Auton Agent Multi-Agent Syst (2012) 25:419–424
DOI 10.1007/s10458-011-9191-4

A brief introduction to agent mining

Longbing Cao · Gerhard Weiss · Philip S. Yu

Published online: 9 May 2012

Abstract Agent mining is an emerging interdisciplinary area that integrates multiagent systems, data mining and knowledge discovery, machine learning and other relevant areas. It brings new opportunities to tackling issues in relevant fields more efficiently by engaging together the individual technologies. It will also bring about symbiosis and symbionts that combine advantages from the corresponding constituent systems. In this editorial, we briefly introduce the concept of agent mining, the main areas of research, and challenges and opportunities in agent mining. Finally, we give an overview of the papers in this special issue.

Keywords Agent mining · Multiagent system · Data mining · Knowledge discovery · Machine learning

1 The concept of agent mining

Agent mining [1,6] refers to the methodologies, technology, tools and systems that synthesize multiagent technology, data mining and knowledge discovery, machine learning and other relevant techniques such as statistics and semantic web for better addressing issues that cannot be tackled by any single technique with the same quality and performance. The concept of agent mining is interchangeable with other concepts such as *agents and data mining*

L. Cao (✉)
Advanced Analytics Institute, University of Technology, Sydney, Australia
e-mail: longbing.cao@uts.edu.au

G. Weiss
Department of Knowledge Engineering, Maastricht University, Maastricht, The Netherlands
e-mail: gerhard.weiss@maastrichtuniversity.nl

P. S. Yu
Department of Computer Science, University of Illinois at Chicago, Chicago, IL, USA
e-mail: psyu@cs.uic.edu

interaction and integration, *agent mining interaction and integration* [5,7,8], or *data mining and multiagent integration* [2,3,12,13].

The interaction and integration between agent technology and data mining and machine learning come from the intrinsic challenges, needs and opportunities faced by the constituent technologies both respectively and mutually. New challenges are appearing with the emergence of new computing mechanisms such as behavior computing, cloud computing, and social computing. Agent mining brings about multi-fold advantages to multiagent systems, data mining, and machine learning, as well as new derived theories, tools and applications that are beyond any individual technology. For instance, in a multiagent system, agents can collaborate with one another to work towards a mutual goal per predefined policies. With data mining and machine learning, the analysis of a counterpart's historical behaviors and the detection of current behaviors can effectively optimize agent collaboration performance and enhance capabilities for tackling exceptions and conflicts. By the same token, cloud analytics may be greatly enhanced by engaging automated cooperation between agent-based computing components. The adaptability of norms and protocols may be enhanced in a social system through learning results from distributed agents in historical and real time. These examples illustrate the potential of agent mining in handling challenges in individual communities, and bringing about new opportunities for creating new technologies, tools and systems that cannot be delivered with any single technology. For example, the collaboration of multiple trading agents, in which each contributes an optimal trading strategy learned from the historical market data, can lead to a better trading performance [2,17].

2 Main areas of research

The research of agent mining started with the involvement of data mining with agents to enhance agent learning [10,19]; thereafter, the main areas of research in agent mining have quickly evolved. Broadly speaking, the research in the relevant communities falls in the following main areas [2,3,5,7,8,12,13,15]:

- *Mutual problems* Agents and data mining share some mutual research issues; for instance, representation of constraints and design of constraint-based systems and constraint-based data mining, the involvement of domain knowledge, knowledge representation, ontology and semantics for system design and content learning, organizational factors in conceptual modeling and system construction, social issues such as security, privacy and trust of data and policies, the role of data, and human–system interaction in interface design and modeling. While these issues are explored from different aspects in their respective fields, they are all relevant to both agents and data mining. For instance, [5] collects papers on ubiquitous intelligence in agent mining and pattern-based agent performance evaluation; in [7], output is reported about integrating workflow into agent-based distributed data mining; [12] includes papers on ontology issues in multiagent distributed learning; in [13], outcomes are reported on evaluating knowledge intensive multi-agent systems, semantic modeling, security issues in distributed agent mining; the relationship between multiagent systems and distributed data mining is introduced in [11].
- *Agent-based data mining* This is also called agent or multiagent-driven data mining, or multiagent data mining, in which agents and agent technology are used to enhance data mining processes and systems. As a major research direction, multiagents are used in data mining for different purposes. Typical research includes the agent-based data mining system, agent-based data warehouse, agents for information retrieval, interface agents for

