

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Multiagent System for Image Mining

Nicksson Ckayo Arrais de Freitas and
Marcelino Pereira dos Santos Silva

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.69353>

Abstract

The overdone growth, wide availability, and demands for remote sensing databases combined with human limits to analyze such huge datasets lead to a need to investigate tools, techniques, methodologies, and theories capable of assisting humans at extracting knowledge. Image mining arises as a solution to extract implicit knowledge intelligently and semiautomatically or other patterns not explicitly stored in the huge image databases. However, spatial databases are among the ones with the fastest growth due to the volume of spatial information produced many times a day, demanding the investigation of other means for knowledge extraction. Multiagent systems are composed of multiple computing elements known as agents that interact to pursuit their goals. Agents have been used to explore information in the distributed, open, large, and heterogeneous platforms. Agent mining is a potential technology that studies ways of interaction and integration between data mining and agents. This area brought advances to the technologies involved such as theories, methodologies, and solutions to solve relevant issues more precisely, accurately and faster. AgentGeo is evidence of this, a multiagent system of satellite image mining that, promotes advances in the state of the art of agent mining, since it relevant functions to extract knowledge from spatial databases.

Keywords: remote sensing, database, image mining, multiagent system, agent, agent mining, data mining, AgentGeo

1. Introduction

Technological advances have provided new ways to collect spatial data: satellites, radars, unmanned air vehicle, balloons, and many others. These instruments caused an enormous accumulation of images data on remote sensing databases for many reasons. These databases are the ones with the fastest growth due to the volume of spatial information produced all day long.

The systematic and intelligent analyses of the remote sensing images provide a unique opportunity for understanding how, when, and where changes take place in our world. Precious information exploited from spatial repositories has been promoting benefits on many areas, such as agricultural [1, 2] (forecast of harvests and soil erosion), hydric [3] (use of water resources and verification of the water quality), urban [4] (urban planning and demographic inferences), forest [5–7] (monitoring deforestation and biomass control), limnology [8] (characterization of aquatic vegetation and identification of water types), meteorology [9] (weather and climate studies), air traffic [10] (information for safety in the air), and national security [11] (military strategic planning of operations and missions).

However, the manual analysis of huge databases is an extremely inconvenient task for human experts. Despite professionals such as physicists, meteorologists, and ecologists trained to analyze spatial data, the semi-automatic and intelligent interpretation of these data can be a useful tool to leverage the monitoring of the earth surface.

Data mining (DM) arises as solution to detect precious patterns semiautomatically and intelligently in huge databases. DM is defined as the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [12, 13]. Image mining (IM) is also a challenging field which extends traditional data mining from structured data to unstructured data such as image data [14]. IM deals with the extraction of implicit knowledge, image data relationships, or other patterns not explicitly stored in the huge image databases such as remote sensing and medical database.

Despite the success in different applications, the research community of DM has dealt with some issues mining methodology, user interaction, efficiency and scalability, diversity of database models, and data mining and society [15]. The efficiency and scalability issue is particularly significant as the amount of data currently available are increasing rapidly day by day. Therefore, it is necessary to investigate new technological resources that improve some of these issues.

A multiagent system (MAS) is composed of multiple computing elements, known as agents that interact to pursuit their goals. Agents have software architecture for decision-making systems that are embedded in an environment. Consequently, this technology has been widely adopted in numerous applications to solve significant issues.

Agent mining is a new area under development that deals with interaction and integration of between data mining and intelligent agents, and aims to join resources to solve relevant problems that cannot be tackled by a single technology with the same quality and performance. This technology provides important resources, and promises to solve particular issues of both technologies involved.

In this chapter, we present an introduction about image mining, multiagent systems, and agent mining, as well as an overview of these areas. Besides that, a tool known as AgentGeo will be presented [16–18]. It is a multiagent system for satellite image mining that uses the agent resources to mine image data in remote sensing databases. AgentGeo improves the analysis and application of satellite image mining when compared to other systems. The agents leverage the process of image mining due to properties such as autonomy, interaction, reaction, and initiative. This system has been developed in Java and its functionalities are the creation, edition

and selection of agents, selection and creation of the environment, and the use of agents to mine the satellite images. These agents can support many tasks of the image mining process, as well as improve the performance of the steps of preprocessing, transformation and feature extraction, and classification and evaluation.

The chapter is structured as follows: Section 2 discusses image mining, Section 3 describes multiagent systems, Section 4 presents the agent mining, Section 5 focuses on AgentGeo, and finally, Section 6 presents conclusion and mentions future research work.

2. Image mining

Data from computing systems are produced constantly, thereby causing the unbridled growth in the institutions, industries, and corporations databases. This explosive development is caused by several factors, including: internet versatility, reduction in the price of data storage devices, improvement of data collection tools, popularity of embedded systems, increasing of online work, among others. In addition, data are being made available in various formats such as video, text, image, and spreadsheet.

The data variety and volume are so immense that relevant information becomes hidden within databases. Unfortunately, it is difficult or even impossible for a human being to detect patterns handling huge and diversified databases. Several specialists such as economists, statisticians, forecasters, and communication engineers worked with the idea that patterns from data can be reached automatically, identified, validated, and used for various purposes [19]. Therefore, the need to assist the specialist in the extraction of knowledge from huge databases originated the knowledge discovery databases (KDD).

KDD is defined as the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [12]. This process is composed by the following steps selection, preprocessing, transformation, data mining, and interpretation/evaluation [12, 13]. Data mining (DM) is a particular step in KDD process, where specific algorithms for extracting patterns are applied [12, 13]. However, the term data mining has become popular in the database field, used by statisticians and data analyst like synonymous for KDD.

Data mining technology is application oriented and incorporates a variety of techniques, tools, and algorithms capable of extracting relevant information from a wide and diversified collection of databases. Image mining (IM) is a potential technology for data mining, and also a challenging field which extends traditional data mining from structured data to unstructured data such as satellite images, medical images, and digital pictures. Structured data patterns are different from unstructured data patterns. Extracted patterns from image databases are not easily interpreted and understood. Consequently, IM is considered more than just an extension of data mining. It is an interdisciplinary endeavor that incorporates knowledge of important areas such as machine learning, image processing, computer vision, data mining, database, and artificial intelligence [20, 21].

In fact, IM is different from computer vision and image processing. The focus of these areas is extracting specific features from a single image, whereas the IM makes efforts for extraction patterns stored in the huge image databases. This implies in all aspects of databases such as the indexing scheme, the storage of images, and the image retrieval [20, 21].

