Applied Soft Computing 26 (2015) 368–377

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Evolutionary collective behavior decomposition model for time series data mining

Zengchang Qin[a], Tao Wan[b,*], Yingsai Dong[c], Yu Du[d]

[a] Intelligent Computing and Machine Learning Lab, School of ASEE, Beihang University, Beijing 100191, China
[b] School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China
[c] Department of Computer Science, University of British Columbia, Canada
[d] Courant Institute of Mathematical Sciences, New York University, USA

ARTICLE INFO

Article history:
Received 8 March 2013
Received in revised form 24 August 2014
Accepted 22 September 2014
Available online 7 October 2014

Keywords:
Minority games
Mixed games
Collective behavior decomposition
Genetic algorithms
Evolutionary mixed games learning

ABSTRACT

In this research, we propose a novel framework referred to as collective game behavior decomposition where complex collective behavior is assumed to be generated by aggregation of several groups of agents following different strategies and complexity emerges from collaboration and competition of individuals. The strategy of an agent is modeled by certain simple game theory models with limited information. Genetic algorithms are used to obtain the optimal collective behavior decomposition based on history data. The trained model can be used for collective behavior prediction. For modeling individual behavior, two simple games, the minority game and mixed game are investigated in experiments on the real-world stock prices and foreign-exchange rate. Experimental results are presented to show the effectiveness of the new proposed model.

1. Introduction

Collective intelligence is a shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making of agents. Collective behaviors can be modeled by agent-based games in which each individual agent follows its own local strategies. Agent-based experimental games have attracted much attention in different research areas, such as psychology [21], economics [4,24] and financial market modeling [6,12,17]. Agent-based models (ABM) of complex adaptive systems (CAS) provide invaluable insight of the highly non-trivial collective behavior of a population of competing agents. Researchers aim to model these systems where the agents involved have similar capabilities, share global information and are competing for limited resources.

In a given complex system populated with a group of agents, it is unrealistic that the whole population follows the same strategy. The basic assumption is the existence of various types of strategies for agents. In this research, we propose a model that the behavior of an agent can be modeled using a simple game theory model. Aggregations of the individual behaviors become the collective behavior of the system. Two learning scenarios are considered in the proposed framework: Learning with *complete information* where all the data of agents' choices and behaviors in each round are available. However in reality, it is always infeasible to obtain all records of agents' choices in each round of the game. Since we have assumed that the collective data are generated from the combination of variant groups of agents' behaviors, how can we decompose the collective data into the combinations of micro-level data. That is referred to as learning with *incomplete information*.

In the proposed framework, the behavior of an agent is modeled by a game and can be determined by a set of parameters. A *genetic algorithm* (GA) can be used to optimize these parameters to get the best approximation of the original system. To further explore collective behaviors, we add variant groups of agents playing different game theory models and proposed a new data mining framework. We estimate the resource-constrained environment parameters to maximize the approximation of the system outputs and test the effectiveness of the proposed model in the real-world stock market. This framework provides a new way to understand the relationship between micro-behaviors and macro-behaviors in complex dynamic systems.

This paper is organized as follows: Section 2 introduces the recent related research. In Section 3, we describe the new model of collective behavior prediction with complete information. In Sections 4 and 5, we investigate the environment with incomplete

* Corresponding author. Tel.: +86 010 82316875; fax: +86 010 82316875.
E-mail addresses: zcqin@buaa.edu.cn, zengchang.qin@gmail.com (Z. Qin), tao.wan.wan@gmail.com (T. Wan), yingsaid@cs.ubc.ca (Y. Dong), ydu1989@gmail.com (Y. Du).

http://dx.doi.org/10.1016/j.asoc.2014.09.036
1568-4946/

information, we use GA to discover the composition of agent behaviors of the system based on the minority game and the mixed game, respectively. In Section 6, we apply the model to predict the real-world financial time series data to verify its effectiveness. Conclusions and discussions are given in the end.

2. Related work

In agent based modeling, repetitive and competitive interactions among agents generate complex behavioral patterns. The aggregation of simple interactions at the micro-level may generate sophisticated structures at the macro-level, providing valuable information about the dynamics of the real-world system [27]. In fact this intricate two-way feedback between micro-structure and macro-structure has been recognized within economics for a very long time [9,15,23]. One of the directions is agent-based financial market modeling [5]. The “bottom-up” models were proposed to use computational and mathematical tools to describe macro features emerging from a composition of individual interacting agents. For example, the Santa Fe Artificial Stock Market, one of the first agent-based financial market platforms born in the late 1980s, looks at financial markets from an agent-based perspective [14]. This virtual market originated from a desire for building a financial market with an ecology of trading strategies, i.e., successful strategies would persist and replicate, and weak strategies would go away. This project makes use of stochastic optimizations tools, such as the genetic algorithms and classifier system, to model the process of learning [10].

