Agent-Based Social Systems

Volume 8

Editor in Chief:

Hiroshi Deguchi, Yokohama, Japan

Series Editors:

Shu-Heng Chen, Taiwan, ROC
Claudio Cioffi-Revilla, USA
Nigel Gilbert, UK
Hajime Kita, Japan
Takao Terano, Japan

For other titles published in this series, go to
www.springer.com/series/7188

ABSS–Agent-Based Social Systems

This series is intended to further the creation of the science of agent-based social systems, a field that is establishing itself as a transdisciplinary and cross-cultural science. The series will cover a broad spectrum of sciences, such as social systems theory, sociology, business administration, management information science, organization science, computational mathematical organization theory, economics, evolutionary economics, international political science, jurisprudence, policy science, socioinformation studies, cognitive science, artificial intelligence, complex adaptive systems theory, philosophy of science, and other related disciplines.

The series will provide a systematic study of the various new cross-cultural arenas of the human sciences. Such an approach has been successfully tried several times in the history of the modern science of humanities and systems and has helped to create such important conceptual frameworks and theories as cybernetics, synergetics, general systems theory, cognitive science, and complex adaptive systems.

We want to create a conceptual framework and design theory for socioeconomic systems of the twenty-first century in a cross-cultural and transdisciplinary context. For this purpose we plan to take an agent-based approach. Developed over the last decade, agent-based modeling is a new trend within the social sciences and is a child of the modern sciences of humanities and systems. In this series the term "agent-based" is used across a broad spectrum that includes not only the classical usage of the normative and rational agent but also an interpretive and subjective agent. We seek the antinomy of the macro and micro, subjective and rational, functional and structural, bottom-up and top-down, global and local, and structure and agency within the social sciences. Agent-based modeling includes both sides of these opposites. "Agent" is our grounding for modeling; simulation, theory, and real-world grounding are also required.

As an approach, agent-based simulation is an important tool for the new experimental fields of the social sciences; it can be used to provide explanations and decision support for real-world problems, and its theories include both conceptual and mathematical ones. A conceptual approach is vital for creating new frameworks of the worldview, and the mathematical approach is essential to clarify the logical structure of any new framework or model. Exploration of several different ways of real-world grounding is required for this approach. Other issues to be considered in the series include the systems design of this century's global and local socioeconomic systems.

Editor in Chief
Hiroshi Deguchi
Chief of Center for Agent-Based Social Systems Sciences (CABSSS)
Tokyo Institute of Technology
4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan

Series Editors
Shu-Heng Chen, Taiwan, ROC
Claudio Cioffi-Revilla, USA
Nigel Gilbert, UK
Hajime Kita, Japan
Takao Terano, Japan

chen.shuheng@gmail.com

Shu-Heng Chen · Takao Terano
Ryuichi Yamamoto
Editors

Agent-Based Approaches in Economic and Social Complex Systems VI

Post-Proceedings of The AESCS International Workshop 2009

Springer

chen.shuheng@gmail.com

Editors
Shu-Heng Chen
Professor
AI-ECON Research Center
Department of Economics
National Chengchi University
64 Sec. 2, Chih-nan Road,
Taipei 116, Taiwan, ROC
chchen@nccu.edu.tw

Takao Terano
Professor
Interdisciplinary Graduate School of Science
and Engineering
Tokyo Institute of Technology
4259 Nagatsuta-cho, Midori-ku
Yokohama 226-8502, Japan
terano@dis.titech.ac.jp

Ryuichi Yamamoto
Assistant Professor
Department of International Business
National Chengchi University
64 Sec. 2, Chih-nan Road,
Taipei 116, Taiwan, ROC
ryuichi@nccu.edu.tw

ISSN 1861-0803
ISBN 978-4-431-53906-3 e-ISBN 978-4-431-53907-0
DOI 10.1007/978-4-431-53907-0
Springer Tokyo Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010942346

© Springer 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

chen.shuheng@gmail.com

Preface

Social science is moving in a direction in which its various constituent parts are sharing a common set of foundations, languages, and platforms. This commonality is making the social sciences unprecedentedly behavioral, algorithmic, and computational. At the turn of the twenty-first century, a group of computer scientists and social scientists worked together to initiate new series of conferences and to establish new academic organizations to give momentum to this emerging integration now known as computational social sciences. One of them is the International Workshop on Agent-Based Approaches in Economic and Social Complex Systems (AESCS), which originated in Japan. The first five AESCS workshops were all organized in Japan – Shimane (2001), Tokyo (2002), Kyoto (2004), and Tokyo (2005, 2007). The sixth was the first one to be held outside Japan. It was hosted by National Chengchi University in Taipei, Taiwan, and co-hosted by the Pacific-Asian Association for Agent-Based Approaches in Social System Sciences (PAAA) as its biennial conference.

On the occasion of AESCS'09 we had 39 presentations, which were delivered in single sessions on November 13 and 14, 2009. In addition to the regular presentations, three keynote speeches were given, by Jeffrey Johnson (Open University, UK), Sobei Oda (Kyoto Sangyo University, Japan), and Takao Terano (Tokyo Institute of Technology, Japan). While most of the time the "agent" in agent-based modeling refers to software agents, the increasing involvement of human agents and their interactions with software agents has given agent-based social modeling a new direction to explore, which is known as experimental agent-based modeling or participatory simulation. To feature this new development, AESCS'09 also offered a one-day tutorial on software for software-agent simulations and human-subject experiments. The tutorials included SOARS (lectures by Hiroshi Deguchi, Manabu Ichikawa, and Hideki Tanuma), Netlogo (Bin-Tzong Chih) and z-Tree (Chung-Ching Tai).

As in the previous five events, we also prepared a post-conference publication to archive selected papers as evidence of the advances in computational social sciences. Fourteen papers were selected to be included in this volume, each being reviewed by three to four referees. These 14 papers were then further grouped into six parts. Of these, Part I, "Agent-Based Financial Markets," and Part II, "Financial Forecasting and Investment," have long-standing positions in the literature.

chen.shuheng@gmail.com

We believe that these two topics will continue receiving attention from scholars as well as the general public, particularly after the recent financial tsunami. Part III, "Cognitive Modeling of Agents," is a new direction in agent-based social sciences. Cognitive capacity, as well as other related measures, has been studied by cognitive psychologists for decades. However, only recently has this constraint sensibly been taken into account in constructing artificial agents so as to, from a microscopic viewpoint, better mimic the human behavior observed in, for example, human-subject experiments, or, from a macroscopic viewpoint, to better understand the emergent complex phenomena. The two chapters included in this part are examples of this kind of work.

Agent-based models of complex adaptive systems obviously provide an alternative way of thinking about policy making in a complex and uncertain environment. The flexibility of agent-based models provides us with tremendous opportunities for policy simulation under various scenarios, from the behavior of stakeholders and interaction networking to environmental uncertainties. This information regarding the landscape of outcomes can be particularly useful in evaluating the potential risk of policy regimes. The three chapters included in Part IV, "Complexity and Policy Analysis," address policies related to pension funds, local taxes, and marketing.

Needless to say, all great challenges currently facing primates and human societies are interdisciplinary. The solutions require not just technology, but also fence crossing among the various social sciences. Agent-based modeling as an integration platform within the social sciences is becoming active in tackling these challenges. Part V, "Agent-Based Modeling of Good Societies," is an example of this development. The three chapters included in this part use agent-based modeling to address the issues of human well-being: peace, greenness, and disaster management, respectively. The remaining two chapters are included in Part VI "Miscellany," which extends the volume to applications to organizations and management and a literature review of the computational social sciences.

We do hope that this volume (AESCS'09), as a continuation of the past decade and the opening of a new decade, can stimulate and motivate more prospective readers, particularly young scholars, to join this growing and exciting area and contribute to the flourishing development of the computational social sciences.

AESCS'09 Workshop Chair	Shu-Heng Chen
AESCS'09 Organizing Committee Chair	Takao Terano
AESCS'09 Program Committee Chair	Ryuichi Yamamoto
AESCS'09 Publication Chair	Hiroshi Deguchi

chen.shuheng@gmail.com

Committees and Chairs of AESCS'09

Organizational Chair

Takao Terano, Tokyo Institute of Technology, Japan

Workshop Chair

Shu-Heng Chen, National Chengchi University, Taiwan

Program Committee Chair

Ryuichi Yamamoto, National Chengchi University, Taiwan

Workshop Coordinator

Chung-Ching Tai, Tunghai University, Taiwan

Tutorial Coordinator

Ke-Hung Lai, National Chengchi University, Taiwan

Program Committee

Chair **Eizo Akiyama**, University of Tsukuba, Japan
Yuji Aruka, Chuo University, Japan
Sung-Bae Cho, Yonsei University, Korea
Mauro Gallegati, Universita′ Politecnica delle Marche, Italy
Nobuyuki Hanaki, GREQAM, Universite de la Mediterranee, and University of Tsukuba, France/Japan
Fumihiko Hashimoto, Osaka City University, Japan
Xuezhong (Tony) He, University of Technology, Sydney, Australia
Reiko Hishiyama, Waseda University, Japan
Masayuki Ishinishi, Ministry of Defense, Japan
Kiyoshi Izumi, National Institute of Advanced Industrial Science and Technology (AIST), Japan
Toshiya Kaihara, Kobe University, Japan
Taisei Kaizoji, International Christian University, Japan
Toshiyuki Kaneda, Nagoya Institute of Technology, Japan
Toshiji Kawagoe, Future University-Hakodate, Japan
Hajime Kita, Kyoto University, Japan
Thomas Lux, University of Kiel, Germany

chen.shuheng@gmail.com

Jun-ichi Maskawa, Seijo University, Japan
Takayuki Mizuno, Hitotsubashi University, Japan
Hideyuki Mizuta, IBM JAPAN, Japan
Yutaka Nakai, Shibaura Institute of Technology, Japan
Yosihiro Nakajima, Osaka City University, Japan
Akira Namatame, National Defence Academy, Japan
Isamu Okada, Soka University, Japan
Isao Ono, Tokyo Institute of Technology, Japan
Tamotsu Onozaki, Aomori Public College, Japan
Philippa Pattison, University of Melbourne, Australia
Stefan Reitz, German Federal Bank, Germany
Aki-Hiro Sato, Kyoto University, Japan
Hiroshi Sato, National Defense Academy, Japan
Yoshinori Shiozawa, Osaka City University, Japan
Keiji Suzuki, Future University-Hakodate, Japan
Keiki Takadama, The University of Electro-Communications, Japan
Shingo Takahashi, Waseda University, Japan
Noriyuki Tanida, Kansai University, Japan
Kazuhisa Taniguchi, Kinki University, Japan
Frank Westerhoff, University of Bamberg, Germany

Sponsors

Department of Economics, National Chengchi University
NCCU Top University Program
National Science Council

chen.shuheng@gmail.com

Contents

chen.shuheng@gmail.com

chen.shuheng@gmail.com

Part I
Agent-Based Financial Markets

chen.shuheng@gmail.com

Comprehensive Analysis of Information Transmission Among Agents: Similarity and Heterogeneity of Collective Behavior

Aki-Hiro Sato

Abstract Recent development of Information and Communication Technology enables us to collect and store data on human activities both circumstantially and comprehensively. In such circumstances it is necessary to consider trade-off between personal privacy and public utility. In the present article I discuss methods to quantify comprehensive states of human activities without private information and propose a measure to characterize global states of societies from a holistic point of view based on an information-theoretic methodology. By means of the proposed method I investigate participants' states of the foreign exchange market during the period of the recent financial crisis which started around the middle of 2008. The results show that drastic changes of market states frequently occurred at the foreign exchange market during the period of global financial crisis starting from 2008.

Keywords Bipartite graph · Degree centrality · Shannon entropies · Kullback–Leibler divergence · Jensen–Shannon divergence · Foreign exchange market

1 Introduction

Recent development of Information and Communication Technology (ICT) enables us to communicate with one another via electronic devices and a ubiquitous environment has been realized everywhere from commerce to education. One can further collect and accumulate large amounts of socio-economic data on human activities, and analyze and visualize them in principle. Based on vast amounts of data from socio-economic systems new types of commercial services and research fields have been emerging. Specifically, several researchers in the fields of sociology, economics, informatics, and physics are focusing on these frontiers and have launched data-oriented sciences in order to understand the collective behavior of human groups [1–11].

A.-H. Sato (✉)
Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto, Japan
e-mail: aki@i.kyoto-u.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_1, © Springer 2011

chen.shuheng@gmail.com

However, since our society, which is the sum total of both internal and external states of individuals, is several orders of magnitude more complicated than each individual, it seems to be impossible for us to capture its real total states even if several humans cooperate to capture them. In other words the "complexity" of our society leads to and/or comes from our nescience.

Therefore it is required that we may develop methodologies to compress or extract information from vast amounts of data on the states in our society with higher degrees of freedom than each individual's degrees of freedom. If technological advances make such unobservable circumstances perceptible then they may provide us with predictability and manageability for our society. We further may be able to find arbitrary opportunities obscured due to the nescience and improve several undesirable circumstances.

According to Heinz von Foerster [12] complexity is not of any properties which observed systems possess but it is to be perceived by observing systems. He asks us about it through the following question: *Are the states of order and disorder states of affairs that have been discovered, or are these states of affairs that are invented*? If the states of order and disorder are discovered then complexity means the property of the observed systems. If invented then it is perceived by the observing systems.