The image mining process is shown in **Figure 1**. Everything starts from an image database where data are being stored. There are often inconsistent data that need to pass through a preprocessing step in order to improve the level of database quality. Image processing techniques are applied on this step, which are mathematical operations to change the pixel values of images, such as filtering, histogram equalization, image subtraction, image restoration, and others [22]. In the transformation and feature extraction, the images undergo some transformations until identified the relevant objects present in these data. Then features from these objects are extracted, such as edge shape, texture, and length. Obtained such features, the mining step can be carried out using data mining techniques to discover significant patterns automatically or semiautomatically. These patterns need to be evaluated and interpreted by a specialist to obtain the knowledge, which can be applied to applications and can be useful on decision-making processes or on problem understanding.

Image mining process is analogous to data mining process. However, there are important differences between relational databases and image databases [20]. Some differences are as follows:

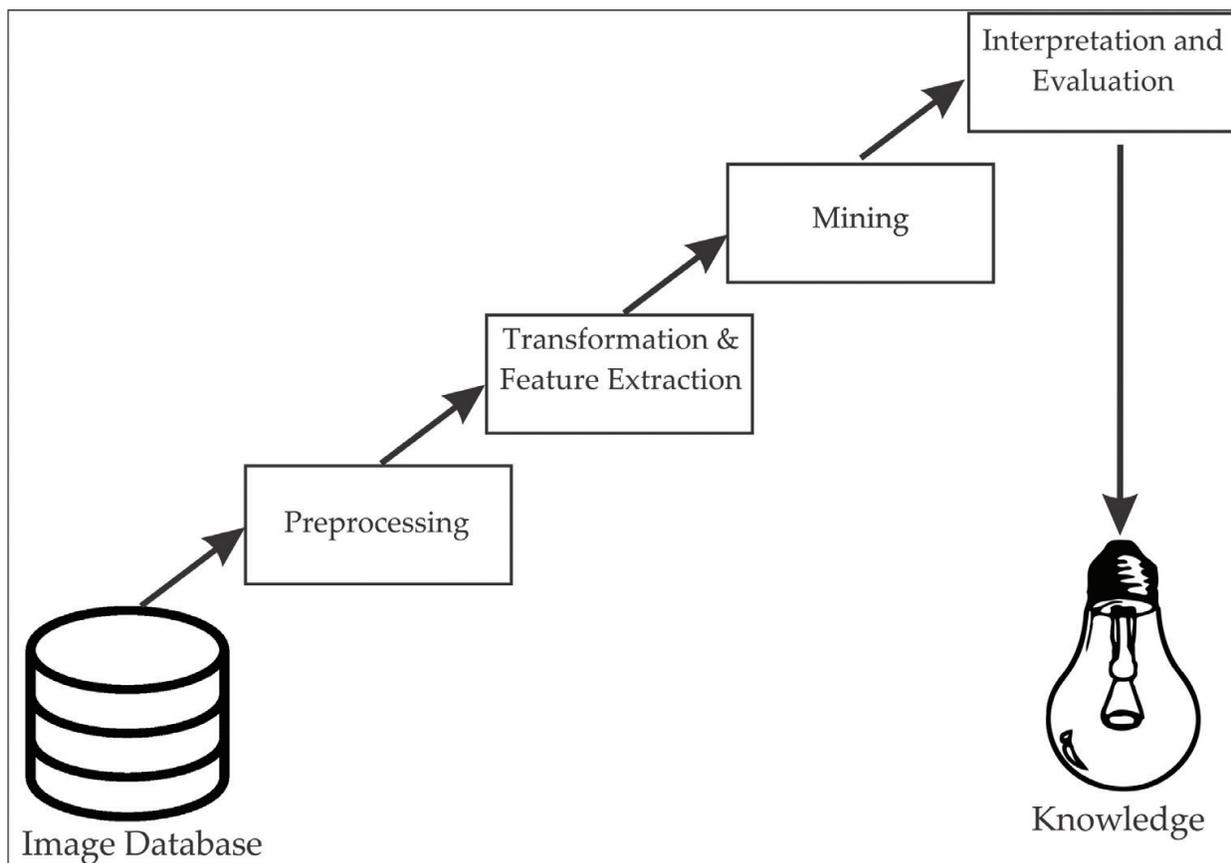


Figure 1. Image mining process [20].

- *Domain dependency*. Real-world activities belong to a given domain, and consequently have specific features and elements. In relational databases, the data values are semantically meaningful. However, at image databases, the identification of elements, their classes, and relationships are linked to the context itself, and a same image can have different information inherent to different domains.
- *Spatial information* (independent versus dependent position). In image databases, a simple image is composed by several elements (pixels). Each pixel is related to its neighbors, often forming a homogeneous region. Due to this, the image miners try to overcome this problem by extracting position-independent features from images before mining useful patterns.
- *Unique versus multiple interpretation*. In relational databases, the data values are easily understood. For example, field person is Paul; we already understood that the field stores a person's name. However, in image databases, an image data may reproduce ambiguous interpretations for the same visual pattern. For example, a simple intensity data can be seen like red, orange, or yellow.

Image mining is a promising and vast field, incorporating mature techniques. Despite the field is under development, there are techniques frequently used for object recognition, image indexing and retrieval, image classification and clustering, association rule mining, and neural network [23–26]. Besides that, image mining has become increasingly important due to its application in many areas such as health, meteorology, aerospace, agriculture, industry, air traffic, spatial research, among others.

2.1. GeoDMA: geographic data mining analyst

Satellite image mining, also known as remote sensing image mining, is an image mining process. Remote sensing image mining deals specifically with the challenge of capturing patterns, processes, and agents present in the geographic space, in order to extract specific knowledge for problem understanding or decision making related to a set of relevant topics, including land change, climate variations, and biodiversity studies. Events like deforestation patterns, weather change correlations, and species dynamics are examples of precious knowledge contained in remote sensing image repositories [27].

The spatial and multiband characteristics of the satellite images differ from the general category of image data. Therefore, remote sensing image mining demands specific image mining tools. GeoDMA is a toolbox for remote sensing image mining that arose based on methodology proposed by Silva et al. [28]. The software incorporates resources for segmentation, feature selection, feature extraction, classification, and multitemporal methods for change detection and analysis of remote sensing data [29].

GeoDMA works as a plugin of the software TerraView GIS [30], which provides the interface for the interpreter to visualize the geographic information stored in databases, to control the database, and also to display the objects' properties [31]. The image mining process in the GeoDMA is shown in **Figure 2**. This process is composed by five steps which described below:

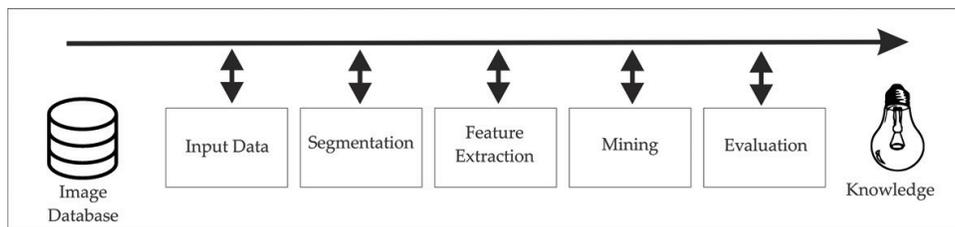


Figure 2. Image mining process at GeoDMA.