interactive data mining, mobile agents for distributed data mining, agents for distributed learning [18] and parallel learning [20], agent-based clustering, information gathering agents, and automated data mining using agents for mediation. In [2], for instance, 10 selected papers address different aspects of agent-based data mining, such as an agent approach to analyzing the data stream in complex uncertain environments, a multiagent system for large data clustering, a multiagent data clustering algorithm, agent mining for an interactive web environment for psychometric diagnostics, an agent framework for anomaly detection in distributed firewalls by data mining, competitive-cooperative automated reasoning from distributed and multiple sources of data, normative multiagent enriched data mining to support e-citizens, using multiagent systems for the formalization of groups, agent-based video content identification, and data mining using watermark-based filtering. In [5], research outcomes are introduced on agent-enriched data mining using an extendable framework, autoclustering, and a self-organizing multiagent system for intrusion detection; [12] selects papers on distributed agent-based data preprocessing, meta-reasoning methods for agent intention modeling, and agent-based community clustering based on ontology annotations; peer-to-peer data mining is discussed in [13].

– *Data mining-driven agents* This typically seeks to empower agents and multiagent systems by data mining and knowledge discovery. Typical research work involves data mining for enhancing agent learning and adaptation, data mining for user modeling, data mining-based personalized agents, data mining for empowering recommendation agents, data mining-driven trading agents and trading strategies, data mining-based agent assistants, data mining for agent norm optimization, and data mining for agent behavior analysis. For example, in [2], selected papers discuss about mining temporal patterns to improve agent behaviors, and equipping agents with commonsense knowledge acquired from search query logs. In [7], outcomes are reported about a data mining approach to identify obligation norms in agent societies, probabilistic modeling of a mobile agent's trajectories, and mining real-time sensory patterns in agents. Findings are introduced in [5] about a sequence mining method for predicting bidding strategies for trading agents. Gorodetsky et al. [12] collects papers on the formation and conservative configuration of agent coalitions, and on predicting agent behaviors. In [13], we can find work on role-based decision mining for multiagent emergency response management.
– *Agent learning* This is typically to integrate machine learning techniques in multiagent systems, and to create intelligent and adaptive agents [19]. Typical research work involves reinforcement learning, especially for complex real world problem-solving, for instance, how to efficiently handle many states, many agents, and continuous strategy spaces in a complex environment, such as with incomplete information, limited communications with other agents, and cooperative or competitive settings. Typical techniques for designing agent learning algorithms fall in computational, descriptive, normative, prescriptive-competitive, and prescriptive-noncooperative areas. Typical progress includes joint action learning [9], Nash-Q learning [14], and Gradient Ascent algorithms [16].
– *Agent mining applications* As many references show, the integration of agents and data mining and the development of agent mining are driven by broad and increasing applications in many areas for a variety of purposes. Typical applications include simulation of artificial immune systems, artificial stock markets with programmatic trading mechanisms, distributed data analysis, personalized service and recommendation in e-commerce, web personalized assistants, network intrusion detection systems, peer-to-peer services and systems, supply chain management, trading agents and systems, web and social networks, robots and games.

3 Challenges and opportunities

Agent mining is experiencing an evolutionary process that typically exhibits such characteristics as (1) from one-way interaction (for instance, agents for data mining) to two-way synergy, and (2) from single side need driven (for example, data mining is used to enhance agent behavior understanding) to mutual need driven (an example is a trading agent empowered by trading patterns detected in markets). As a new area, there are huge opportunities open for us to explore which permit us, in particular, to address three fundamental questions.

- How can agents and data mining interact and be integrated?
- What methodologies are necessary and suitable for interacting, integrating and complementing agents and data mining?
- What are the lifecycle, processes and work mechanisms for interacting, integrating and complementing agents and data mining?

In line with the existing principal areas of research, we highlight the following key research challenges and opportunities that address the above three fundamental problems in agent mining.