In this article, following Foerster's definition of the complexity, we suppose that the degree of order and disorder is relatively determined by the degrees of freedom of an observed system and an observing system. One of the most dominant reasons why we recognize the complexity in observed systems is because finiteness of periods when and abilities where we are able to observe the systems and limitations of our memory and a priori knowledge on them lead to our bounded rationality or nescience. Therefore, if we can overcome nescience with comprehensive data on the observed systems by means of massive computation then we will be able to make complexity change to simplicity. Moreover, if we can invent the definition of the states of order and disorder then we have an ability to perceive the states of order and disorder. Therefore it is important for us to have a conceptual framework for coping with complexity in human societies.

Our own knowledge is a part of *the whole of knowledge*, and *the whole of knowledge* consists of each part of our own knowledge. Therefore it seems to be impossible for us to comprehend the whole of knowledge because of our finiteness. However it is possible for us to know it by intuition. This intuition for *the whole of knowledge* is a "comprehensive" perspective which is expected to lead us to a holistic point of view.

The aim of the present article is to propose methods to quantify and visualize attentions of participants in groups whilst protecting the anonymity of agents from a holistic point of view. Specifically I focus on the foreign exchange market and attempt to comprehensively visualize market states of the foreign exchange market with high-resolution data recorded in an electronic brokerage system.

The rest of this paper is organized as follows. In Sect. 2 a literature survey is carried out through recent studies on data-centric socio-economic sciences. In Sect. 3 I propose an agent-based model of a society consisting of N kinds of groups where M participants exchange contexts, and propose methods to capture states of

chen.shuheng@gmail.com

participants in a practical manner. In Sect. 4 I show results of empirical analysis on states of the foreign exchange market by means of the proposed methods with high-resolution data. Section 5 is devoted to conclusions and used to address future works.

2 Literature Survey

Recently several researchers in a wide spectrum of fields have paid a remarkable amount of attention to massive amounts of comprehensive data. For example, search engines of Web services need massive data about hyperlink connections among Web pages, and electronic commerce systems need to cover various kinds of products. Due to the development of ICT, the Advanced Information Society has already emerged globally and it has eventually made our world to be smaller and smaller. For such circumstances the concept of information explosion has been proposed [13, 14]. This concept is that the total amount of information created by individuals exceeds the individuals' information processing capability. According to recent studies on the information explosion, it is predicted that the total amount of information created by human beings will reach over 1Zbyte/year around 2010.

Studies based on vast amounts of socio-economic data have several branches. Here five kinds of recent studies (financial market data, demographic data, traffic flow data, POS data, and Web-commerce data) are surveyed for the purpose of finding directions to cope with the complexity of human societies.

A large amount of data on financial markets is available because the electronic matching systems of financial markets are spreading all over the world due to the development of ICT. Recent trading is done through electronic platforms and settlement operation is done through electronic clearing systems. Financial market data can be collected through a direct API or through the historical data centers of data providers. Applications of statistical mechanics to finance by means of statistical physics, agent-based modeling and network analysis have progressed during the last decade [1–4].

The launch of E-Stat database by the Japanese government [5] provides us with new technology for data-based understanding of our country. Specifically, based on demographic data everyone can understand the state of our country from a viewpoint of population in principle. Furthermore, real-time demographic data are also available since the technologies to collect human activities via each personal mobile phone have been established [6]. In the near future we will be able to visualize real-time demographics both comprehensively and circumstantially.

Recently, several car navigation companies have launched autonomous sensory navigation services in Japan. As a result, these companies can collect real-time car traffic data via each car navigation terminal. Moreover, by collecting data from many cars one can find roads and points where traffic jams are occurring. Without constructing new infrastructure to collect traffic states they can accumulate real-time traffic data due to the development of Integrated Transport Systems (ITS). Based on

chen.shuheng@gmail.com

such data comprehensive analysis of traffic flows can be conducted in order to cope with traffic jams [7]. Recent developments of traffic measurement technologies have been driving the theoretical development of traffic control and modeling [8].

POS is an abbreviation for "point-of-sales" and all the department stores and supermarkets have introduced this kind of system in order to ring up the amounts at the cash registers. As a result, retail sales can be managed in real-time and data centric operations can be done. On the basis of massive amounts of data, marketing methods have been developed. The statistical properties of expenditure in a single shopping trip show a power-law distribution [9]. A comprehensive analysis of retail sales is one of the prominent directions to be followed in order to bridge between microeconomics and macroeconomics.

Web-based commerce systems enable us to purchase everything from books to electronic equipment, via web sites. The details of consumers and goods can be stored on the data-base engine of each web site. If we can use such data, then we may, in principle, capture real-time demand and supply of all the items which are traded via web sites. Analyzing massive amounts of data on items which are sold via web commerce systems is expected to open a window to new economic theory and service engineering [10, 11].

The common properties of these studies seem to overcome the complexity in socio-economic systems by using massive amounts of data and vast computations. Copious amounts of data on human activities are collected by means of ICT and vast amounts of computation for such data are conducted for the purposes of searching, matching, visualizing and extracting.

Specifically in the literature of Cloud Computing, users practically use their service provider's computers via a rich network infrastructure and store their data on the provider's storage. Therefore, in principle, it is possible for us to integrate collected personal data and to even understand the states of order and disorder of the affairs based on vast amounts of data and of computation.

On the other hand, in such circumstances we have to consider the trade-off between personal privacy and public utility. In this context almost all advanced countries have Private Information Protection Laws. Hence, protection of personal privacy is one of most important issues involved in dealing with personal data in Cloud Computing. If it is permitted for us to secondarily use such data, then it can be a useful infrastructure for us to capture our society and to circulate our knowledge from a comprehensive point of view. In the following sections we consider a method of capturing the global states of affairs whilst protecting personal information.

3 Model and Methods

In this section I discuss methodologies to quantify the total states of affairs created by participants' activities and discriminate them for different observation periods from a comprehensive point of view. Specifically methods to characterize an indicative index with data from all the components of which the system consist.

chen.shuheng@gmail.com

The fundamental ideas are to focus on relative frequencies within group activities, and to quantify and to discriminate patterns based on their relative frequencies. Since the relative frequencies of group activities are quantities which can be computed from the number of activities for an observation period it is possible to count the actions of agents without any knowledge of private information. By repeatedly quantifying the total states of human society for each observation period we will be able to capture changes of circumstances.

According to the definition of information by Gregory Bateson [15] information is defined as *a difference which makes a difference*. In other words information is an ideal element (=meaning) of which a message changes a receiver's state. It is known that Bateson's definition characterizes information from qualitative aspects as compared with Shannon's quantitative definition of information. Moreover it is possible that agents create messages based on their inner state and receive messages from other agents in multi-agent systems. In such systems the production of messages and state variations are repeatedly done. The structure and dynamics of chaining perceptions and actions through communication among agents seem to be related to the stability of societies.

Consider N groups consisting of M agents as shown in Fig. 1. In the context of web commerce systems, agents correspond to consumers and groups correspond to goods or shops. In the context of financial markets agents correspond to traders, and groups are the commodities which are traded by agents. In the context of blog systems agents are bloggers, and groups are contents or communities. If one knows the relative activities of agents on each group then one can know the shares of those groups in society. Though the complete structure of human society is unknown and

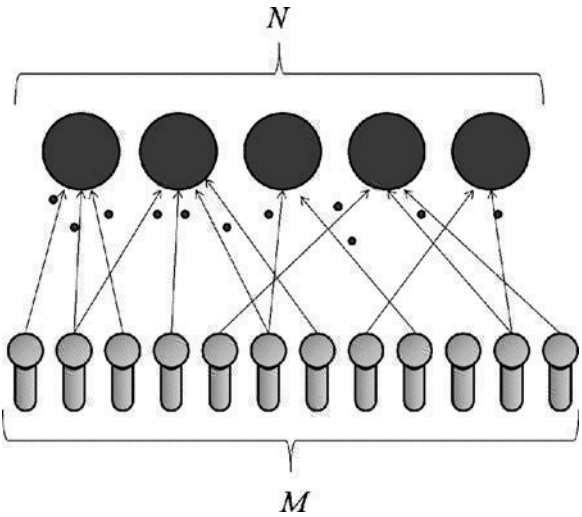


Fig. 1 A conceptual illustration of human society consisting of $\boldsymbol{N}$ groups where $\boldsymbol{M}$ agents exchange messages

chen.shuheng@gmail.com

unobservable, the shares of groups are computable from observations. Such kinds of activity data can be comprehensively collected by observing the arrivals of submissions on host computers.

Let $f_i(T)$ $(i = 1, \ldots, N)$ be the number of observations for the actions of agents in the ith group for the observation period $T : [T\Delta, (T+1)\Delta)(\Delta > 0)$. Then a relative frequency of the ith group's activities can be estimated as

$$p_i(T) = \frac{f_i(T)}{\sum_{i=1}^{N} f_i(T)}. \tag{1}$$

Namely, the relative frequency $p_i(T)$ may approximate the degree of centrality c_i during an observation period T [16]. On the other hand, suppose that random walkers hop from one node to another node with the same probability on a network having an adjacency matrix C_{ij}. Then a stationary residence probability of random walkers at the ith node is equivalent to the ith node's degree of centrality which is defined as

$$c_i = \frac{\sum_{j=1}^{N} C_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij}}. \tag{2}$$

Recently Wilhelm and Hollunder proposed a method to characterize directed weighted networks with several nodes [17]. They consider the normalized weight of the flux between two nodes as the probability for a symbol in the transmitter signal corresponds to the sum of all influxes to/effluxes from a given node. They also propose information-theoretic measures for the normalized weight in order to characterize the shape of networks. In the case of undirected unweighted networks, the normalized weight is equivalent to the degree of centrality as shown in (2).

Since a relative frequency of actions during an observation period T may approximate to a probability for agents to be in a group, its information entropy during the observation period and the information divergence between two observation periods may characterize the state of the system at the period and quantify a distance between two states of the system. If we employ Shannon entropy as information entropy and Kullback–Leibler divergence as divergence then we can describe them as

$$S(T) = -\sum_{i=1}^{N} p_i(T) \log\left(p_i(T)\right) \tag{3}$$

$$KL(T_1, T_2) = \sum_{i=1}^{N} p_i(T_1) \log \frac{p_i(T_1)}{p_i(T_2)} \tag{4}$$

Obviously one has $0 \leq S(T) \leq \log N$, and $S(T) = \log N$ when $p_i(T) = 1/N$. Further, one has that $KL(T_1, T_2) \geq 0$ is satisfied and $KL(T_1, T_2) = 0$ if and only if

chen.shuheng@gmail.com

$p_i(T_1) = p_i(T_2)$ for any-i. In this definition the Shannon entropy is regarded as a special case of the Kullback–Leibler divergence. When we put $p_i(T_0) = 1/N$, we have $KL(T, T_0) = \log N - S(T)$. Namely the Shannon entropy is equivalent to the Kullback–Leibler divergence between the activity state on the network at T and uniform activity state on an undirected fully-connected network.

Furthermore in order to compare the shapes of probability distribution we can choose a divergence from f-divergence, Kullback–Leibler divergence, Jensen–Shannon divergence, and so on [18]. In the case of the f-divergence the similarity of centralities between two observation periods is defined as follow.

Let $f(u)$ be a convex function satisfying $f(1) = 0$. Then the similarity between group states on the T_1-th observation period and those on the T_2-th observation period is defined as

$$D_f(T_1, T_2) = \sum_{i=1}^{N} p_i(T_1) f\left(\frac{p_i(T_2)}{p_i(T_1)}\right) \tag{5}$$

As an alternative divergence the Jensen–Shannon divergence [18] can be adopted.

$$JS(T_1, T_2) = -\sum_{i=1}^{N} \frac{p_i(T_1) + p_i(T_2)}{2} \log \frac{p_i(T_1) + p_i(T_2)}{2} - \frac{1}{2} \sum_{k=1}^{2} S(T_k) \tag{6}$$

One has $JS(T_1, T_2) \geq 0$ and $JS(T_1, T_2) = 0$ if and only if $p_i(T_1) = p_i(T_2)$ for any i.

4 Empirical Analysis

In this section I show results of empirical analysis by means of the proposed methods with high resolution data of the foreign exchange market. The analysis is conducted by using high-resolution data collected by ICAP EBS platform (ICAP EBS Data Mine Level 1.0) [19]. In the exchangeable currency pairs consisting of 24 currencies and five precious metals,[1] 47 kinds of currencies pairs[2] are included in

[1] USD (United States Dollar), CHF (Swiss francs) EUR (Euro), JPY (Japanese Yen), NZD (New Zealand Dollar), AUD (Australia Dollar), GBP (British Sterling) CAD (Canadian Dollar), SEK (Swedish Krona), SGD (Singapore Dollar), HKD (Hong Kong Dollar), NOK (Norwegian Krone), ZAR (South African Rand), MXN (Mexico Peso), DKK (Danish Krone), CZK (Czech Koruna), PLN (Poland New Zloty), HUF (Hungarian Forint), ISK (Iceland Krone), RUB (Russian Rouble) SKK (Slovakia Koruna), TRY (Turkey Lira), THB (Thailand Baht), RON (Romanian Leu), BKT (Basket Currency of USD and EUR), ILS (Israeli Shekel), XAU (Gold), XAG (Silver), XPD (Palladium), XPT (Platinum), and SAU (Small amount of Gold).