2.1.1. *Input data*

This step is responsible for image selection. GeoDMA can only manipulate one image at a time. Therefore, it is necessary to define an image as input for segmentation process.

2.1.2. *Segmentation*

Segmentation is a process where an image pass through various transformations to detect its regions or objects. The level of detail to a segmentation process depends on the application purpose. This means that a segmentation process should stop when all important objects or regions of the application have been detected [29]. Therefore, segmentation is one of the most challenging tasks into digital image processing. The GeoDMA provides four segmentation algorithms as follows:

- Region growing approach based on Ref. [32]
- Segmentation approach based on Ref. [33]
- Chessboard segmentation
- Algorithm based on Ref. [34].

2.1.3. *Feature extraction*

This process aims to extraction attributes as well as spectral and spatial features of objects from the images. Spectral features relate all pixel values inside a region, therefore include metrics for maximum and minimum pixel values, or mean values such as amplitude, dissimilarity, and pixels mean. Spatial features measure the shapes of the regions, including height, width, or rotation [31].

2.1.4. *Mining*

Mining is the process where the algorithms to find a set of models (or functions) are defined, which describe and distinguish classes or concepts. The GeoDMA provides two ways for mining [31]:

- Supervised classification using decision trees based on Ref. [35] and
- Unsupervised classification using self-organizing maps (SOM) [31].

Supervised classification can be divided in two process training and classification. Training is a process supervised by specialists, where data class to known objects is identified, and models to classification of objects are designed. Classification is a process where the build models are used to detect objects that are still unknown. Unsupervised classification is a process that searches for interpretable patterns in data and describe them forming regions known as cluster. The search is based on spectral features such as variance, mean, and light intensity.

2.1.5. Evaluation

The output of GeoDMA is a thematic map. At image mining process, specialists should analyze whether the results are satisfying to application. In the decision tree classification model, the specialist should check whether the regions classified by the built models are valid. In the classifier SOM, generally the result produces more clusters than the desired patterns. In this case, the specialist is responsible for label the patterns according to the application. However, if the results are not satisfactory in both process, previous tasks may be executed again [31].

3. Multiagent system

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators [36]. Perception refers to input information received by agents at a certain moment. In general, the choice of its action at any instant may depend on the entire sequence of observed perceptions. These perceptions can occur through the physical world, via graphical interface, a collection of agents, the internet, or perhaps all combined [37]. The environment, which typically is both computational and physical, might be open or closed, and might or might not contain other agents.

Agents have properties such as autonomy, social ability, reactivity, and proactiveness [37]. Autonomy is the ability to analyze the environment and take their own decisions without the intervention of humans or other agents, controlling their acts and internal state. Social ability is the agent's ability to communicate with other agents or even with human beings using some kind of communication language. Reactivity is the agent's ability to respond in a timely manner, given a history of perception. Proactiveness (or initiative) refers to the agent's capacity of taking initiative in order to achieve their goals. These properties make agents of a technology capable of the following: cooperate in solving problems; share expertise; work in parallel on common problems; develop and implemented modularly; be fault tolerant through redundancy; represent multiple viewpoints and the knowledge of multiple experts; and mainly be reusable.

We can consider four basic kinds of agents: simple reflex agents, model-based reflex agents, goal-based agents, and utility-based agents [36].

- *Simple reflex agents.* Agents that select their actions based on current perception, ignoring the rest of the perceptual history.

- *Model-based reflex agents*. They are a simple reflex agent, but with some differences. These agents maintain some sort of internal state that depends on the perception history and thereby reflects at least some of the unobserved aspects of the current state.
- *Goal-based agents*. These agents know some sort of information about their goals. Based on this information and with internal state, these agents analyze and take their actions.
- *Utility-based agents*. These agents know some sort of information about their goals and have internal state. Besides that, there is a utility function that measures goal performance. This way, these agents choose actions to maximize its utility function.

Agents can learn new concepts and techniques, they can adapt to the needs of different users, they can anticipate the needs of the user, besides others abilities. During the last years, agents have become a powerful technology, which have been adopted in several applications as a solution to solve complex issues that cannot be solved by humans. In the area of remote sensing, for example, a complex task is to analyze remote sensing images, a human being is able to analyze a single remote sensing image, but analyze a significant amount of this data is unlikely because of limited human ability to reason and interpret huge information volumes.

However, when a simple agent cannot solve the problem, a MAS can be implemented. These systems have been studied since 1980s, but were only recognized in the mid of the 1990s. At that time, scientific and industrial interests raised due to the need of exploiting information and modern computing platforms, as well as distributed, open, large and heterogeneous ones [38]. A MAS is formed by two or more agents that interact between them to solve some specific problem. In general, the agents act on behalf of users with different goals. However, agents also may have the same goals, this is determined by the purpose of the system. The interaction in the system occurs through exchanging messages, and it is determined by the ability to coordinate, cooperate, and negotiate between agents [38].

Coordination is a property which aims performing some activity in a shared environment. The degree of coordination is determined by the extent to which they avoid extraneous activity by reducing resource contention, avoiding livelock and deadlock, and maintaining applicable safety conditions [39]. In general, we consider a relevant degree of coordination, when agent activities activities agents are well balanced inside of the environment as well as the operations are being distributed and involved among agents without any failure. There are some reasons why multiple agents need to be coordinated [40], which are as follows:

- Their goals may be conflicting.
- Their goals may be interdependent.
- Agents may have different capabilities and different knowledge.
- Their goals can quickly achieve if different agents work together in a coordinated way.

Cooperation is coordination among nonantagonistic agents, while negotiation is coordination among competitive or simply self-interested agents [39]. At the MAS, agents can cooperate with each other to general goals of the system, or they can compete for their individual goals. Both features must be determined according to the general purpose of the agents into the application.

In the MAS, a key issue is how the agents will communicate. In particular, the communication among processes has long been an important research problem in computer science. In fact, concurrent processes need to be synchronized if there is a possibility that they can interfere with one another in a destructive way [41].

