minho.

ACKNOWLEDGMENTS

The authors gratefully acknowledge District Police Department of Braga, especially to the Superintendent Alfredo Farinha Ferreira, for all the support and incentive supplied to the development of this work.

REFERENCES

1. **Aronoff, S.** Geographic Information Systems: a management perspective. WDL Publications, Ottawa, 1989.
2. **Bailey, T. C.** A review of statistical spatial analysis in geographical information systems. in Fotheringham, S. and Rogerson, P. (Eds.), Spatial Analysis and GIS, Taylor & Francis, London, 13-44, 1994.
3. **Belo, O., and Neves, J.** Cooperation Among Opportunistic Agents on a Distributed Environment. in Zhang, C. and Lukose, D. (Eds.), Lectures Notes in Artificial Intelligence (1087). Distributed Artificial Intelligence Architecture and Modelling, Springer-Verlag, 1996.
4. **Belo, O., and Neves, J.** The BEABLE system: a distributed computational environment for multi-agent systems implementation, in Proceedings of the MASTA'97 - Multi-Agents Systems: Theory and Applications, Coimbra, Portugal, 1997,
5. **Bond, A., and Gasser, L.** (Eds.) Readings in Distributed Artificial Intelligence, Morgan Kaufmann Publishers, 1988.
6. **Burrough, P. A.** Principles of Geographical Information Systems for Land Resources Assessment. Oxford University Press, Oxford, 1986.
7. **Castelfranchi, C.** To Be or Not To Be an "Agent". in Muller, J., Wooldridge, M., and Jennings, N. (Eds.), Lectures notes in Artificial Intelligence (1193). Intelligent Agents III. Agent Theories, Architectures, And Languages, Springer-Verlag, Budapest, Hungary, 37-39, 1996.
8. **Ekblom, P.** Getting the best out of Crime Analysis, Home Office, Crime Prevention Unit, Paper 10, 1988.
9. **Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R.** (Eds.) Advances in Knowledge Discovery and Data Mining, The MIT Press, Massachusetts, 1996.
10. **Franklin, S., and Grasser, A.** Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents. in Muller, J., Wooldridge, M., and Jennings, N. (Eds.), Lectures notes in Artificial Intelligence (1193). Intelligent Agents III. Agent Theories, Architectures, And Languages, Springer-Verlag, Budapest, Hungary, 21-35, 1996.
11. **Hirschfield, A., Brown, P., and Todd, P.** GIS and the analysis of spatially-referenced crime data: Experiences in Merseyside, U.K. International Journal of Geographical Information Systems, Vol. 9, Nr. 2, 1995, 191-210.
12. **Klösgen, W., and Zytkow, J. M.** Knowledge Discovery in Databases Terminology. in Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (Eds.), Advances in Knowledge Discovery and Data Mining, The MIT Press, Massachusetts, 1996.
13. **Maguire, D. J.** An overview and definition of GIS. in Maguire, D. J., Goodchild, M. F., and Rhind, D. W. (Eds.), Geographical Information Systems: Principles and Applications, Longman Scientific & Technical, Harlow, 9-20, 1991.
14. **Mitchell, D.** Crime, Policing and GIS: An Emerging Technology?, Publication 5/94, Association for Geographic Information, 1994.
15. **Muller, J.** Control Architectures for Autonomous and Interacting Agents: A Survey. In Cavedon, L., Rao, A., and Wobcke, W. (Eds.), Lectures notes in Artificial Intelligence (1209). Intelligent Agents Systems - Theorical and Practical Issues, Springer-Verlag, Cairns, Australia, 1-26, 1996.
16. **Oliveira, E., and Mouta, F.** A Distibuted AI Architecture Enabling Multi-Agent Cooperation, in Proceedings of the 6th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Edinburgh, Scotland, 1993,
17. **Read, T., and Oldfield, D.** Local Crime Analysis, Home Office, Crime Prevention and Detection, Paper 65, 1995.
18. **Wittig, T.** (Eds.) ARCHON: an architecture for multi-agent systems, Ellis Horwood Limited, 1992.
19. **Wooldridge, M., and Jennings, N.** Intelligent Agents: theory and practice. Knowledge Engineering Review, Vol. Nr. 19, 1995,
20. **Wooldridge, M., and Jennings, N.** (Eds.) Lecture Notes in Artificial Intelligence (890). Intelligent Agents. Proceedings of the ECAI-94 Workshop on Agent Theories, Architectures and Languages, Springer-Verlag, 1995.

Maribel SANTOS
maribel@di.uminho.pt

Maribel Santos is a research assistant at the Department of Informatics, University of Minho, where she's also a PhD student in Informatics.

Among her interests are geographic information systems, data mining, knowledge discovery in databases and intelligent agents.

Maribel is graduated in Systems Engineering by the University of Minho, where she also got a MSc degree in Informatics.

University of Minho
Departamento de Informática
Escola de Engenharia
Campus de Azurém
4800 Guimarães
Portugal
Tel: +351.53 510259
Fax: +351.53 510250
URL: http://www.uminho.pt

Orlando BELO
obelo@di.uminho.pt

Orlando Belo is an auxiliary professor at the Department of Informatics, University of Minho, where he develops research activity in the fields of Knowledge Engineering, Multi-Agent Systems and Data Warehousing.

Orlando is graduated in Systems Engineering by the University of Minho, where he also got a PhD degree in Informatics.

University of Minho
Departamento de Informática
Escola de Engenharia
Largo do Paço
4709 Braga Codex
Portugal
Tel: +351.53 604470
Fax: +351.53 604471
URL: http://www.uminho.pt

Luís AMARAL
amaral@ci.uminho.pt

Luís Amaral is an associate professor at the Department of Informatics, University of Minho, where he's also director of Business Informatics course, executive board of APSI (Information Systems Portuguese Association) and senior editor of Sistemas de Informação Journal.

Among his interests are information systems planing, information systems management and business informatics.

Luís is graduated in Systems Engineering by the University of Minho, where he also got a PhD degree in Informatics.

University of Minho
Departamento de Informática
Escola de Engenharia
Campus de Azurém
4800 Guimarães
Portugal
Tel: +351.53 510259
Fax: +351.53 510250
URL: http://www.uminho.pt