[2] The data include records of BID/OFFER for USD/CHF, EUR/USD, USD/JPY, EUR/JPY, EUR/CHF, NZD/USD, AUD/USD, GBP/USD, USD/CAD, AUD/NZD, EUR/GBP, XAU/USD, XAG/USD, EUR/SEK, CHF/JPY, XPD/USD, XPT/USD, USD/SGD, USD/HKD, EUR/NOK, USD/ZAR, USD/MXN, EUR/DKK, EUR/CZK, EUR/PLN, EUR/HUF, EUR/ISK, USD/RUB, GBP/JPY, EUR/SKK, USD/PLN, GBP/CHF, AUD/JPY, USD/TRY, USD/THB, NZD/JPY, CAD/JPY, ZAR/JPY, EUR/ZAR, EUR/RUB, EUR/CAD, GBP/AUD, USD/SEK, EUR/AUD, EUR/RON, BKT/RUB, and USD/NOK.

chen.shuheng@gmail.com

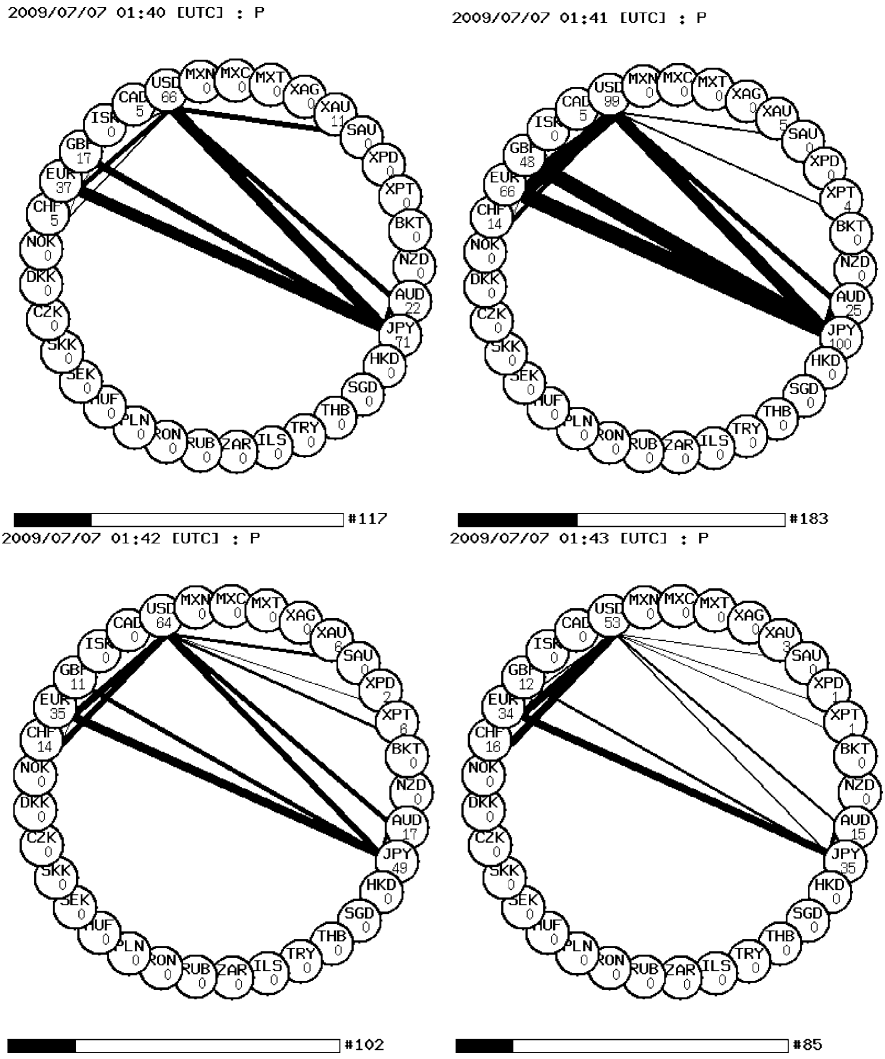


Fig. 2 Network illustrations of the number of quotation activities from 1:40 to 1:44 (UTC) on 7th July 2009. A node (the number in each node shows the number of quotations/transactions) represents currency and a weighted link represents the flow of each currency pair

the data set with 1-s resolution during a period from June 2008 to December 2009.[3] Figure 2 shows weighted network illustration visualizing the number of quotations for each currency pair within 1-min over the whole market. Nodes represent currencies, and weighted links between two kinds of currencies the number of quotations (left) and transactions (right). It is found that the activities of quotations temporally

[3] Totally 123,958,633 records are found in the dataset.

chen.shuheng@gmail.com

vary and several characteristic patterns. It is further confirmed that there are several thick links connected with USD, EUR, GBP, and JPY called major currencies or hard currencies and that precious metals (XAU and XPT) have a connection with only USD.

The relative occurrence rates of quotations and transactions are counted for each week. Since the relative occurrence rates of transactions have been triennially reported by BIS [20], the usefulness of this quantity has been widely recognized in international finance. Emergence of ICT allows us to obtain the high-resolution shares for currencies and currency pairs. This technological advantage is able to improve our perceptual resolution.

Suppose one can observe quotations/transactions about the i-th currency and the j-th currency and count their arrival of quotations or occurrence of transactions for each currency pair on a brokerage system with an interval of $\Delta(>0)$. The quotation/transaction activity is defined as the number of quotations/transactions which market participants enter into the electronic broking system per Δ. We define $f_{ij}(t;T)$ as the quotation activities between the i-th currency and j-th currency $(i,j=1,2,\ldots,N)$ in $[(TS+t)\Delta,((TS+t+1)\Delta))(t=0,1,\ldots,S-1)$ on the T-th observation period. In this analysis we adopt the definition that the activities should be counted in symmetric way $f_{ij}(t,T)=f_{ji}(t,T)$ and assume the condition that there is no self-dealing $f_{ii}(t,T)=0$. Then the density of quotations between the i-th currency and j-th currency can be estimated as

$$A_{ij}(T)=\frac{\sum_{t=0}^{S-1} f_{ij}(t,T)}{\sum_{t=0}^{S-1}\sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}(t,T)}. \tag{7}$$

Obviously it has probabilistic properties, so that, $\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(T)=1$, $A_{ii}(T)=0$, and $0\leq A_{ij}(T)\leq 1$.

Under the assumption that the attention of market participants to the exchangeable currency pairs can be approximated as the centrality of currency pairs, $A_{ij}(T)$ can be empirically estimated by using quotation/transaction frequencies calculated from high-resolution data without knowledge on network structure of market participants. The reason why the centrality is adopted in order to quantify the attention of market participant is because currency pairs (currencies) which are quoted/traded by many participants are focused by many participants. Moreover relative occurrence rates of the i-th currency on the T-th observation period are defined as

$$K_i(T)=\sum_{j=1}^{N} A_{ij}(T), \tag{8}$$

where it has also probabilistic properties, so that, $\sum_{i=1}^{N} K_i(T)=1$ and $0\leq K_i(T)\leq 1$.

chen.shuheng@gmail.com

Since both $A_{ij}(T)$ and $K_i(T)$ may be regarded as fingerprints representing the market states on the observation period T, their shape may describe market states at T. Furthermore since they are probability distributions from their definition, the similarity between them can be evaluated by means of several kinds of information-theoretical divergences.

In order to estimate a total diversification of quotation/transaction activities in a financial market one can adopt the normalized Shannon entropy of the centralities [21] for currencies/currency-pairs defined as

$$H_{cp}(T) = -\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(T)\log A_{ij}(T)}{\log N(N-1)}, \tag{9}$$

$$H_{c}(T) = -\frac{\sum_{j=1}^{N} K_{j}(T)\log K_{j}(T)}{\log N}. \tag{10}$$

If one currency-pair (currency) is exclusively traded then $H_{cp}(T)$ $(H_c(T))$ takes the minimum value 0. Contrarily every currency-pair (currency) is equivalently traded then $H_{cp}(T)(H_c(T))$ takes the maximum value 1. Therefore $H_{cp}(T)$ $(H_c(T))$ can be one of candidates to measure monopolization of currency-pairs (currencies) in the foreign exchange market.

Then the similarity between market states on the T_1-th observation period and those on the T_2-th observation period is defined as

$$D_{cp}^{(f)}(T_1,T_2) = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(T_1) f\left(\frac{A_{ij}(T_2)}{A_{ij}(T_1)}\right), \tag{11}$$

$$D_{c}^{(f)}(T_1,T_2) = \sum_{i=1}^{N} K_{i}(T_1) f\left(\frac{K_{i}(T_2)}{K_{i}(T_1)}\right). \tag{12}$$

Then they have the following properties:

$$D_{cp}^{(f)}(T_1,T_2) \geq 0, \tag{13}$$

$$D_{cp}^{(f)}(T_1,T_2) = 0 \text{ iff } A_{ij}(T_1) = A_{ij}(T_2) \forall i,j, \tag{14}$$

$$D_{c}^{(f)}(T_1,T_2) \geq 0, \tag{15}$$

$$D_{c}^{(f)}(T_1,T_2) = 0 \text{ iff } K_{i}(T_1) = K_{i}(T_2) \forall i. \tag{16}$$

If we choose $f(u) = -\log u$, then they give the Kullback–Leibler divergence

$$D_{cp}^{(KL)}(T_1,T_2) = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(T_1)\log\frac{A_{ij}(T_1)}{A_{ij}(T_2)}, \tag{17}$$

$$D_{c}^{(KL)}(T_1,T_2) = \sum_{i=1}^{N} K_{i}(T_1)\log\frac{K_{i}(T_1)}{K_{i}(T_2)}. \tag{18}$$

As an alternative symmetric divergence the Jensen–Shannon divergence was introduced [22]. They are defined as

chen.shuheng@gmail.com

$$D_{cp}^{(JS)}(T_1,T_2) = H_A\left(\frac{1}{2}\sum_{k=1}^{2} A_{ij}(T_k)\right) - \frac{1}{2}\sum_{k=1}^{2} H_A\left(A_{ij}(T_k)\right), \tag{19}$$

$$D_{c}^{(JS)}(T_1,T_2) = H_K\left(\frac{1}{2}\sum_{k=1}^{2} K_i(T_k)\right) - \frac{1}{2}\sum_{k=1}^{2} H_K\left(K_i(T_k)\right), \tag{20}$$

where $H_A(A_{ij})$ and $H_K(K_i)$ are respectively denoted as the Shannon entropies defined as

$$H_A(A_{ij}) = -\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\log A_{ij}, \tag{21}$$

$$H_K(K_i) = -\sum_{i=1}^{N} K_i \log K_i. \tag{22}$$

From (19) and (20) we can confirm that they have the following properties:

$$D_{cp}^{(JS)}(T_1,T_2) = D_{cp}^{(JS)}(T_2,T_1), \tag{23}$$

$$D_{cp}^{(JS)}(T_1,T_2) \geq 0, \tag{24}$$

$$D_{cp}^{(JS)}(T_1,T_2) = 0 \text{ iff } A_{ij}(T_1) = A_{ij}(T_2)\forall i,j, \tag{25}$$

$$D_{c}^{(JS)}(T_1,T_2) = D_{c}^{(JS)}(T_2,T_1), \tag{26}$$

$$D_{c}^{(JS)}(T_1,T_2) \geq 0, \tag{27}$$

$$D_{c}^{(JS)}(T_1,T_2) = 0 \text{ iff } K_i(T_1) = K_i(T_2)\forall i. \tag{28}$$

Figure 3 shows the normalized Shannon entropies of quotation activities for currency pairs and for currencies, and of transaction activities for currency pairs and currencies for each week. The normalized Shannon entropies obtained from the number of quotations changed around the week beginning from 21st July 2008, from 15th December 2008, 23rd February 2009. Since larger values of the normalized Shannon entropies indicate multi-poly states of currency exchange, these values of quotations for a period from August 2008 to February 2008 and for a period from August 2009 seem to show that during these periods different kinds of currencies are quoted more equally than during the other periods. The normalized Shannon entropies computed from the number of transactions changed around the week beginning from 10th November 2008, from 22th December 2008, and from 9th March 2009. The values of quotations are more sensitive than the values of transactions since they are more volatile than ones of transactions. Since the transactions proceed out of the quotations, changes of quotations may be regarded as an indicative signs. The first peaked values of quotations around August 2008 coincides with the start of the latest global financial crisis, the second peak around the end of 2008 may correspond to infectious banking turmoil triggered by global financial crisis, the rapid decrease of February 2009 shows the termination of global financial turmoil due to rescue package of G20. Phases of the foreign exchange market seem

chen.shuheng@gmail.com

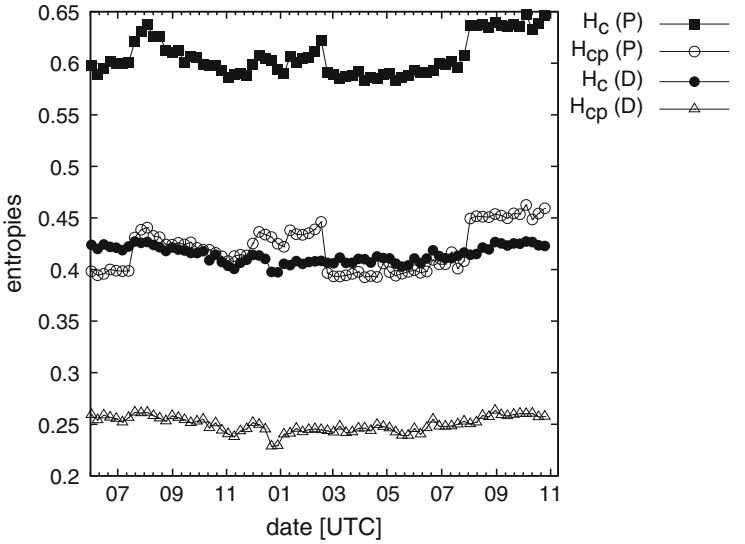


Fig. 3 The normalized Shannon entropies of quotation activities (P) and transaction occurrences (D) for currency pairs (cp) and currencies (c) for a period between June 2008 and December 2009. (H_{cp} (P)) and currencies (H_c (P)), and of transaction activities for currency pairs (H_{cp} (D)) and currencies (H_c (D))

to drastically change from August 2009 since the values rapidly increase and maintained until November 2009.