For example, being P1 and P2 two processes, which have access to some shared variable V, when P1 begins to update the value of V, P2 may act at the same moment, but without interfere in the P1 acts. Such communication among processes is like communication among agents. Among other reasons, agents communicate in order to coordinate actions more effectively, to distribute more accurate models of the environment, and to learn subtask solutions from one another [42]. This communication can be implemented through a determined programming language (for example, Java) or can be used an agent communication language like knowledge query and manipulation language (KQML) [41]. There are two types of communication in the multiagent systems [42]:

- *Direct communication*: agents are able to communicate with each other using direct message exchange mechanisms between them. These mechanisms may be constrained in terms of throughput, latency, locality, and agent class.
- *Indirect communication*: consists of indirect transfer of information between agents. For example, when an agent wants to send a message to other agent, it relies on the mediating agent who is responsible for the exchange of information within the environment.

Multiagent systems are increasingly being implemented in several applications such as industry, distributed applications, applications for the internet, games, air traffic control, and teaching environment (e.g., distance education sites). MAS have become more and more important in many aspects of computer science such as distributed artificial intelligence, distributed computing systems, robotics, and artificial life. Some reasons to implement a MAS are

- When the system is complex and the human being cannot or is unable to predict the behavior of that system.
- When it is expensive to keep a team of specialists working.
- When the activity involved put humans at risk.
- When the decision-making process requires performance, agents can solve the problem quickly using parallel processing.
- When it is necessary to ensure information privacy.

4. Multiagent system for image mining

Given the overview of multiagent systems and image mining, we have seen that the areas of agents and data mining emerged separately. Both independent research streams have been created and originally evolving with separate aims and objectives. The area of agents, for example, aims to study the autonomous and independent behavior of agents, and data mining, more

comprehensively, dealt with the KDD process. Despite emerged separately, several similar aspects of these areas appear such as user-system interaction, human roles and involvement, constraints, dynamic modeling, life-cycle and process management, domain factor, and organizational and social factors [43].

Agents is a powerful technology, generally used to solve complex problems in distributed environments where agents can cooperate, coordinate, and communicate their activities in order to reduce the complexity of the problem. Agent research focuses on theoretical, methodological, technical, experimental, and practical issues and the means to handle system complexities [44]. Agent technologies have been contributing to many diverse domains such as software engineering, user interfaces, e-commerce, information retrieval, robotics, computer games, education and training, ubiquitous computing, and social simulation.

Data mining is an application-oriented technology that employs techniques, tools, and algorithms capable of extracting relevant information (or patterns) semi-automatically and intelligently from a massive and diversified collection of datasets. Data mining has been used in web mining services, text mining, medical data mining, meteorological data mining, governmental services, fraud detections, securities, and bioinformatics.

Agents and data mining deal with their specific problems and limitations. Both areas face critical challenges that the other technology might contribute. Agents can leverage the KDD process on data selection, extraction, preprocessing, and integration, and they are an excellent choice for peer-to-peer, parallel, and distributed computing. Agents can bridge the gap between humans and software systems by acting as interfaces that can sense and affect human-mining needs or multisource mining [44]. In the same way, the knowledge acquired through data mining processes provides more stable, predictable, and controllable models for dispatching and planning, and can be used for learning on multiagent systems.

Therefore, agents are elements that can leverage the data mining process, and data mining can contribute significantly to agent's area. A few years ago, researchers have studied means of joining forces between agent and data mining technologies. These studies have given rise to a new research field which became known as agent mining [44–49] or agent-mining interaction and integration [43, 50–52]. Agent mining is the most popular term.

Agent mining refers to the methodologies, principles, techniques, and applications for the integration and interaction of agents and data mining, as well as the community that focuses on the study of the complementarity between these two technologies, for better addressing issues that cannot be tackled by a single technology with the same quality and performance [44, 49].

Agent-mining area is under development; therefore, some issues demands research on theoretical, technological, and methodological aspects. This area follows two fronts of research, which are

- Agent-driven distributed data mining (otherwise known as multiagent-driven data mining, and multiagent data mining): studies ways to use agents to enhance data mining processes and systems. Agents can be used in data mining for different purposes such as agent-based data mining system, agent-based data warehouse, agents for information retrieval, mobile agents for distributed data mining, among others [49].

- Data mining driven agents: investigates issues related to the proper and formal representation of extracted knowledge models from data mining applications such as collaborative learning in multiagents, data mining driven agent learning, reasoning, adaptation and evolution, data mining-driven multiagent communication, planning and dispatching, data mining agent intelligence enhancement [52].

According to Seydim [53], in several steps for knowledge discovery, an agent can be used to automate the individual tasks, including data preprocessing, data mining as well as search for patterns of interest using learning and intelligence in classification, clustering, summarization, and generalization. For example, in the data mining step, an agent can perform automatic sensitivity analysis to determine which parameters should be used in learning. This would reduce the dependency of having domain experts available to examine the problem every time something changes in the environment. The great advantage of using agents in automation of data mining is indicated as their possible support for online transaction data mining [53].

This way, agents can also support and enhance the image mining process in some steps:

- Preprocessing: When an image is added in a database or when it is defined the image database, an agent can perceive these events and automatically examine the data. If the agents perceive anomalies on these images, it can automatically use digital image processing techniques to deal with them. Agent acts can be taken based on rules built by a specialist. This reduces the dependence of having domain experts available for analyzing several images.
- Transformation and feature extraction: agents can be subordinated to transformation and feature extraction task. They can be trained to work together in a much faster way on data transformation, on detecting objects and on the segmentation process.
- Mining: agents can discover significant patterns automatically or semi-automatically in image databases. For example, on classification process, agents can be trained by specialists and can perform intelligently the classification process. Besides that, agents can specialize in one standard and the process may be carried out by several agents simultaneously. This certainly increases the performance and accuracy of the classification process.
- Interpretation and evaluation: knows itself that the evaluation of knowledge is realized usually performed by specialists. In general, specialists have experience in data analysis of a particular domain. In fact, agents can learn with specialists about a particular domain, that is, they can perform this task to support or even substitute the specialists.

Agent mining is a new area with huge opportunities that brings several advantages to multiagent systems, data mining, and machine learning, as well as new derived theories, tools and applications that are beyond any individual technology. The following section presents AgentGeo, a multiagent system of satellite image mining. The tool employs agent to leverage the image mining process, and uses the knowledge of image mining process to agent learning.

5. AgentGeo

AgentGeo emerged based on the ideas proposed by [16, 17], which used agents to mine remote sensing images. This tool is being developed in Java language, and its initial version works along with GeoDMA and TerraView. This system brings advances in the state-of-the-art of multi-agent systems and image mining, presenting relevant resources and precious functionalities. For example, this system implements functions such as creation, edition and selection of agents, selection and creation of the environment and use of agents for image mining. Moreover, it is capable of performing the classification process with multiple images at a time, differently from GeoDMA that is limited to only one image during the entire image mining process.

5.1. System architecture

The system architecture is shown in **Figure 3**.