- *Theoretical foundation in agent mining* This involves the identification and exploration of fundamental issues in enabling the interaction and integration between agents, data mining, machine learning, and relevant technologies. Typical challenges are which methodologies are suitable for agent mining research and applications, how to model and involve human intelligence and human roles in agent mining systems and applications, how to model and involve organizational, environmental and social factors in agent mining design, rules, policies and systems, and how to narrow the gaps in developing actionable agent mining tools and systems.
- Issues in agent-enhanced data mining. In data mining, agents can play an essential role in tackling emerging challenges such as isolation, distribution, mobility, dynamics and large scale data sources in sophisticated enterprise environments, and the privacy and security protection of data and resulting learning outcomes. In particular, there are opportunities with cloud computing and Internet of Things, as well as direct human interaction and intervention with multiple data mining models and systems, and the need for combining and coordinating multiple data sources, multiple models, and multiple feature sets on demand to form a global problem-solving solution.
- *Issues in data mining-empowered agents* Data mining and machine learning can complement and strengthen agent intelligence and multiagent systems in many aspects; for example, making agents and agent-based problem-solving truly automated, adaptive, dynamic, inductive and personalized, enabling agent-oriented rules, norms, policies, processes and reasoning to be more efficient, robust and proactive by identifying hidden information in a large scale of historical data and behaviors in agent networks, and by integrating historical, real-time and future information of self and other agents in a dynamic and distributed manner at the run time, the combination of autonomy with hidden knowledge, the combination of reasoning with learning, and the processing of uncertain, noisy, incomplete, imbalanced, high frequency, large scale, high dimensional and heterogeneous data for agents and agent organizations.
- *Issues in agent learning* Today the agent learning field is dominated by work on reinforcement learning and, specifically, by research conducted at the intersection of reinforcement learning and game theory [19]. New opportunities appear in transfer learning, swarm intelligence, evolutionary learning, multi-view learning, multi-strategy learning, and parallel inductive learning. Agent learning in complex systems and complex environments may

be more demanding when it is positioned in enterprise applications. In addition, the accumulation and availability of big data, big analytics and intelligence meta-synthesis [4] will further challenge and also enable the learning capability in a complex environment.

– *Agent mining tools and applications* Data mining has gained wide acceptance, with applications widely seen in many industry sectors, such as fraud detection, risk management, and anomaly detection for online business, e-commerce, telecommunication, banking and capital investment. This is largely due to the availability of several commercial and research platforms, and data-centric quantitative requirements and validation. In agent mining, there is an urgent and essential need to create effective tools, platforms and applications to support the exploration of the above research issues from both research case study and practice perspectives.

4 Organization of special issue

This special issue on Agent Mining reflects a partial picture of the most recent research and development in agent mining. Through a very rigorous and competitive selection process, four papers are included in this special issue. They involve the following areas of research discussed in Sect. 2: agent-based data mining, data mining-driven agents, agent learning, and agent mining applications.

Chaimontree, Atkinson and Coenen address the issue of agent-based clustering through engaging agent negotiation. They propose a framework for multi-agent-based clustering which enhances clustering by agents. This framework enables agents to negotiate to improve on an initial clustering configuration. Experiments on incorporating k-Means and KNN into the framework have shown an overall positive effect. It is demonstrated that the negotiation between agents improves the clustering performance.

Wardeh, Coenen and Bench-Capon propose a mechanism for multiagent-based classification from argumentation experience. It is based on the observation that classification can be conducted as a process whereby a group of agents "argue" about the classification of a given case according to their experience as recorded in individual local data sets. The arguments are mined and expressed in the form of association rules, which are viewed as generalisations of an individual agent's experience. This paper shows that the agent argumentation from experience can enhance classification performance.

The paper by Hang and Singh addresses the issue of data mining-enhanced agents. The graph similarity is utilized for personalized recommendations for agent networks. They map an agent network to a structure graph, and therefore the agent recommendation problem to the graph similarity problem, and define the similarity measurement as a mutually reinforcing relation. The resulting personalized recommendations (with similarity score reflecting trust) are provided to a requester in an agent network. This paper demonstrates the potential of exploring data analytics for strengthening agent recommendations.