Minority game (MG) [3] is arguably the simplest model in agent-based modeling research. In such a game, an odd number (N) of agents successively compete to be in the minority, which can be regarded as a simplified version of *El Farol Bar Problem* [2] created in 1994 based on a bar in Santa Fe, New Mexico. It was one of the best known experimental game models in economics, in which a number of people decide weekly whether go to the El Farol bar to enjoy live music in the risk of staying in a crowd place or stay at home. Formally: N people decide independently to go or stay each week, they have two actions: go if they expect the attendance to be less than $\alpha N (0 < \alpha < 1)$ people or stay at home if they expect it will be overcrowded. There is no prior communication among the agents; the only information available is the numbers who came in past a few weeks. In the minority game, there are an odd number of players and each must choose one of two choices independently at each turn, players who end up on the minority side win. The minority game has the property that no single deterministic strategy may be adopted by all participants in equilibrium [3].

As a new tool for learning complex adaptive systems, the minority game has been applied to various areas especially in financial market modeling [12,13]. However, there are some weakness of using the basic MG model in real-world market data analysis. One critical weakness is that all agents have the same memory length so that the diversity of agents is limited. In order to cover such limitations, several MG variants were proposed:

(1) Grand Canonical Minority Game (GCMG): The GCMG [11] model is an extension of the classical MG, wherein an agent is rewarded for being in the minority, and it has the possibility not to trade, therefore allowing for a fluctuating agent number invested in the market.
(2) Grand Canonical Majority Game (GCMjG): In the GCMjG [20] model, an agent is rewarded for being in the majority instead of in the minority.
(3) Delayed Grand Canonical Majority Game (delGCMjG): In the delGCMjG [1] model, an agent is rewarded similarly to an agent in the GCMjG, but the return following the decision is delayed by one time step to reflect the more realistic market property.
(4) Delayed Grand Canonical Minority Game (delGCMG): The delGCMG [27] model is the analog of the delGCMjG, except for the minority payoff, whereby each agent is rewarded according to how the return at the next time step is compared with her decision taken at the previous time step. In other words, the delGCMG is a delayed GCMG.
(5) Mixed Game: In the real-world markets, some agents play the minority game, which are referred to as “foundation traders” who hope to maximize their profits; while others are just “trend chasers” who choose what the majority do (i.e., majority game). In order to establish an agent-based model which more closely approximate the real market, Gou [7] modifies the MG model by dividing agents into two groups: one group play the minority game and the other group play the majority game, thus this system is referred to as a mixed game model.

Extensive research in econophysics [19] has been done on agent-based experimental games from the perspective of interdisciplinary disciplines such as physics, mathematics and complexity science. For example, Sysi-Aho [25] proposed a genetic algorithm based adaptation mechanisms within the framework of the minority game, and found that the adaptation mechanism leads the market system fastest and nearest to maximum utility or efficiency. Gou [8] studied how the change of mixture of agents in the mixed game model can affect the change of average winnings of agents and local volatilities of the artificial stock market. Wang [26] proposed an extended minority game model called the market-directed resource allocation game (MDRAG) and investigated the influence of agents' decision-making capacity toward the efficiency, stability and predictability of the market system and its phase structure.

Unfortunately, fewer research focus on exploring macro-level collective behavior prediction by understanding the emergent properties of macro-level behavior from micro-level behavior. We can rarely see that agent-based models were put into practice of real market predictions, e.g., predicting fluctuation of the stock prices. In this paper, we assume that the collective data are generated from the combination of micro-behaviors of variant groups of agents employing different strategies. We then model and estimate the resource-constrained environment parameters to maximize the approximation of the system outputs to the real-world test data.

3. Agent-based minority game

3.1. Strategies of agents

Minority games [3] is one of the simplest game theory models. Suppose an odd number of N agents decide between two possible options, say to attend Room A or B in each round of the game. Formally, in round t ($t = 1, 2, \ldots, T$), each agent i takes an action $a_i(t)$ for $i = 1, 2, \ldots, N$ to choose between A and B, or formally:

$$a_i(t) = \begin{cases} A & \text{agent } i \text{ chooses room } A \\ B & \text{agent } i \text{ chooses room } B \end{cases} \quad (1)$$

At the t round of the game, agents belonging to the minority group win. The winning outcome can be represented by a binary function $w(t)$. If A is the minority side, i.e., the number of agents choosing Room A is no greater than $(N-1)/2$, we define the winning outcome $w(t) = 0$; otherwise, $w(t) = 1$. The winning outcomes