The similarity of market states between two observation periods is computed for each week. The Jensen–Shannon divergence is employed in order to compute similarities since it has resistance characteristic for zero probabilities.

Figures 4 and 5 show similarities calculated from (19) (left) and (20) (right) for the numbers of quotations, and those calculated from (19) (left) and (20) (right) for the numbers of transactions, respectively. A color on each pixel shows a magnitude of similarity of shares between 2 weeks. A black pixel shows that the distribution of quotations on a horizontal week is similar to one on a vertical week.

On the basis of quotation activities it is found that the shares of currency pairs show drastic changes around the week beginning from 21st July 2008, from 29th September 2008, from 1st December 2008, from 16th February 2009, and from 1st August 2009. Furthermore it is found that the states during a period from 14th July 2008 to 23rd February 2009 are different from other periods. This period corresponds to the period when the normalized Shannon entropies for currency pairs and currencies remained larger values than before 14th July 2008 and after 23rd February 2009 (see Figs. 4 and 5). It is confirmed that the states before 14th July 2008 and those after 23rd February 2009 were very similar from Fig. 3.

The transaction activities for currency pairs slightly differ from their quotation activities. From Fig. 5 the situations of the week beginning from 8th September 2008, the weeks from 15th December 2008 to 29th December 2008, the weeks from

chen.shuheng@gmail.com

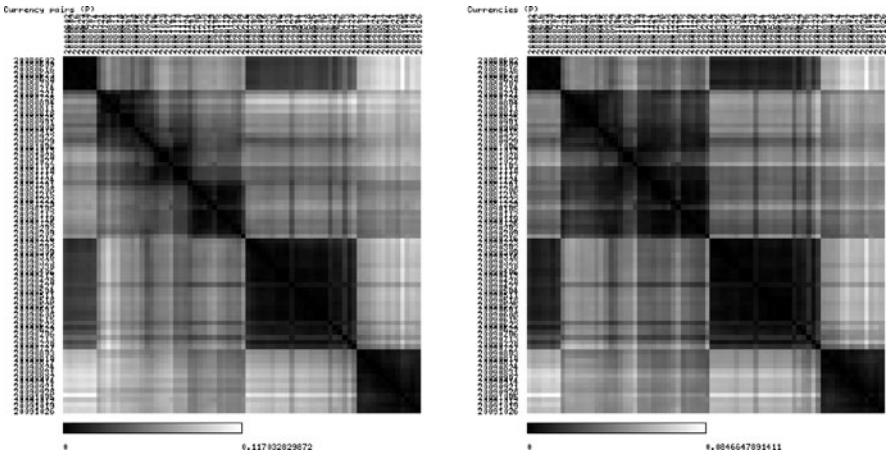


Fig. 4 The similarities of market states between observation periods based on activities of currency pairs (*left*) and of currencies (*right*) computed from quotations. The *black pixel* represents a similar relation, and the *white pixel* represents a dissimilar relation

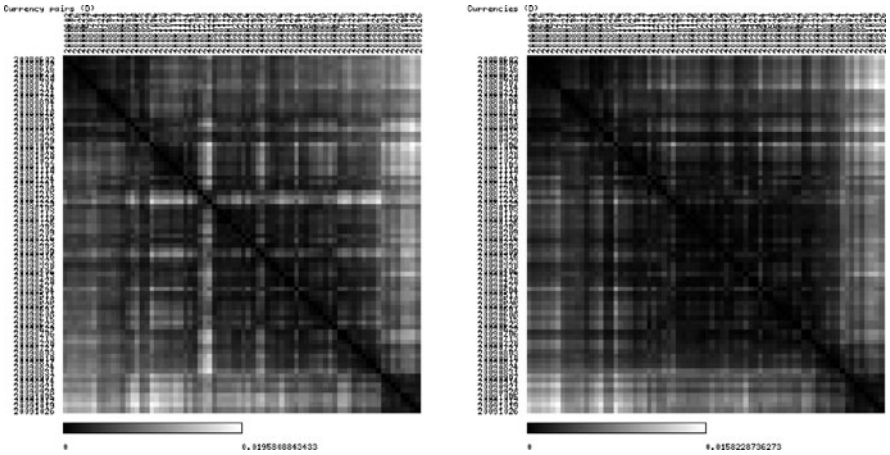


Fig. 5 The similarities of market states between observation periods based on activities of currency pairs (*left*) and of currencies (*right*) computed from transactions. The *black pixel* represents a similar relation, and the *white pixel* represents a dissimilar relation

9th March 2009 to 23rd March 2009, and the week beginning from 4th May 2009 are slightly different from other weeks. Specifically it is found that the transaction activities show drastic changes around the week beginning from 29th September 2008, from 25th December 2008, from 9th March 2009, and from 4th May 2009. The transaction activities for currencies show drastic changes at the week beginning from 6th October 2008, the week 24th November 2008, the weeks from 15th December 2008 to 29th December 2008, the week 4th May 2009, and the week

chen.shuheng@gmail.com

7th September 2009. It is found that these periods are confirmed from the normalized Shannon entropies for transaction activities on the basis of the changes of H_{cp} (D) and H_c (D) as shown in Fig. 4.

These peculiar periods detected throughout the empirical analysis on similarity of shares between pairwise weeks almost correspond to the beginning of the global financial crisis, stimuli by emergency economic package of G20 countries, and the end of the global financial crisis in March 2009, respectively.

Since the values of similarity computed from quotation activities during a period before June 2008 are not different from those during a period from May 2009 to July 2009 as shown in Fig. 4, shares of quotations seemed to be equal during these periods. However the values of similarity after August 2009 are different from ones during any periods. This may lead to that from August 2009 the foreign exchange market entered into new stage where none has experienced during a period from June 2008 to August 2009.

5 Conclusions

I proposed a method to quantify the total states of human society based on relative frequencies calculated from group activities. I conducted empirical analysis by means of the proposed method with a vast amount of data on quotations and transactions in the foreign exchange market. It was found that drastic changes of participant distributions frequently occur on the foreign exchange market during the period between July 2008 and August 2009. The uncertainty of quotation/transaction activities were measured based on the normalized Shannon entropy of the relative occurrence rates. Furthermore the method to quantify the similarity of market states between two observation periods by using their Jensen–Shannon divergence was proposed. It was confirmed that the drastic changes of the normalized Shannon entropy and the Jensen–Shannon divergence coincided with important events of the latest global financial crisis. It is concluded that the changes of quoted/traded shares of currencies in the foreign exchange market can be an important indicator of global economy.

The proposed methodology is workable if comprehensive data on human activities within groups are obtained. It can further measure states of societies whilst protecting personal privacy rights as fundamental human rights, since they do not need any personal information.

In my opinion, one of the most prominent directions of studies on human society is to develop a methodology to realize fair trade-off between public utility and personal privacy. From a comprehensive point of view we should try to tackle massive amounts of data by means of vast amounts of computations. The holistic point of view will be able to provide us with precious insights on the states of affairs which are derived from chance and necessity of our society whilst protecting our personal privacy. Such kinds of measurement systems will be part of our infrastructure in the era of information explosion.

chen.shuheng@gmail.com

Acknowledgments This work was supported by the Grant-in-Aid for Young Scientists (B) (#21760059) from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). The author would like to express his gratitude to the Kyoto University Global COE program "Informatics Education and Research Center for Knowledge-Circulating Society". Furthermore the author is thankful for valuable comments by Prof. Shinji Shimojo, Mr. Makoto Nukaga, and Prof. Hideaki Aoyama.

References

1. Takayasu H (2002) The advent of econophysics. Springer, Tokyo
2. Mantegna RN, Stanley HE (2000) An introduction to econophysics: correlations and complexity in finance. Cambridge University Press, Cambridge
3. Sornette D (2003) Why stock markets crash: critical events in complex financial systems. Princeton University Press, Princeton
4. Sato A-H, Hołyst JA (2008) Characteristic periodicities of collective behavior at the foreign exchange market. Eur Phys J B 62:373–380
5. Portal Site of Official Statistics of Japan by National Statistics Center. http://www.e-stat.go.jp
6. González MC, Hidalgo CA, Barabási A-L (2008) Understanding individual human mobility patterns. Nature 453:479–482
7. Antoniou C, Ben-Akiva M, Koutsopoulos HN (2006) Dynamic traffic demand prediction using conventional and emerging data sources. IEE Proc Intell Transp Syst 153:97–104
8. Helbing D (ed) (2007) Managing complexity: insights, concepts, applications. Springer, Berlin
9. Mizuno T, Toriyama M, Terano T, Takayasu M (2008) Pareto law of the expenditure of a person in convenience stores. Physica A 387:3931–3935
10. Lambiotte R, Ausloos M (2006) Endo-vs. exogenous shocks and relaxation rates in book and music sales. Physica A 362:485–494
11. Deschâtres F, Sornette D (2005) Dynamics of book sales: endogenous versus exogenous shocks in complex networks. Phys Rev E 72:016112
12. von Foerster H (1984) Observing systems. Intersystems Publications, Seaside, CA
13. New IT Infrastructure for the Information-explosion Era (July 2005 – March 2011) URL: http://www.infoplosion.nii.ac.jp/info-plosion/ctr.php/m/IndexEng/a/Index/
14. Korth HF (1997) Database research faces the information explosion. Commun ACM 40:139–142
15. Bateson G (1972) Steps to ecology of mind. University of Chicago Press, Chicago
16. Noh JD, Rieger H (2004) Random walks on complex networks. Phys Rev Lett 92:118701
17. Wilhelm T, Hollunder J (2007) Information theoretic description of networks. Physica A 385:385–396
18. Amari S, Nagaoka H (2000) Methods of information geometry. AMS and Oxford University Press, Oxford
19. The data is purchased from ICAP EBS. http://www.icap.com
20. Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in 2007. http://www.bis.org/publ/rpfxf07t.pdf
21. Sato A-H (2009) Detecting environmental changes through high-resolution data of financial markets. In: Nakamatsu K, Phillips-Wren G, Jain LC, Howlett RJ (ed) New advances in intelligent decision technologies: results of the first KES international symposium IDT'09 (studies in computational intelligence), pp 595–603. Springer, Berlin
22. Lin J (1991) Divergence measures based on the Shannon entropy. IEEE Trans Inf Theory 37:145

chen.shuheng@gmail.com

Examining the Effects of Traders' Overconfidence on Market Behavior

Chia-Hsuan Yeh and Chun-Yi Yang

Abstract Much attention has been paid in the past decade to how traders' psychological factors affect market properties. Overconfidence is one of most important characteristics of traders. Under an agent-based modeling framework, this paper examines how traders' overconfidence affects market properties. The preliminary results have shown that overconfidence increases market volatility, price distortion, and trading volume. Some stylized facts such as the fat-tail of the return distribution and volatility clustering would be more evident.

Keywords Rationality · Behavioral finance · Overconfidence · Artificial stock market · Agent-based modeling · Genetic programming

1 Introduction

It is well-known that modern financial economic theory relies heavily on the assumption that the representative agent in the market behaves rationally and has rational expectations. Under this assumption, it is shown that asset prices fully reflect all available information and always reflect their intrinsic value. In this situation, future price movements cannot be predicted on the basis of past information. Any financial regulation imposed on the market should generate no substantial effects but result in delayed revelation of the information. Milton Friedman is one of the strongest advocates for supporting the rational expectations approach.

Examining the efficiency of real financial markets has been an interesting topic in the past three decades. Many studies have questioned the validity of the efficient market hypothesis (EMH) in real financial markets and have provided the theoretical foundations or empirical evidence to show the existence of market inefficiency. De Long et al. [9] point out that noise traders may survive in the long run and exert an impact on price dynamics. Kogan et al. [13] further indicate that irrational traders can persistently maintain a large impact even though their relative wealth

C.-H. Yeh (✉) and C.-Y. Yang
Department of Information Management, Yuan Ze University, Chungli, Taoyuan 320, Taiwan
e-mail: imcyeh@saturn.yzu.edu.tw; ckhyde.yang@gmail.com

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_2, © Springer 2011

chen.shuheng@gmail.com

becomes quite small. Lo and MacKinlay [14], Campbell and Shiller [4], Brock et al. [3], and Neely et al. [15] all find evidence of predictability and profitability in financial markets. In addition, the increasing empirical evidence has indicated that traditional asset pricing models such as the capital asset pricing model (CAPM), arbitrage pricing theory (APT), and intertemporal capital asset pricing model are unable to provide explanations regarding the stylized facts. Financial markets usually experience several anomalies, such as event-based return predictability, short-term momentum, long-term reversal, and high volatility of asset prices relative to fundamentals where bubbles and crashes never cease. These phenomena cannot be purely explained by the changes in fundamentals. Given these findings, it is reasonable to reexamine the theory of finance based on imperfect rationality.