The image mining process begins in TerraView, which provides the structure to insert and view the images. In fact, the database are created, and several images are inserted. GeoDMA is responsible for processing the data, therefore it receives a single image at a time, and segment and extract the characteristics of each one. These processed images are stored in the image database. AgentGeo connects to the same image database, and receives as input the processed data. So, the environment for the agents can be created, that is, the user defines the image process through GeoDMA. After that, the mining agents responsible for automatically mining these images are defined; that is an automatic process performed by agents and has as output the mined images. These images are stored in the image database, and can be visualized through TerraView. However, AgentGeo provides statistical data about the mining performed by agents, such the number of segments classified by agents. Each process presented may be performed several times.

5.2. Agent's structure

There are two agent types implemented at AgentGeo: the simple agents and a monitor agent. These agents have different properties, features, goals, functions, and behaviors. The simple agents are simple reflex ones, that is, they select their actions based on current perception, ignoring the rest of the perceptual history. Agent's perception occurs in the moment that a user defines the environment and selects the agents on the AgentGeo. Agents have a degree of autonomy, when they are pursuing their particular goals on mining process. The user is not able to forecast the agent actions, because they autonomously decide which goals to pursue.

Besides that, simple agents compete for their goals at the environment just looking to their own well-being, and when their goals conflict with other agents, a monitor agent is responsible for coordinating environmental disputes. For example, two simple agents "A" and "B" are in conflict because of resource at the environment. Soon, the monitor agent perceives the conflict, and takes initiative to finish it. Despite the monitor agent also be a simple reflex agent, we can notice that the behavior of the monitor agent is different from the simple agent. The monitor agent is cooperative and it has the ability to communicate with other agents.

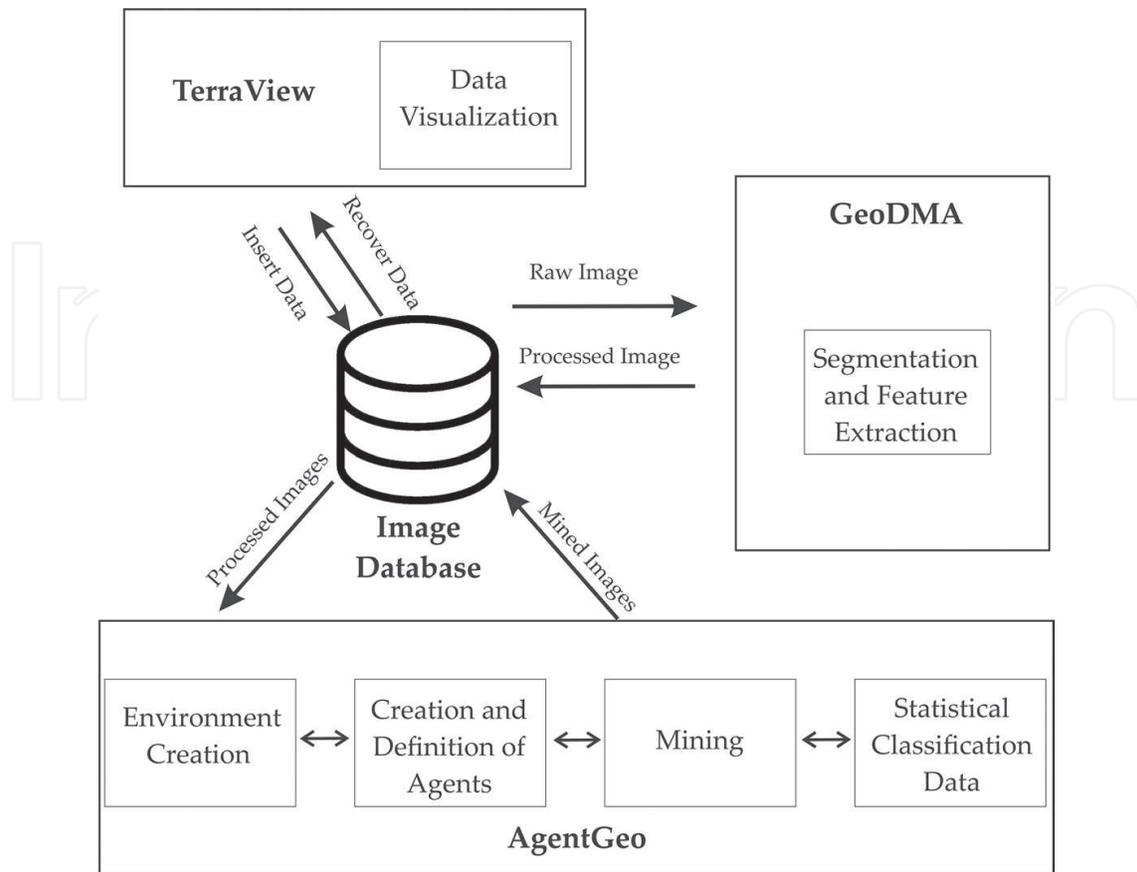


Figure 3. Image mining system architecture.

The communication at the environment occurs indirectly as shown in **Figure 4**. Simple agents only store information about their goals, whereas the monitor agent has the overview of the environment and knows the other agents. Thus, this communication architecture partially solves the communication, coordination, and negotiation problems and considerably reduces the complexity of the MAS.

The agents are implemented through a thread structure and they can perceive and act simultaneously in a certain environment. The simple agents are built by the user, and their internal structure consists of: a description (agent's name), a knowledge base (information about their goals obtained in the training step), and a metadata (extra information about the agent for the usability). This development has four steps as illustrated in **Figure 5**.

Any agent needs a knowledge base to reason about their acts. That base is built through the training phase carried out in GeoDMA and after that comes the segmentation and extraction of features. The agent's knowledge base is formed by a decision tree structure. Firstly, it is defined the image database, and the images are preprocessed on the segmentation and features extraction step. At training step, segment samples that are known to users referring to a specific class are selected by them, according to the agent goals referring to a specific class according to the agent goals. After that, the users can generate a decision tree using GeoDMA, which provides a resource to build it adapted for spatial data mining using the supervised algorithm C4.5 [35].