Finally, in Emele, Norman, Sensoy and Parsons's paper about learning strategies for task delegation, an agent decision-making mechanism is proposed by combining decision theory, machine learning and ontology reasoning, in which models of other agents are refined through evidence from past dialogues and domain knowledge, and these models are used to guide future delegation decisions. As a result, the dialogical outcomes can be significantly improved. This paper shows that agent learning and decision-making can be greatly improved from analyzing historical and domain information.

Acknowledgments Thanks are given to all the authors, reviewers and editors who have helped to make this special issue possible.

References

1. Agent Mining Special Interest Group. www.agentmining.org.
2. Cao, L. (Ed.). (2009). *Data mining and multiagent integration*. Berlin: Springer.
3. Cao, L., Zhang, Z., Gorodetsky, V., & Zhang, C. (2008). Editor's introduction: Interaction between agents and data mining. *International Journal of Intelligent Information and Database Systems, Interscience, 2*(1), 1–5.
4. Cao, L., Dai, R., & Zhou, M. (2009a). Metasynthesis: M-space, M-interaction and M-computing for open complex giant systems. *IEEE Transactions On Systems, Man, and Cybernetics: Part A, 39*(5), 1007–1021.
5. Cao, L., Gorodetsky, V., Liu, J., Weiss, G., & Yu, P. S. (Eds.) (2009b). *Agents and data mining interaction*. LNCS (Vol. 5680). Berlin: Springer.
6. Cao, L., Gorodetsky, V., & Mitkas, P. (2009c). Agent mining: The synergy of agents and data mining. *IEEE Intelligent Systems, 24*(3), 64–72.
7. Cao, L., Bazzan, A., Gorodetsky, V., Mitkas, P., Weiss, G., & Yu, P. S. (Eds.) (2010). *Agents and data mining interaction*. LNCS (Vol. 5980). Berlin: Springer.
8. Cao, L., Bazzan, A., Symeonidis, A., Gorodetsky, V., Weiss, G., & Yu, P. S. (Eds.) (2011). *Agents and data mining interaction*. LNCS (Vol. 7103). Berlin: Springer.
9. Claus, C., & Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multi-agent systems. In *Proceedings of the 15th international conference on artificial intelligence* (pp 746–752).
10. Davies W. (1994). Agent-based data-mining. http://logic.stanford.edu/~wdavies/Papers/first-year-report.ps.
11. Giannella, C., Bhargava, R., & Kargupta, H. (2004). Multi-agent Systems and distributed data mining. *Lecture Notes in Computer Science,* 1–15.
12. Gorodetsky, V., Liu, J., & Skormin, V. (Eds.) (2005). *Autonomous intelligent systems: Multi-agents and data mining*. LNCS (Vol. 3505). Berlin: Springer.
13. Gorodetsky, V., Zhang, C., Skormin, V., & Cao, L. (Eds.) (2007). *Autonomous intelligent systems: Multi-agents and data mining*. LNCS (Vol. 4476). Berlin: Springer.
14. Hu, J., & Wellman, M. P. (2003). Nash q-learning for general-sum stochastic games. *Journal of Machine Learning Research, 4*, 1039–1069.
15. Mohammadian, M. (2003). *Intelligent agents for data mining and information retrieval*. Hershey: Idea Group Publishing.
16. Singh, S. P., Kearns, M. J., & Mansour, Y. (2000). Nash convergence of gradient dynamics in general-sum games. In *UAI 00: Proceedings of the 16th conference on uncertainty in articial intelligence* (pp 541–548).
17. Symeonidis, A., & Mitkas, P. (2005). *Agent intelligence through data mining*. Berlin: Springer.
18. Tozicka, J., Rovatsos, M., & Pechoucek, M. (2007). A framework for agent-based distributed machine learning and data mining, In *Proceedings of the 6th international joint conference on autonomous agents and multiagent systems* (pp 1–8).
19. Tuyls, K., & Weiss, G. (2012). Multiagent learning: Basics, challenges, and prospects, *AI Magazine,* 33.
20. Weiss, G. A. (1998). Multiagent perspective of parallel and distributed machine learning. In *Proceedings of Agents '98* (pp 226–230).