Actually, studying economics and finance from the perspective of imperfective rationality has both theoretical and empirical foundations. The reason for economists holding the assumption of rational expectations is that economic systems without this restriction may produce numerous outcomes so that prediction is impossible. Simon [18] argues that agents possess imperfect information or knowledge regarding the environment and that they also have limited ability in processing information. Therefore, bounded rationality is a more reasonable and more appropriate description regarding agents' behavior than perfect rationality. The empirical evidence from cognitive psychology also supports the view that agents do not behave rationally. As mentioned in [12], traders in financial markets exhibit several phenomena that deviate from perfect rationality such as overreaction toward salient news, underreaction toward less salient news, anchoring, loss aversion, mental accounting, herding, and overconfidence.

In the past two decades, research studies devoted to financial economics have considered models that deviate from full rationality. One branch focuses on the effects of noise traders, e.g. [6–9, 17]. Their findings have demonstrated that the presence of noise traders can generate substantial effects which are quite different from those observed in the market populated by rational traders alone. The other branch focuses on the consequences resulting from traders' psychological biases. This line of research has been an important issue in the field of behavioral finance. Actually, the importance of this research trend that takes the behavioral characteristics into account has been noticed. As mentioned in [10]:

> Finally, given the demonstrated ingenuity of the theory branch of finance, and given the long litany of apparent judgment biases unearthed by cognitive psychologists ([5]), it is safe to predict that we will soon see a menu of behavioral models that can be mixed and matched to explain specific anomalies. (p. 291)

DeBondt and Thaler [5] state that perhaps the most robust finding in the psychology of judgment is that people are overconfident. In [19], it is pointed out that traders' overconfidence may be due to an "anchoring and adjustment" process. The anchor has a major influence so that the adjustment is usually insufficient. Therefore, traders have tight subjective probability distributions. This phenomenon is also evidenced in the empirical literature on judgment under uncertainty. Benos [2] then believes that selection and survivorship biases may also be sources of overconfidence and successful traders usually overestimate their own contribution to their success. Such a reasoning is supported by the attribution theory, e.g. [1], which

chen.shuheng@gmail.com

describes that individuals usually attribute outcomes that support the validity of their decisions to high ability, and outcomes that are inconsistent with the decisions to external noise.

However, theoretical results rely heavily on specific assumptions regarding the characteristics of traders as well as the market environments, and the information structures. Since many factors are involved, and traders' behavior may generate externalities on others, there would be a clearer and more concrete picture regarding the effects of traders' psychological biases if a heterogeneous-agent framework were to be employed. In fact, Hirshleifer [12] mentions that:

> The great missing chapter in asset-pricing theory, I believe, is a model of the social process by which people form and transmit ideas about markets and securities. (p. 1577)

Under a well-controlled heterogeneous-agent environment where traders' psychological factors are considered, we are able to examine the market phenomena from the perspective of a micro-foundation. However, such a framework would be too complicated so that analytical results would be difficult to derive. Therefore, a simulated framework composed of many heterogeneous and bounded-rational traders whose learning behavior is appropriately represented would be a better architecture. In this paper, we provide an agent-based artificial financial market to examine the effects of traders' overconfidence on several stylized facts such as volatility clustering and fat tails for the return series.

The remainder of this paper is organized as follows. The basic framework of the model which includes the market environment, the traders' learning behavior, and the mechanism of price determination are described in Sect. 2. Section 3 presents the simulation design and the results. Section 4 concludes.

2 The Model

2.1 Market Structure

The basic framework of the artificial stock market considered in this paper is the standard asset pricing model with many heterogeneous traders. All traders are characterized by bounded rationality in which they are equipped with adaptive learning behavior represented by the genetic programming (GP) algorithm. In the framework of GP, traders are freely allowed to form various types of forecasting functions which may be fundamental-like or technical-like rules in different time periods.

Our framework is very similar to that used in [21]. However, to calibrate the model so that it is able to fit different time horizons of real financial markets, we follow the design proposed in [11].

Consider an economy with two assets. One is the risk free asset called money which is perfectly elastically supplied. Its gross return is $R = 1 + r/K$, where r is a constant interest rate per annum and K represents the trading frequency measured over 1 year. For example, $K = 1$, 12, 52, and 250 stand for the trading periods of a year, month, week, and day, respectively. The other asset is a stock with a stochastic

chen.shuheng@gmail.com

dividend process (D_t) not known to traders. The trader i's wealth at $t+1$, $W_{i,t+1}$, is given by

$$W_{i,t+1} = RW_{i,t} + (P_{t+1} + D_{t+1} - RP_t)h_{i,t}, \tag{1}$$

where P_t is the price (ex dividend) per share of the stock and $h_{i,t}$ denotes the shares of the stock held by trader i at time t. Let R_{t+1} be the excess return at $t+1$, i.e. $P_{t+1} + D_{t+1} - RP_t$, and $E_{i,t}(\cdot)$ and $V_{i,t}(\cdot)$ are the forecasts of trader i regarding the conditional expectation and variance at $t+1$ given his information up to t (the information set $I_{i,t}$), respectively. Then we have

$$E_{i,t}(W_{t+1}) = RW_{i,t} + E_{i,t}(P_{t+1} + D_{t+1} - RP_t)h_{i,t} = RW_{i,t} + E_{i,t}(R_{t+1})h_{i,t}, \tag{2}$$

$$V_{i,t}(W_{t+1}) = h_{i,t}^2 V_{i,t}(P_{t+1} + D_{t+1} - RP_t) = h_{i,t}^2 V_{i,t}(R_{t+1}), \tag{3}$$

Assume that all traders follow the same constant absolute risk aversion (CARA) utility function, i.e. $U(W_{i,t}) = -exp(-\lambda W_{i,t})$, where λ is the degree of absolute risk aversion. At the beginning of each period, each trader myopically maximizes the one-period expected utility function subject to (1). Therefore, trader i's optimal share of stock holding, $h_{i,t}^*$, solves

$$\max_h \{E_{i,t}(W_{t+1}) - \frac{\lambda}{2} V_{i,t}(W_{t+1})\}, \tag{4}$$

that is,

$$h_{i,t}^* = \frac{E_{i,t}(R_{t+1})}{\lambda V_{i,t}(R_{t+1})}. \tag{5}$$

If it is supposed that the current stock holding for trader i is at the optimal level, i.e. $h_{i,t}^* = h_{i,t}$, then the trader's reservation price, $P_i^{\mathfrak{R}}$, can be derived.

$$P_i^{\mathfrak{R}} = \frac{E_{i,t}(P_{t+1} + D_{t+1}) - \lambda h_{i,t} V_{i,t}(R_{t+1})}{R}. \tag{6}$$

2.2 Learning of Traders

According to (6), it is shown that traders' reservation prices rely on their conditional expectations and variances. We adopt the functional form for $E_{i,t}(\cdot)$:

$$E_{i,t}(P_{t+1} + D_{t+1}) = \begin{cases} (P_t + D_t)\left[1 + \theta_0 \tanh\left(\frac{\ln(1+f_{i,t})}{\omega}\right)\right] & \text{if } f_{i,t} \geq 0.0, \\ (P_t + D_t)\left[1 - \theta_0 \tanh\left(\frac{\ln(|-1+f_{i,t}|)}{\omega}\right)\right] & \text{if } f_{i,t} < 0.0, \end{cases} \tag{7}$$

where $f_{i,t}$ is evolved using GP based on $I_{i,t}$.[1]

[1] Regarding the formation of a function by means of GP as well as the implementation of GP, the reader should refer to [20].

chen.shuheng@gmail.com

The modeling of traders' conditional variances also plays an important role. Let $\sigma^2_{i,t}$ denote $V_{i,t}(R_{t+1})$. Here we consider the following form of the conditional variance:

$$\sigma^2_{i,t} = (1-\theta_1-\theta_2)\sigma^2_{i,t-1} + \theta_1(P_t + D_t - u_{t-1})^2 + \theta_2[(P_t + D_t) - E_{i,t-1}(P_t + D_t)]^2, \tag{8}$$

where

$$u_t = (1-\theta_1)u_{t-1} + \theta_1(P_t + D_t). \tag{9}$$

Traders update their own estimated conditional variance of the active rule at the end of each period.

Each trader's overconfidence level is modeled as the degree of underestimation about the conditional variance. Therefore, the conditional variance shown in (8), $\sigma^2_{i,t}$, is replaced by $\Omega^2_{i,t}$:

$$\Omega^2_{i,t} = \gamma(t)\sigma^2_{i,t}, \tag{10}$$

where

$$\gamma(t) = \begin{cases} \gamma_1, & \text{if profit} > 0, \\ \gamma_2, & \text{if profit} < 0, \\ 1, & \text{if profit} = 0, \end{cases} \tag{11}$$

and profit is defined by $W_{i,t} - W_{i,t-1}$ which measures the performance of the trader i's investment profile, the allocation of both the risky and the risk-free assets. The values of γ_1 (γ_2) should be smaller (greater) than 1, and $|1-\gamma_1| > |\gamma_2 - 1|$. The last condition is used to model the behavior of biased self-attribution.

Each trader possesses several models, say N_I, which are represented by GP. The performance of each forecasting model is indicated by the value of strength which is defined by

$$s_{i,j,t} = -\Omega^2_{i,j,t}, \tag{12}$$

where $s_{i,j,t}$ is the strength of the jth model for trader i in period t. Traders learn to make better forecasts through an adaptation process that abandons the model with the poorest performance and generates a new one by means of an evolutionary process devised in GP. The evolutionary process takes place every N_{EC} period (evolutionary cycle) for each trader asynchronously. Traders' learning works as follows. At the beginning of each evolutionary cycle, each trader randomly chooses N_T out of N_I models. The one with the highest strength value is selected as the model he uses in these periods of this evolutionary cycle. At the end of each evolutionary cycle, the model with the lowest strength is replaced by the model which is created by means of crossover, mutation, or immigration.

A simplified double auction (DA) is employed as the trading mechanism. Each period is decomposed into N_R rounds. At the beginning of each round, a new random permutation of all traders is performed to determine the order of their bid and ask.

chen.shuheng@gmail.com

Each trader, based on his own reservation price and current best bid or ask, makes a decision regarding his offer. If a bid (ask) exists, any subsequent bid (ask) must be higher (lower) than the current one. For the sake of simplicity, only a fixed amount of stock (Δh) is traded in each transaction. The last transaction price (closing price) in each period is recorded as the market price for this period.[2]

3 Simulations

One of the reasons why the results obtained in an agent-based financial market model may be convincing has to do with whether or not the model can replicate the stylized facts. The basic statistical properties of the Dow Jones Industrial Average Index (DJIA), Nasdaq Composite Index, and the S&P 500 are summarized in Table 1. The third and fourth columns show the minimum and maximum returns in percentage terms, respectively. The market volatility in terms of the average of absolute returns is described in the fifth column. The sixth column is the kurtosis K. It is evident that the kurtosis of all markets is greater than 3, which is an indication of fat tails. The tail index α which is a more reliable estimator of a fat tail is presented in the seventh column. The α value is obtained based on 5% of the largest observations. The smaller that the α value is, the fatter the tail is. The Hurst exponent is employed to examine whether a time series follows a random walk or whether it possesses underlying trends. The value of the Hurst exponent (H) lies between 0 and 1. A random series has the value of 0.5, while $0.5 < H < 1$ ($0 < H < 0.5$) implies a time series with persistence (anti-persistence). The Hurst exponents of the raw returns and absolute returns are shown in the last two columns, respectively. In Table 1, the raw return series of all markets are close to the random series. By contrast, the absolute return series exhibit strong signs of volatility clustering. These phenomena can be also observed in Fig. 1 which displays the basic properties of the Nasdaq. Figure 1 is the time series plot during 1972–2007. The distribution of the returns is presented in the third panel of Fig. 1 in which the black curve is the normal distribution with the same variance. It is clear that the return distribution of the Nasdaq possesses higher probabilities around the mean and the tails than those of the normal distribution. In addition, at the 5% significance level,

Table 1 Stylized facts of financial markets

Series	Period	$r_{\min}$	$r_{\max}$	$\lvert r\rvert$	K	α	H_r	$H_{\lvert r\rvert}$
DJIA	1972–2007	−29.22	9.21	0.72	77.03	4.25	0.53	0.96
Nasdaq	1972–2007	−12.80	12.41	0.79	13.53	3.39	0.57	0.97
S&P 500	1972–2007	−25.73	8.34	0.70	53.96	4.69	0.53	0.97

[2] For a more detailed implementation, please refer to [21].