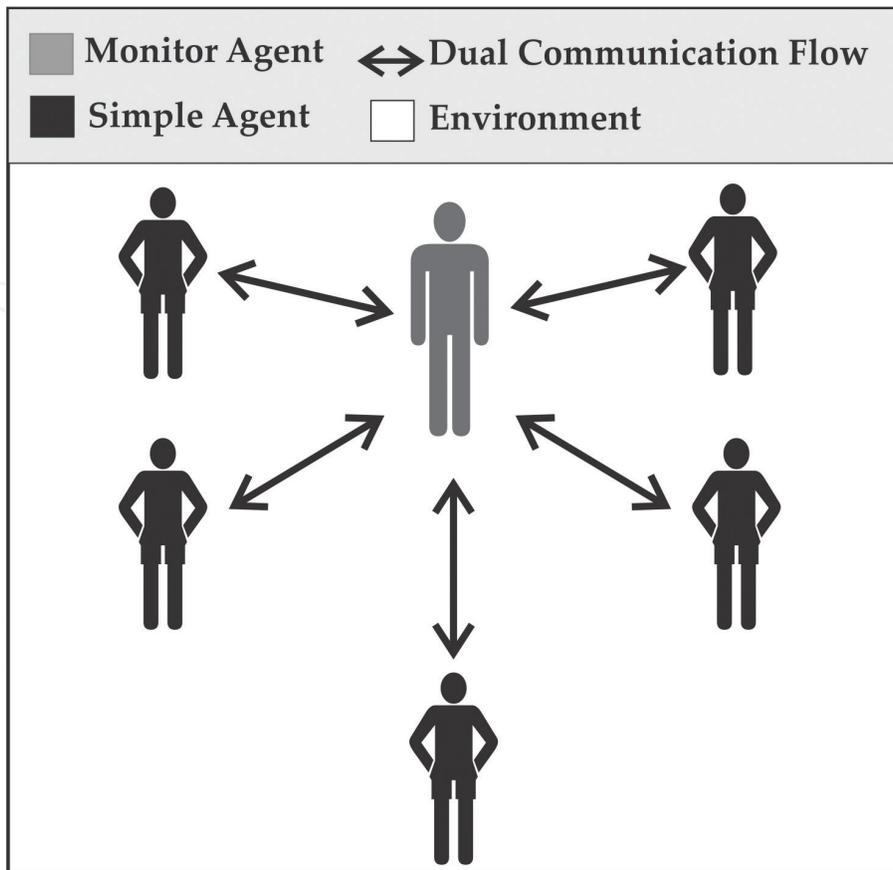


Figure 4. Communication architecture at AgentGeo.

With the decision tree model generated by GeoDMA, the mining step happens automatically. That process is performed by means of thresholds referring to the spatial and spectral attributes of the segments present in the image. All the steps are performed several times in order to have a consistent model. If the user identifies that the results of the model are not satisfactory, he can return to previous steps and perform them again. Otherwise, the model will serve as the knowledge base for the agent, which integration step occurs when the user creates the agent within AgentGeo, that is, the user informs name, metadata, and the knowledge base to the agent.

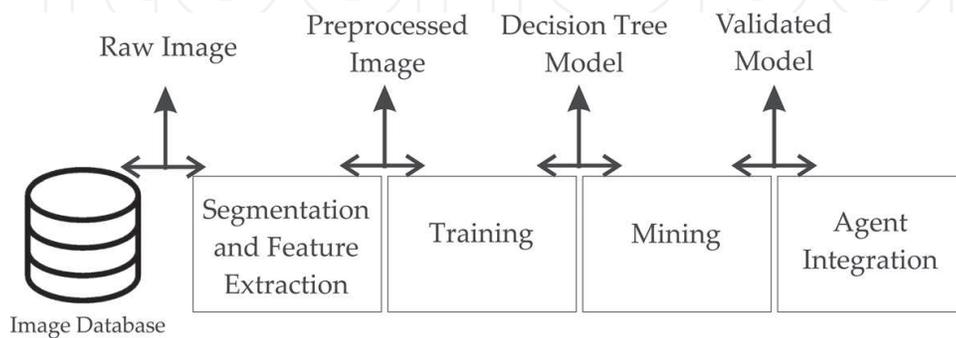


Figure 5. Steps for agent construction at AgentGeo.

All agents of AgentGeo are stored on "agentes.dat" file and can be used whenever necessary. For instance, consider the development of an agent that aims to detect water bodies. The image database is defined, and the images are preprocessed one at a time using GeoDMA. The users select several samples of water bodies and several samples of what is not water bodies. A decision tree model is generated, and the mining process occurs at GeoDMA. After the user analyzed and evaluated positively the models, he can create the agent in AgentGeo.

5.3. Case study

In this section, we will briefly describe two case studies presented in Ref. [17]. The purpose is remote sensing images mining to detect exposed fields, vegetation fields, and water body patterns in Rio Grande Do Norte state, and water bodies patterns in 15 cities of the Ceara state, Brazil. Therefore, three agents were created to detect these patterns, using the methodology for agents construction presented in 5.2 Section.

The image database is formed by LANDSAT-8 satellite images, which are available at <<https://earthexplorer.usgs.gov/>>. At first, image database is build using TerraView and satellite images are inserted. Second, the study region was delimited as well as the area of Rio Grande do Norte state and cities of the Ceara state using digital image processing techniques embedded in TerraView. Third, the images were segmented using the region growing algorithm implemented in GeoDMA, and the spatial and spectral features were extracted from the image database. Fourth, the AgentGeo is loaded, the environment is defined, and the mining agents, implemented in the tool, are selected. At last, the mining results are evaluated by visual inspection using Google Maps images.

This study shows that the methodology of AgentGeo is effective, reaching 92.66% of accuracy on first study and 95.04% on second study. It is important to make it clear that the results of the mining process would be the same if you used only GeoDMA for the mining process. However, the process would occur with one image at a time, which is different from AgentGeo approach, that is capable of mining multiple images, improving the performance of mining process, and keeping the results precision.

6. Conclusion

In this work, we proposed an approach on image mining, multiagent system, agent mining, and presented AgentGeo. In the image mining section, similarities and differences between image mining, data mining, and image mining process were presented. GeoDMA is a toolbox for remote sensing image mining. At the multiagent system section, concepts, proprieties, features, behaviors, structure, and applications about agents and multiagent system were described. Multiagent system for image mining section presented an overview about the new area known as agent mining; we presented and suggested improvements in the integration and interaction of agents and data mining, and of multiagent and image mining. Finally, AgentGeo was introduced, a multiagent system for image mining which uses agent mining to exploit remote

sensing databases. This promising tool brings advances in the state-of-the-art in multiagent systems and image mining, due to relevant resources that leverage the image mining process.

As future work, we hope to integrate more resources into AgentGeo to provide operational advantages, optimization, and innovation. We concluded that agents can be used in several steps in image mining process, that is, we can use them and build modules at AgentGeo for the steps of preprocessing, transformation and feature extraction, and interpretation and evaluation. We can also expand the studies by creating other agents with different goals. For example, agents to road mining, to cloud mining, to deforestation mining, or to mine any spatial object.