chen.shuheng@gmail.com

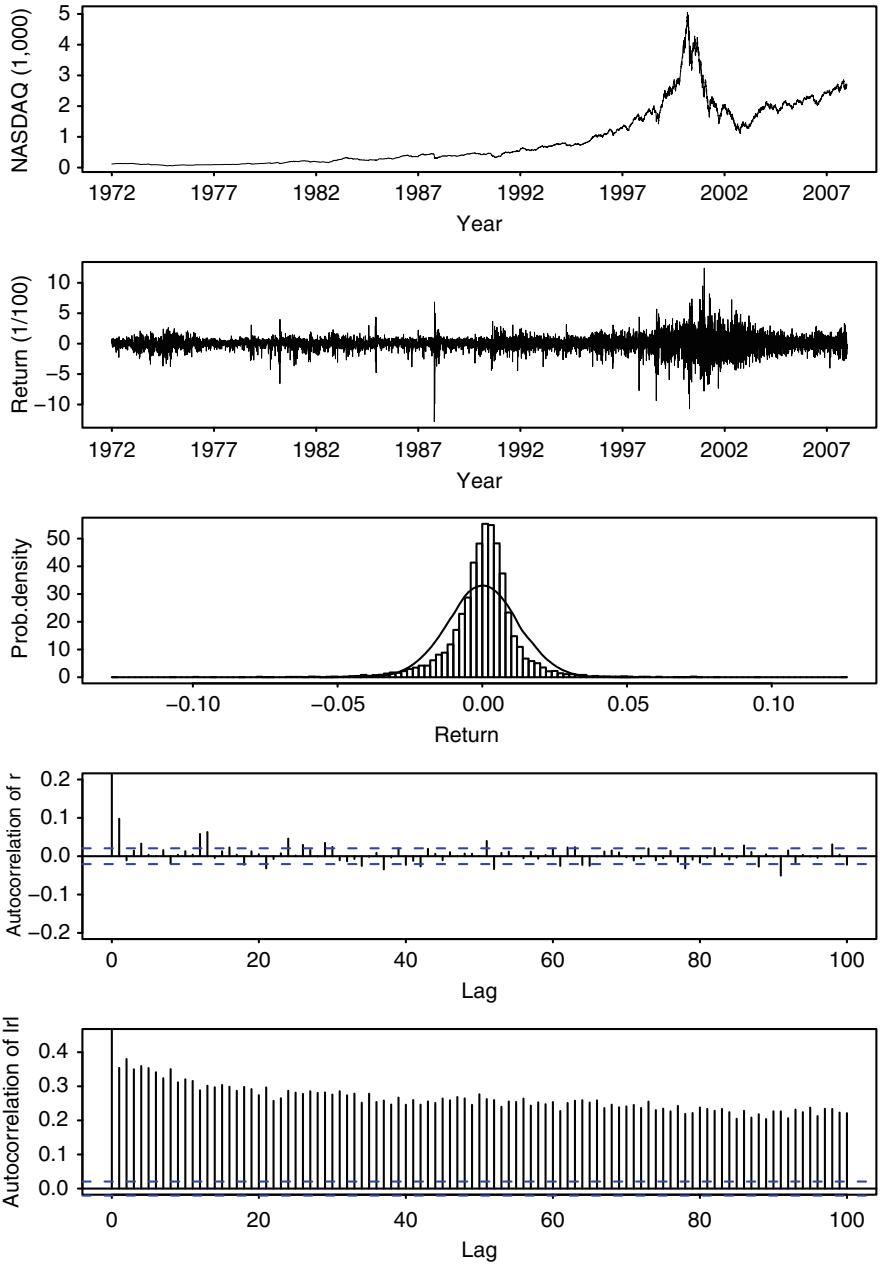


Fig. 1 Time series properties of Nasdaq

the insignificant autocorrelation features of the raw returns for most lag periods and the significant autocorrelation features of the absolute returns are exhibited in the last two panels.

chen.shuheng@gmail.com

To calibrate the model for mimicking the stylized facts of the daily data in the financial markets observed in Table 1 and Fig. 1, we adopt the same setup as [11]. The annual interest rate r is set as 5%, i.e. the daily interest rate $r_d = 0.05/250 = 0.02\,\%$. The daily dividend process is assumed to follow a normal distribution with mean $\overline{D} = 0.02$ and variance $\sigma_D^2 = 0.004$. The other parameters used in this model are shown in Table 2. Under full information and homogeneous expectations, the homogeneous REE price is given below

$$P_f = \frac{1}{R-1}(\overline{D} - \lambda \sigma_D^2 h) \tag{13}$$

where h is the average of the shares of the stock for each trader. Therefore, the fundamental price is 90.0. Short selling and buying on margin are prohibited.

Table 2 Parameters for simulations

The stock market	
Shares of the stock (h) for each trader	1
Initial money supply for each trader	\$100
Interest rate (r, r_d)	(0.05, 0.0002)
Stochastic process (D_t)	N($\overline{D}$, σ_D^2)=N(0.02, 0.004)
Amount for each trade (Δh)	1
Maximum shares of stock holding	10
Number of rounds for each period (N_R)	50
Number of periods (N_P)	20,000
Traders	
Number of traders (N)	100
Number of strategies for each trader (N_I)	20
Tournament size (N_T)	5
Evolutionary cycle (N_{EC})	5
λ	0.5
θ_0	0.5
ω	15
θ_1	0.01
θ_2	0.001
γ_1	0.99
γ_2	1.005
Parameters of genetic programming	
Function set	{`if-then-else;and,or,not;`$\geq,\leq,=$ `+,-,`$\times$`,%,sin,cos,abs,sqrt`}
Terminal set	$\{P_{t-1},\ldots,P_{t-5},D_{t-1},\ldots,D_{t-5}\}$
Selection scheme	Tournament selection
Tournament size	2
Probability of creating a tree by immigration	0.1
Probability of creating a tree by crossover	0.7
Probability of creating a tree by mutation	0.2

chen.shuheng@gmail.com

Table 3 Statistical properties of the calibrated model

	$r_{\min}$	$r_{\max}$	$\lvert r\rvert$	P_D	K	V	σ_V	α	H_r	$H_{\lvert r\rvert}$
Minimum	−13.85	11.80	0.31	21.16	18.57	157.18	18.50	1.91	0.47	0.90
Median	−20.11	17.32	0.49	39.11	49.23	168.26	21.15	3.41	0.52	0.91
Average	−23.38	20.76	0.50	40.56	45.33	167.81	21.56	3.38	0.52	0.92
Maximum	−45.11	32.10	0.78	63.12	81.11	173.70	26.13	4.96	0.57	0.94

The information set that each trader uses to form his expectations consists of the stock price and dividend history up to the last five periods.

Table 3 summarizes the basic statistical properties for 20 simulations and Fig. 2 displays the time series properties of a typical run. In comparison with the results obtained in real financial markets, our model fits these stylized facts very well. The fifth, seventh, and eighth columns of Table 3 are the price distortion (P_D), the trading volume and its standard deviation, respectively. Price distortion which measures the degree of price deviation from the fundamental price is defined as

$$P_D = \frac{100}{N_P} \sum_{t=1}^{N_P} \left| \frac{P_t - P_f}{P_f} \right| \tag{14}$$

Based on the parameters shown in Table 2, we examine the consequences of overconfident traders. Traders' overconfidence is represented by the way in which they underestimate their conditional variances. Each trader's overconfidence level is determined by two parameters, γ_1 and γ_2. In this paper, we choose $\gamma_1 = 0.99$ and $\gamma_2 = 1.005$. The results of 20 simulation runs are presented in Table 4, and the time series properties of a typical run are plotted in Fig. 3.

For most of the runs, our simulated market with overconfident traders still provides a good fit of the stylized facts. For example, the return distribution displays the property of a fat-tail. The autocorrelation of the raw return series is insignificant and that for the absolute returns is quite significant. Overconfidence makes markets exhibit richer dynamics and stronger characteristics. First, the price dynamics is more volatile and the scale of the bubble and crash is larger. There is no doubt that price distortion would be more serious. In the market without overconfident traders, the median of market volatility is 0.49%. By contrast, it is 1.90% in the market composed of overconfident traders. Second, from comparing the second panels of Figs. 2 and 3, it is clear that overconfidence results in larger return variation. In Table 3, the median of the minimum (maximum) of returns is −20.11% (17.32%) among 20 runs, while it is −47.30% (47.24%) when traders are overconfident. Overconfidence also causes more significant volatility clustering. This can be evidenced from the higher autocorrelation of absolute returns. Third, from Tables 3 and 4, we observe that trading volume as well as its volatility increase when traders are overconfident. Basically, our findings confirm the analytical results derived in [2] and [16] where they conclude that overconfidence results in increased price volatility and trading volume. However, our results are obtained based on an environment with many heterogeneous traders. Such an outcome has important implications. First, the

chen.shuheng@gmail.com

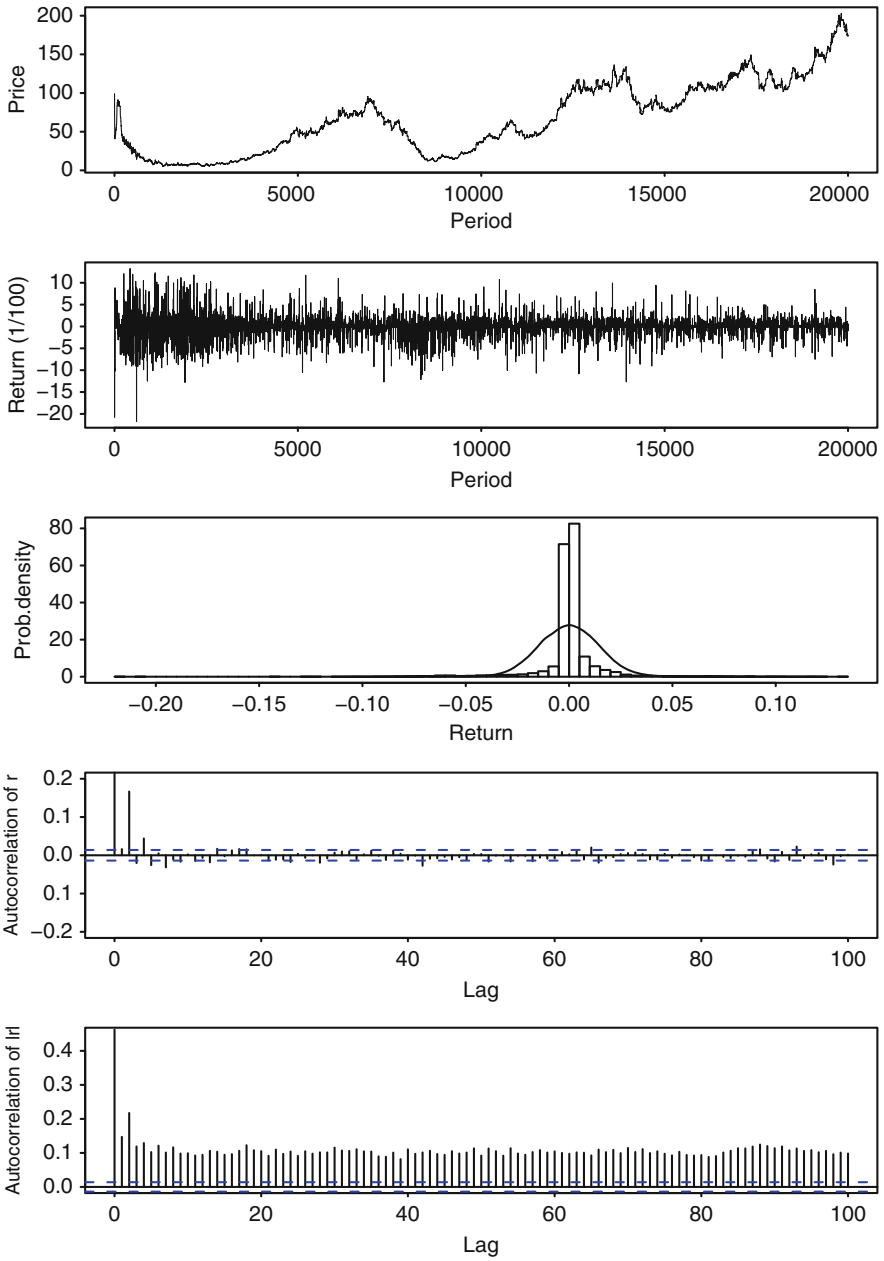


Fig. 2 Time series properties of the calibrated model

usefulness of the agent-based approach is validated. Second, we are able to examine the consequence of overconfident traders under a more realistic framework in which traders are heterogeneous in many respects.

chen.shuheng@gmail.com

Table 4 Statistical properties of the model with overconfident traders

	$r_{\min}$	$r_{\max}$	$\lvert r\rvert$	P_D	K	V	σ_V	α	H_r	$H_{\lvert r\rvert}$
Minimum	−16.08	19.23	0.73	45.19	9.47	161.15	30.36	1.32	0.50	0.82
Median	−47.30	47.24	1.90	98.26	23.54	177.11	49.73	3.68	0.56	0.95
Average	−57.95	963.36	5.97	96.28	209.09	179.81	47.52	3.60	0.59	0.93
Maximum	−99.81	5400.00	32.97	127.56	1128.09	222.91	55.00	7.10	0.77	0.97