Author details

Nicksson Ckayo Arrais de Freitas^{1,2*} and Marcelino Pereira dos Santos Silva¹

*Address all correspondence to: nickssonarraais@gmail.com

1 Rio Grande do Norte State University (UERN), Mossoró, RN, Brazil

2 Federal Rural University of the Semi-arid Region (UFERSA), Mossoró, RN, Brazil

References

- [1] Aguiar DA, Rudorff BFT, Silva WF, Adami M, Mello MP. Remote sensing images in support of environmental protocol: Monitoring the sugarcane harvest in Sao Paulo State, Brazil. *Remote Sensing*. 2011;**3**(12):2682-2703. DOI: 10.3390/rs3122682
- [2] Alba JMF, Schroder VF, Nóbrega MRR. Land cover change detection in Southern Brazil through orbital imagery classification methods. In: Escalante B, editor. *Remote Sensing—Applications*. 1st ed. InTech; 2012. pp. 99-116. DOI: 10.5772/36940. Available from: <https://www.intechopen.com/books/remote-sensing-applications/strategies-of-change-detection-in-southern-brazil-by-orbital-imagery-classification-methods>
- [3] Ding Z, Qi N, Dong F, Jinhui L, Wei Y, Shenggui Y. Application of multispectral remote sensing technology in surface water body extraction. In: *International Conference Audio, Language and Image Processing (ICALIP)*; 2016. pp. 141-144. DOI: 10.1109/ICALIP.2016.7846565
- [4] Matthieu KWK, Maeyer M, Wolff E. The mapping of the urban growth of Kinshasa (DRC) through high resolution remote sensing between 1995 and 2005. In: Escalante B, editor. *Remote Sensing—Applications*. 1st ed. InTech; 2012. pp. 463-478. DOI: 10.5772/38435. Available from: <https://www.intechopen.com/books/remote-sensing-applications/the-mapping-of-the-urban-growth-of-kinshasa>
- [5] Sousa WRN, Couto MS, Castro AF, Silva MPS. Evaluation of desertification processes in Ouricuri-PE through trend estimates of times series. *IEEE Latin America Transactions*. 2013;**11**(1):602-606. DOI: 10.1109/TLA.2013.6502869

- [6] Pizaña JMG, Hernández JMN, Romero NC. Remote sensing-based biomass estimation. Marghany M, editor. *Environmental Applications of Remote Sensing*. InTech; 2016. pp. 1-40. DOI: 10.5772/61813. Available from: <https://www.intechopen.com/books/environmental-applications-of-remote-sensing/remote-sensing-based-biomass-estimation>
- [7] Maciel AM, Silva MPS, Escada MIS. Mining frequent substructures from deforestation objects. In: *IEEE International. Geoscience and Remote Sensing Symposium (IGARSS)*. 2012. pp. 6745-6748. DOI: 10.1109/IGARSS.2012.6352557
- [8] Cho HJ, Mishra D, Wood J. Remote sensing of submerged aquatic vegetation. In: Escalante B, editor. *Remote Sensing—Applications*. 1st ed. InTech; 2012. pp. 297-308. DOI: 10.5772/35156. Available from: <https://www.intechopen.com/books/remote-sensing-applications/remote-sensing-of-submerged-aquatic-vegetation>
- [9] Ashraf A, Rustam M, Khan SI, Adnan M, Naz R. Remote sensing of the glacial environment influenced by climate change. In: Marghany M, editors. *Environmental Applications of Remote Sensing*. 1st ed. InTech; 2016. pp. 99-129. DOI: 10.5772/62134. Available from: <https://www.intechopen.com/books/environmental-applications-of-remote-sensing/remote-sensing-of-the-glacial-environment-influenced-by-climate-change>
- [10] Zhang L, Zhang Y. Airport detection and aircraft recognition based on two-layer saliency model in high spatial resolution remote-sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2016;**10**(4):1511-1524. DOI: 10.1109/JSTARS.2016.2620900
- [11] Ardouin JP, Lévesque J, Roy V, Chestein YV, Faust A. Demonstration of hyperspectral image exploitation for military applications. In: Escalante B., editor. *Remote Sensing—Applications*. 1st ed. InTech; 2012. pp. 493-517. DOI: 10.5772/37681. Available from: <https://www.intechopen.com/books/remote-sensing-applications/demonstration-of-hyperspectral-image-exploitation-for-military-applications>
- [12] Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases. *AI Magazine*. 1996;**17**(3):37-52
- [13] Fayyad U. Data mining and knowledge discovery in databases: Implications for scientific databases. In: *Ninth International Conference on Scientific and Statistical Database Management*; August 11-13. 1997; Olympia, Washington, USA. IEEE; 1997. pp. 2-11. DOI: 10.1109/SSDM.1997.621141
- [14] Cao X, Wang S. Research about image mining technique. In: *Communications and Information Processing, Aveiro, Portugal*. Berlin Heidelberg: Springer-Verlag; 2012. pp. 127-134. DOI: 10.1007/978-3-642-31965-5_15
- [15] Han J, Pei J, Kamber M. *Data Mining: Concepts and Techniques*. 3rd ed. Waltham, MA, USA: Elsevier; 2011. p. 703
- [16] Moura CDG, De Freitas NCA, Silva MPS. Agentes de Mineração de Imagens. In: *Proceeding of the XV Brazilian Symposium on GeoInformatics (GEOINFO)*, SP-Brazil; 2014. pp. 1-6

- [17] De Freitas NCA, Filho PPR, Moura CDG, Silva MPS. AgentGeo: Multi-agent system of satellite images mining. *IEEE Latin America Transactions*. 2016;**14**(3):1343-1351. DOI: 10.1109/TLA.2016.7459619
- [18] De Freitas NCA, Moura CDG, Silva MPS. Sistema multiagente para mineração de imagens de satélite. In: XVII Brazilian Symposium on Remote Sensing, PB-Brazil; 2015. pp. 7351-7358
- [19] Witten IH, Frank E, Hall MA, Pal CJ. *Data Mining: Practical Machine Learning Tools and Techniques*. 4th ed. Cambridge, MA, United States: Todd Green; 2017. p. 621
- [20] Hsu W, Lee ML, Zhang J. Image mining: Trends and developments. *Journal of Intelligent Information Systems*. 2002;**19**(1):7-23. DOI: 10.1023/A:1015508302797
- [21] Zhang J, Wynne H, Lee ML. Image mining: Issues, frameworks and techniques. In: *Proceedings of the Second International Conference on Multimedia Data Mining*. Springer-Verlag; 2001. pp. 13-20
- [22] Gonzales RC, Woods RE. *Digital Image Processing*. 3rd ed. Upper Saddle River, NJ, USA Pearson Education, Prentice Hall; 2008. p. 976
- [23] Dey N, Karâa WBA, Chakraborty S, Banerjee S, Salem MMA, Azar AT. Image mining framework and techniques: A review. *International Journal of Image Mining*. 2015;**1**(1):45-64. DOI: <http://dx.doi.org/10.1504/IJIM.2015.070028>
- [24] Kaur P, Kaur K. Review of different existing image mining techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2014;**4**(6):518-524
- [25] Shukla VS, Vala J. A Survey on image mining, its techniques and application. *International Journal of Computer Applications*. 2016;**133**(9):12-15. DOI: 10.5120/ijca2016907978
- [26] Tripathi A, Jangir H. A study on image mining methods and techniques. *International Journal of Innovative Research in Computer and Communication Engineering*. 2016;**4**(4):7047-7053. DOI: 10.15680/IJIRCCE.2016.0404136
- [27] Silva MPS, Câmara G, Escada MIS. Image mining: Detecting deforestation patterns through satellites. In: Rahman H, editor. *Data Mining Application for Empowering Knowledge Societies*. UK/USA: IGI Global; 2009. pp. 54-75. DOI: 10.4018/978-1-59904-657-0.ch004
- [28] Silva MPS, Câmara G, Souza RCM, Valeriano DM, Escada MIS. Mining patterns of change in remote sensing image databases. In: *Fifth IEEE International Conference on Data Mining*; November 2005; Houston, Texas, USA. Houston, Texas, USA: IEEE; 2005. p. 8. DOI: 10.1109/ICDM.2005.98
- [29] Korting TS, Fonseca LMG, Câmara G. GeoDMA—Geographic data mining analyst. *Computers & Geosciences*. 2013;**57**:133-145. DOI: 10.1016/j.cageo.2013.02.007
- [30] INPE. São José dos Campos, SP: INPE [Internet]. 2010 [Updated: TerraView 4.1.0]. Available from: www.dpi.inpe.br/terraview [Accessed: March 1, 2017]