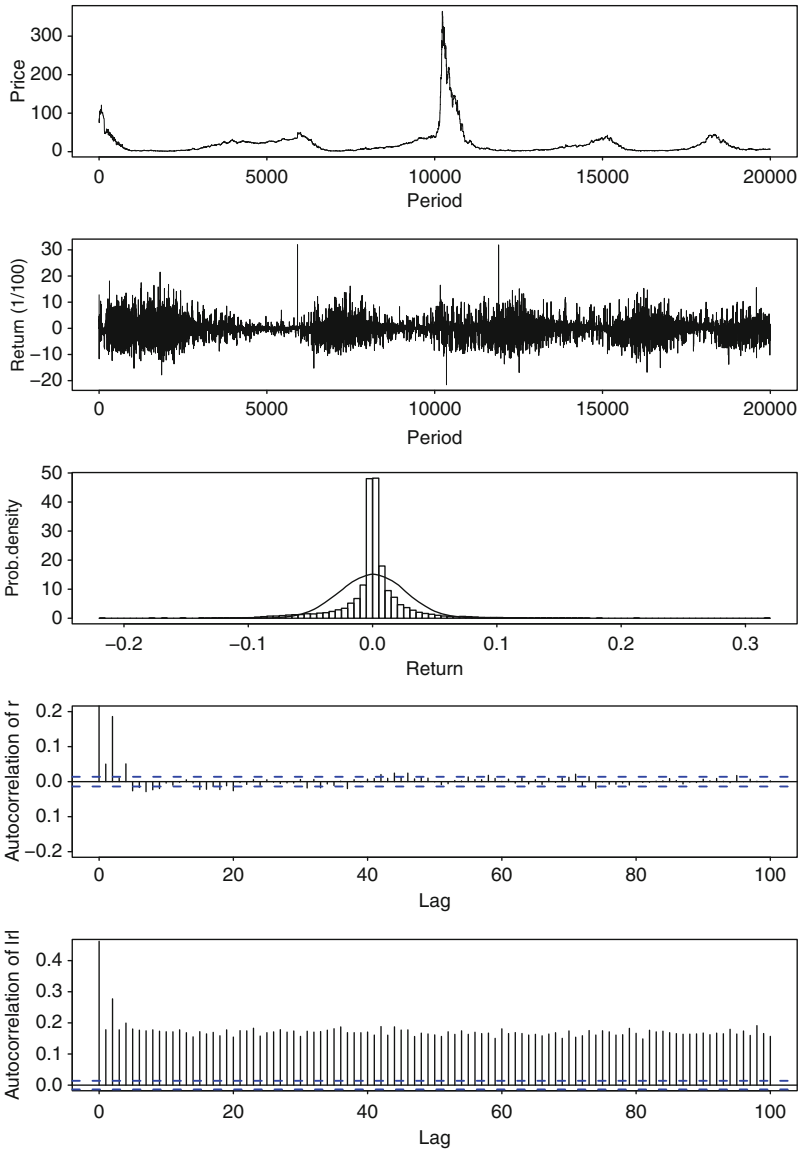


Fig. 3 Time series properties of the model with overconfident traders

chen.shuheng@gmail.com

4 Conclusion

Examining the effects of traders' overconfidence has been an important issue in behavioral finance and cognitive psychology. This paper develops an agent-based artificial financial market which consists of many heterogeneous and bounded-rational traders to examine the impacts of traders' overconfidence on market properties. Traders' learning behavior is modeled by a genetic programming algorithm. In such a framework, we show that the existence of overconfidence results in higher market volatility, price distortion, and trading volume. These results are consistent with those obtained in the theoretical framework. In addition, overconfidence also induces more significant stylized facts. Of course, our results crucially rely on the design of traders' overconfident behavior. Further investigation regarding this issue would be necessary in future studies.

Acknowledgement Research support from NSC Grant no. 98-2410-H-155-021 is gratefully acknowledged.

References

1. Bem DJ (1965) An experimental analysis of self-persuasion. J Exp Soc Psychol 1:199–218
2. Benos AV (1998) Aggressiveness and survival of overconfident traders. J Financ Mark 1:353–383
3. Brock W, Lakonishok J, LeBaron B (1992) Simple technical trading rules and the stochastic properties of stock returns. J Finance 47:1731–1764
4. Campbell JY, Shiller R (1988) The dividend-price ratio and expectations of future dividends and discount factors. Rev Financ Stud 1:195–227
5. DeBondt WFM, Thaler RH (1995) Financial decision-making in markets and firms: a behavioral perspective. In: Jarrow RA, Maksimovic V, Ziemba WT (eds) Finance, handbooks in operations research and management science, vol 9, pp 385–410. North Holland, Amsterdam
6. De Long JB, Shleifer A, Summers LH, Waldmann RJ (1989) The size and incidence of the losses from noise trading. J Finance 44:681–696
7. De Long JB, Shleifer A, Summers LH, Waldmann RJ (1990) Positive feedback investment strategies and destabilizing rational speculation. J Finance 45:379–395
8. De Long JB, Shleifer A, Summers LH, Waldmann RJ (1990) Noise trader risk in financial markets. J Polit Econ 98:703–738
9. De Long JB, Shleifer A, Summers LH, Waldmann RJ (1991) The survival of noise traders in financial markets. J Bus 64:1–19
10. Fama EF (1998) Market efficiency, long-term returns, and behavioral finance. J Financ Econ 49:283–306
11. He X-Z, Li Y (2007) Power-law behaviour, heterogeneity, and trend chasing. J Econ Dyn Control 31:3396–3426
12. Hirshleifer J (2001) Investor psychology and asset pricing. J Finance 56:1533–1597
13. Kogan L, Ross SA, Wang J, Westerfield MM (2006) The price impact and survival of irrational traders. J Finance 61:195–229
14. Lo AW, MacKinlay AC (1988) Stock prices do not follow random walks: evidence from a simple specification test. Rev Financ Stud 1:41–66
15. Neely C, Weller P, Dittmar R (1997) Is technical analysis in the foreign exchange market profitable? A genetic programming approach. J Financ Quant Anal 32:405–426

chen.shuheng@gmail.com

16. Odean T (1998) Volume, volatility, price, and profit when all traders are above average. J Finance 53:1887–1934
17. Shleifer A, Summers LH (1990) The noise trader approach to finance. J Finance 4:19–33
18. Simon HA (1957) Models of Man. Wiley, New York, NY
19. Tversky A, Kahneman D (1974) Judgement under uncertainty: heuristics and biases. Science 185:1124–1131
20. Yeh C-H (2007) The role of intelligence in time series properties. Comput Econ 30:95–123
21. Yeh C-H (2008) The effects of intelligence on price discovery and market efficiency. J Econ Behav Organ 68:613–625

chen.shuheng@gmail.com

Part II
Financial Forecasting and Investment

chen.shuheng@gmail.com

Short Time Correction to Mean Variance Analysis in an Optimized Two-Stock Portfolio

Wenjin Chen and Kwok Y. Szeto

Abstract The effect of short time correlation in stock prices on a two-stock portfolio under the framework of Mean Variance Analysis is investigated. The theory of Markowitz on portfolio management, based on a long time scale analysis of return and variance, is first optimized over a selection of pair of stocks from the Hang Seng Index and then corrected by the return of short time scales of the stocks. Several choices of short time returns, from 1 to 5 days in the past, are studied. The cumulative return is highest when the returns of the two-stock portfolio in the past 2 days are included in the correction to the "modified Sharpe ratio". The testing data cover the period between Jan 10, 2007 and July 21, 2009 for 24 blue chip stocks from the Hang Seng Index. The strategy is compared to the average return of these 24 stocks as well as to the Hang Seng Index in the same period. We conclude that our strategy has a positive return over most of the days of the testing period, including a very stable positive performance in the period of market crash. The variation of the cumulative return of our strategy is less than both the average returns of the chosen stocks or the Hang Seng Index, thereby providing a portfolio with a smaller risk but still attractive return. This strategy of time dependent mean variance analysis to include both the long and short time scale data appears to be a good investment scheme for conservative investors who prefer stable return even during market downturn.

Keywords Mean variance analysis · Conservative strategy · Portfolio management · Short term correction

1 Introduction

Proper resource allocation in the context of financial portfolio management is of continuous interest ever since the seminal work on the mean variance analysis by Markowitz [1] more than 50 years ago. His mathematical analysis for a given set of

W. Chen and K.Y. Szeto (✉)
Department of Physics, The Hong Kong University of Science and Technology,
Clear Water Bay, Hong Kong SAR, China
e-mail: wenjin@ust.hk; phszeto@ust.hk

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_3, © Springer 2011

chen.shuheng@gmail.com

financial time series produces a simple and elegant solution for resource allocation, such as in the fraction of money invested in each constituent stock in the two-stock portfolios by specifying the investment frontier and the risk tolerable by the investor. However, in real practice, one does not have a static picture of the mean nor the variance as they are always time dependent. The topic of resource allocation thus remains active and many works have been published with references to mean variance analysis [2–7]. In real application of portfolio management, the time varying nature of the mean and variance thus requires usage of time series analysis, such as pattern recognition [8, 9], genetic algorithm [10–12], neural network [13], or even fuzzy rule [14, 15]. However, it will be important to develop a general framework under which various forecasting techniques can be fruitfully applied along with the established theory of mean variance analysis. Since the original theory of Markowitz assumes a known mean, variance, and covariance of the constituent stocks in the multi-stock portfolio. In practice, these quantities can be computed using past data, although the values obtained should be updated on a daily basis. If we consider the inherent time-dependent nature of these quantities, we see that they can be approximated by some constants if we consider them in a sufficiently long time scale, so that the trend of various quantities are basically unaffected by short term fluctuation. This point of view on the importance of long term behavior in resource allocation is in an acute contrast to the point of view on time series forecasting, where the predictive power of a forecast relies heavily on an intelligent data-mining algorithm, applied not on the long or medium term data, but on the news and fluctuation of the market in the past few days. It will be desirable to bring together these two points of view, so that we have a general platform to construct a resource allocation algorithm, with the definition of the long time scale and short time scale given by the user. Although this seems contradictory and difficult, we will provide in this paper a simple connection between these two points of views. Of course, in practice, what we mean by short and long time scale should be based on empirical studies on the selected stocks. The important platform we developed here is to provide a decision mechanism based on time dependent mean variance analysis that involve the two pre-defined time scales.

In a recent work in our group [16], we have investigated a multi-agent system of stock traders, each making a two-stock portfolio using the mean-variance analysis. The results of this work show that there exists portfolio with low risk and high return, in spite of the random nature of the stock price and the unknown mechanism between the price variations of individual stock. Indeed, in all the works on portfolio management involving stocks, a common goal is to pursue high return, low risk and consistent performance. In this paper, the goal is to extend our previous work to take into account the time dependence on our trading strategy, so that the return is high, risk is low, and most important of all, do not incur great loss even in crash. The time dependent nature of our trading strategy also should avoid frequent transaction, as we cannot afford the great loss incurred by the transaction cost in the long run. To meet all these requirements, we design our investment strategy by taking into account the "Sharpe ratio" using mean-variance analysis in the long time scale, as well as the stock price fluctuation in the short term to make a correction factor on the long term analysis.

chen.shuheng@gmail.com

As in all decision making scheme, we introduce a parameter G which value we can compute and use for the trading decision, such that when G is above a certain critical value, certain action such as selling all stocks will be triggered. In this paper, we set this critical value for G artificially high to avoid frequent transactions that accumulate large loss in the long run, thereby providing a low risk portfolio even during the financial tsunami in 2008. Our investment strategy gives good performance (better than the Hang Seng Index) in a test period between 2007 and 2009, covering both the bull and the bear market, avoiding crash and taking advantage of the recovery. In Sect. 2, we review the investment strategy of two-stock portfolio in the context of Markowitz theory. The optimal resource allocation is time dependent. In Sects. 3 and 4, short term related trading mechanism is introduced. Numerical results are summarized and discussed in Sect. 5.

2 Mean-Variance Analysis for the Long Time Scale

We first consider the resource allocation problem of a portfolio consisting of two stocks and cash. Let's denote the expected return $U(t)$ and variance $Var(t)$ by

$$U(t) = \frac{1}{Sample\ Size} \sum_{k=t-Sample\ Size+1}^{t} r(k) \tag{1}$$

$$Var(t) = \frac{1}{Sample\ Size - 1} \sum_{k=t-Sample\ Size+1}^{t} (r(k) - U(t))^2 \tag{2}$$

where $r(t) = \frac{p(t)-p(t-1)}{p(t-1)}$ is the daily rate of return and $p(t)$ is the daily closing price of the stock. The sample size is chosen to be 50 days, which we consider to be sufficiently long so that the mean and variance are rather smooth function of time. In our study of a two-stock portfolio, the expected return and variance for stock pair (1,2) are given by

$$U_{12}(t,x) = U_1(t)x(t) + U_2(t)y(t) \tag{3}$$

$$Var_{12}(t,x) = Var_1(t)x^2(t) + Var_2(t)y^2(t) + 2Cov_{12}x(t)\cdot y(t) \tag{4}$$

where x and y are the fraction of the portfolio invested in stock 1 and in stock 2, respectively. Note that the constraint $x+y=1$, with $x,y \in (0,1)$ implies that these quantities are function of t and x only. The covariance Cov_{12} of the two stocks is defined as

$$Cov_{12}(t) = \frac{1}{Sample\ Size} \sum_{k=t-Sample\ Size+1}^{t} (r_1(k) - U_1(t))(r_2(k) - U_2(t)) \tag{5}$$

while the standard deviation of the two-stock portfolio is expressed as:

$$\sigma_{12}(t,x) = \sqrt{Var_{12}(t,x)} \tag{6}$$

chen.shuheng@gmail.com

To analyze this two-stock portfolio, we make use of a version of the Sharpe ratio defined as:

$$F_{12}(t,x) = U_{12}(t,x)/\sigma_{12}(t,x) \tag{7}$$

Note that this ratio is a function of x, so that we can find its maximum in the range of $x \in (0,1)$. In order to achieve maximum return per unit fluctuation or risk, we maximize $F_{12}(t,x)$ with respect to x and denote this maximum value as $F_{12}^{*}(t)$ and the corresponding resource allocation at $(x(t),y(t)) = (x^*, y^* = 1 - x^*)$. Note that this resource allocation of the portfolio refers to time t and stock pair (1,2). One may use exhaustive search with preset precision to obtain this value of $(x^*, 1-x^*)$ where the maximum of $F_{12}(t,x)$ occurs. In real application, one should use some efficient search algorithm to obtain this time dependent optimal resource allocation value $\left\{x^*{}_{ij}(t) | i = 1,\ldots N, j > i\right\}$ for all possible pair of stocks.