- [31] Korting TS. GEODMA: A toolbox integrating data mining with object-based and multi-temporal analysis of satellite remotely sensed imagery [thesis]. Sao Jose dos Campos, SP: INPE, Brazil; 2012. p. 97. Available from: <http://urlib.net/8JMKD3MGP7W/3CCH86S>
- [32] Bins LSA, Fonseca LMG, Erthal GJ, Ii FM. Satellite imagery segmentation: A region growing approach. In: Brazilian Symposium on Remote Sensing; April 14-19, 1996; Brazil. Salvador: INPE; 1996. pp. 677-680
- [33] Baatz M, Schape A. Multiresolution Segmentation—an Optimization Approach for High Quality Multi-Scale Image Segmentation. AGIT-Symposium Salzburg; 2000. pp. 12-23
- [34] Korting TS, Fonseca LMG, Câmara G. A geographical approach to self-organizing maps algorithm applied to image segmentation. In: Advanced Concepts for Intelligent Vision Systems. Ghent, Belgium: Springer Berlin Heidelberg; 2011. pp. 162-170. DOI: 10.1007/978-3-642-23687-7_15
- [35] Quinlan JR. Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research*. 1996;4(1):77-90. DOI: 10.1613/jair.279
- [36] Russel ST, Norvig P. *Artificial Intelligence: A Modern*. 3rd ed. United States of America: Pearson Education; 2010. p. 1132
- [37] Wooldridge M, Jennings N. *Intelligent agents: Theory and practice*. *The Knowledge Engineering Review*. 1995;10(2):115-152. DOI: 10.1017/S0269888900008122
- [38] Wooldridge M. *An Introduction to MultiAgent Systems*. 2nd ed. United Kingdom: John Wiley & Sons Ltd; 2009. p. 462
- [39] Weiss G. *Multiagent Systems: A Modern Approach to Distributed Modern Approach to Artificial Intelligence*. Cambridge, Massachusetts and London, England: The MIT Press; 2000. p. 620
- [40] Bellifemine F, Caire G, Greenwood, D. *Developing Multi-Agent Systems with JADE*. 1st ed. United Kingdom: John Wiley & Sons Ltd; 2007. p. 285
- [41] Wooldridge M. *An introduction to Multiagent Systems*. 1st ed. United Kingdom: John Wiley & Sons; 2002. p. 348
- [42] Panait P, Luke S. Cooperative Multi-Agent Learning: The State of the Art. In: *Autonomous Agents and Multi-Agent Systems*. Springer Science; 2005;11(3):387-434. DOI: 10.1007/s10458-005-2631-2
- [43] Silva CVS, Ralha CG. Detecção de Cartéis em Licitações Públicas com Agentes de Mineração de Dados. *Revista Eletrônica de Sistema de Informações*. 2011;10(1):1-19. DOI: 10:5329/RESI.2011.1001
- [44] Cao L, Gorodetsky V, Mitkas PA. Agent mining: The synergy of agents and data mining. *IEEE Computer Society*; 2009;24(3):64-72. DOI: 10.1109/MIS.2009.45
- [45] Cao L, Gorodetsky V, Mitkas PA. Agents and data mining. *IEEE Computer Society*. 2009;24(3):16-17. DOI: 10.1109/MIS.2009.54

- [46] Cao L, Luo D, Zhang C. Ubiquitous intelligence in agent mining. In: Cao L, Gorodetsky V, Liu J, Weiss G, Philip SY, editors. Agents and Data Mining Interaction. Berlin Heidelberg: Springer; 2009;24(3):23-35. DOI: 10.1007/978-3-642-03603-3_3
- [47] Cao L. Introduction to agent mining interaction and integration. In: Cao L, editor. Data Mining and Multi-Agent Integration. USA: Springer; 2009. pp. 3-36. DOI: 10.1007/978-1-4419-0522-2_1
- [48] Cao L, Zhang C, Yu PS, Zhao Y. Agent-driven data mining. In: Cao L, Zhang C, Yu PS, Zhao Y, editors. Domain Driven Data Mining. USA: Springer; 2010. pp. 145-169. DOI: 10.1007/978-1-4419-5737-5_7
- [49] Cao L, Weiss G, Yu PS. A brief introduction to agent mining. Autonomous Agents and Multi-Agent Systems. 2012;25(3):419-424. DOI: 10.1007/s10458-011-9191-4
- [50] Ralha CG, Silva CVS. A multi-agent data mining system for cartel detection in Brazilian. Expert Systems with Applications. 2012;39(14):11642-11656
- [51] Ralha CG. Towards the integration of multiagent applications and data mining. In: Cao L, editor. Data Mining and Multi-Agent Integration. USA: Springer; 2009. pp. 37-46. DOI: 10.1007/978-1-4419-0522-2_2
- [52] Cao L, Luo C, Zhang C. Agent-mining interaction: An emerging area. In: Gorodetsky V, Zhang C, Skormin VA, Cao L, editors. Autonomous Intelligent Systems: Multi-Agents and Data Mining. St. Petersburg, Russia, June 3-5: Berlin Heidelberg: Springer; 2007. pp. 60-73. DOI: 10.1007/978-3-540-72839-9_5
- [53] Seydim AY. Intelligent agents: A data mining perspective. TechReport. 1999. p. 19