3 Short Term Correction

In the mean-variance analysis described above, we have chosen the sample size of 50 days, corresponding to a rather long period for stock price prediction. In general, the evaluation of the mean and variance is more reliable when the sample size is larger. However, the fluctuation on the short time scale usually has a bigger effect on the stock price on the day of prediction, than the choice of the period that defines the long time scale. To account for this factor, we consider effect of price fluctuation in a 2-day period just before the day of prediction. As this kind of short term price analysis is absent in the above mean-variance analysis, we propose to consider the following renormalization scheme for the time dependent mean variance analysis for long time scale in Sect. 2. We start with our resource allocation vector $(x(t),y(t)) = (x^*,y^*)$ obtained in the long time scale (50 days), and introduce a 2 days correction factor as follow:

$$r_{12}(t-1) = r_1(t-1)x^*(t) + r_2(t-1)y^*(t) \tag{8}$$

$$r_{12}(t-2) = r_1(t-2)x^*(t) + r_2(t-2)y^*(t) \tag{9}$$

where r_{12} is the estimated profit of the two-stock portfolio for the stock pair (1,2) using the proportion (x^*,y^*) determined by long term mean variance analysis at time t. We can generalize and incorporate short term correction factor for n-days, $r_{12}(t-i) = r_1(t-i)x^*(t) + r_2(t-i)y^*(t)$, $i = 1,\ldots,n$. For now, we will use only 2 days correction, $r_{12}(t-1)$ and $r_{12}(t-2)$, as reference to reflect the short term trend of the stock price, assuming that during these past 2 days, the portfolio allocation remains *unchanged* at (x^*,y^*). (In our numerical work, we have considered a series of short time scale correction factor, ranging from 1 to 5 days.) This short term correction poses the following decision making process for the trading strategy:

1. Should we change the selected choice of stock pair (1,2) to some other pair of stock?

chen.shuheng@gmail.com

2. What is the new portfolio allocation, in case of original pair of stock, or the new pair?

We now discuss the decision making process based on the combination of the long term mean-variance analysis and the short term price trend.

4 Decision Mechanism

On each trading day, we first calculate for each stock pair (i,j) its best initial "Sharpe ratio", denoted by $F_{ij}^{*}(t)$. This calculation also gives the corresponding resource allocation vector $v_{ij}^{*}(t) = \left(x_{ij}^{*}(t), y_{ij}^{*}(t) = 1 - x_{ij}^{*}(t)\right)$. Next, we compute for each stock pair its estimated profits for the past 2 days: $r_{ij}(t-1)$ and $r_{ij}(t-2)$ given in (8) and (9). Now, we hypothesize that the short term yield produces a multiplicative correction to the original "Sharpe ratio" $F_{ij}^{*}(t)$ to yield a "renormalized Sharpe ratio" $s_{ij}(t)$ defined by

$$s_{ij}(t) = F_{ij}^{*}(t) \exp\left\{\prod_{k=1}^{n}\left(\frac{r_{ij}(t-k)}{U_{ij}^{*}(t)}\right)\right\}, \quad n=2 \tag{10}$$

where $U_{ij}^{*}(t) \equiv U_{ij}\left(t, x_{ij}^{*}(t)\right)$. The factor inside the exponential in (10) can be considered as a logarithmic correction to the ratio $s_{ij}(t)/F_{ij}^{*}(t)$ by the price trend in the short term of the past 2 days. The price trend $[r_{ij}(t-1), r_{ij}(t-2)]$ are themselves normalized by the factor $U_{ij}^{*}(t)$, which is the long term expected return of the stock pair (i,j) provided by mean variance analysis. This renormalized "Sharpe Ratio" incorporates both a long term trend of 50 days and a short term correction of 2 days. In principle, one should consider various combinations of the "long term" and "short term" effects on each of the chosen pair of stocks. Our formulation can be readily generalized to include different length of the two time scales once we define what is long term and short term. We leave this investigation to another paper.

Now a given set of N important stocks provides $M = N(N-1)/2$ distinct pairs. Thus, we can calculate the renormalized Sharpe ratio $s_{ij}(t)$ for a given pair (i,j) on a given trading day t, and obtain a set $\boldsymbol{K}$ of M values of $s_{ij}(t)$. For M sufficiently large, these M values form a distribution function $f(s,t)$ of $s_{ij}(t)$. Let the mean and standard deviation of this $f(s,t)$ distribution be $\bar{s}(t)$ and $\sigma_s(t)$. Let the maximum value of the M members in the set $\boldsymbol{K}$ be

$$G(a,b,t,x_{ab}^{*}) = \max\left(s_{ij}(t) \,\middle|\, i = 1,\ldots,N; j < i,\right) \tag{11}$$

Here (a,b) is the pair of stocks that achieve the maximum renormalized Sharpe ratio, which is denoted by G. Note that this G is given by (11) for the stock pair (a,b) and the corresponding portfolio allocation vector $v_{ab}^{*}(t) = \left(x_{ab}^{*}(t), y_{ab}^{*}(t)\right)$.

chen.shuheng@gmail.com

The decision criterion is now given by the critical value with a threshold factor θ

$$G_c(t) \equiv \bar{s}(t) + \theta \sigma_s(t); \quad \theta = 3 \tag{12}$$

- When $G\left(a,b,t,x^*_{ab}\right) > G_c(t)$, meaning that the switch to the new pair of stocks (a,b) is likely to give a higher "Sharpe ratio", we will change our portfolio from the original to the new pair (a,b). The resource allocation is also set by the new value $v^*_{ab}(t)$. Of course, if the stock pair (a,b) is the same as the original stock pair, no action is required. If we do not have any stock, we will use all our cash to buy the stock pair (a,b) with $x^*_{ab}(t)$ for stock a and $\left(1 - x^*_{ab}(t)\right)$ for stock b.
- On the other hand, if $G\left(a,b,t,x^*_{ab}\right) < G_c(t)$, it should be safer to keep cash and we will sell *all* the stocks we have, or in the case that we only have cash, we only keep cash and no action is required.

In this trading strategy, we have reduced the number of transaction substantially, thereby avoiding the transaction cost that can reduce the return of the portfolio greatly in the long run. Trading action is required only when the optimum "Sharpe ratio" $G\left(a,b,t,x^*_{ab}\right)$ is greater than the critical value, $G_c(t)$. When we begin our investment with a given cash reserve, we only buy stock and form a portfolio with stock pair (a,b), which has a high return and low risk. As soon as the optimum value $G\left(a,b,t,x^*_{ab}\right)$ falls below the critical value $G_c(t)$, we will keep cash. Therefore, we expect our trading strategy to yield high profit with low risk some time. When there is no such opportunity, cash is preferred. Thus, our strategy is generally more conservative than many stock portfolios that keep cash less often.

5 Simulation Result

To perform numerical test of our theory, we select $24(=N)$ stocks that make up the Hang Seng Index. These stocks were all in the Hang Seng Index during the period between Jan 10, 2007 and July 21, 2009 (Table 1). With these 24 stocks we have a collection of $M = 276$ distinct pairs of stocks that can be the candidate of the optimum two-stock portfolio in the context of mean-variance analysis. The initial condition for the simulation is that the portfolio contains only cash. The transaction cost is 0.1% of the stock price for each selling or buying. We numerically calculate the resource allocation of our money on optimum stock pair over the period of Jan 10, 2007 and July 21, 2009, which consists of 620 days for possible trading.

Table 1 The stocks whose real price data are used in the study

0001.HK	0002.HK	0003.HK	0004.HK	0005.HK	0006.HK	0011.HK	0012.HK
0013.HK	0016.HK	0019.HK	0023.HK	0066.HK	0101.HK	0144.HK	0267.HK
0291.HK	0293.HK	0330.HK	0494.HK	0762.HK	0883.HK	0941.HK	1199.HK

chen.shuheng@gmail.com

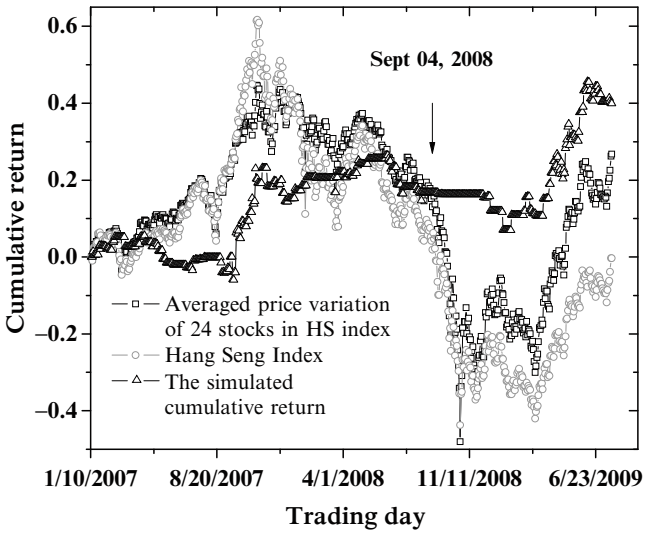


Fig. 1 The cumulative return by the proposed trading strategy in comparison with the stock average price variation and the Hang Seng Index over the same period

To have a benchmark test, we also compute the cumulative return for each stock as well as the simple averaged stock price for the 24 stocks,

$$c_i(t) = \sum_{k=0}^{t} r_i(k); \quad \bar{c}(t) = \frac{1}{N}\sum_{i=1}^{N} c_i(t) \tag{13}$$

Here $c_i(t)$ is the price for stock i at time t divided by the stock price on the first day. By averaging the stock profit variation in (13), one gets the curve denoted by open squares in Fig. 1.

We also include the Hang Seng Index (which consist other stocks, besides these 24 stocks) in the testing period (open circle). The portfolio cumulative value is calculate by the product of initial portfolio value (set as 1) and the ratio of each day's closing price divided by the closing price of the day before that day, so that at time t, the portfolio cumulative value is simply the ratio of the portfolio value at time t to initial portfolio value. The portfolio cumulative value is shown in Fig. 1 by the open triangles.

From Fig. 1, we see that before Sept 04, 2008, the cumulative return, overcoming the 0.1% transaction cost, does not deviate much from the averaged stock price variation. The return is less than a portfolio that tracks the Hang Seng Index, since we tend to keep cash unless the stock market rises rather sharply. Nevertheless, the return is still impressive, making about 30% increase in mid 2008 before the tsunami. The shape of the blue curve looks similar to the Hang Seng Index, though with generally less fluctuation. This implies that the proposed strategy can predict the portfolio with above-average return at least in most of the "good" days. Indeed, if our strategy has no intelligence, for example, in the case of a random transaction,

chen.shuheng@gmail.com

the transaction cost will incur great lost (each day when trading happens, one loses 0.1%, and over a period of 620 days, the loss can be really great). During the months of financial tsunami (September–October 2008), the stock price and Hang Seng Index nosedive while our portfolio strategy is more or less immune to the crash, because we mainly keep cash. Our strategy initiates a timely withdrawal from the stock market, so that large loss is avoided. From November 2008 onward, the cumulative return is growing in pace with the stock price variation, indicating that during that rising period, the strategy can keep up with an early sign of bull market and take profits by investing in pair of good stocks, at a good combination provided by mean variance analysis with two time scales. As a result, the overall performance of this strategy is better than the stock price variation as well as the Hang Seng Index, over the "bad" period. In summary, our investment strategy provides good and stable return in good times, while in the crash period, it keeps cash before the arrival of bad time due to a conservative decision mechanism using short term correction to the traditional long term mean variance analysis. Overall, it is a conservative strategy that can avoid penalty of transaction cost, avoid crash, and provide stable positive return in bull market or the period of recovery.

To further investigate the short term correction of the "Sharpe ratio" expressed in (10), we change the number of days in the short term correction to the "Sharpe ratio". With other conditions the same, the cumulative return as a function of n is shown in Fig. 2.

In Fig. 2, the short term correction length is the number of days considered in the short term correction of the "Sharpe ratio". The zero short term correction length refers to the original long term mean variance analysis. The cumulative return is peaked at 2 days correction, and will decrease with larger n. Actually, when n is large, the cumulative return eventually decreases to the long term value of 0 day correction. This is consistent to our formulation since when n is large, we eventually use the long term result and then there should be no more correction to the modified Sharpe ratio.

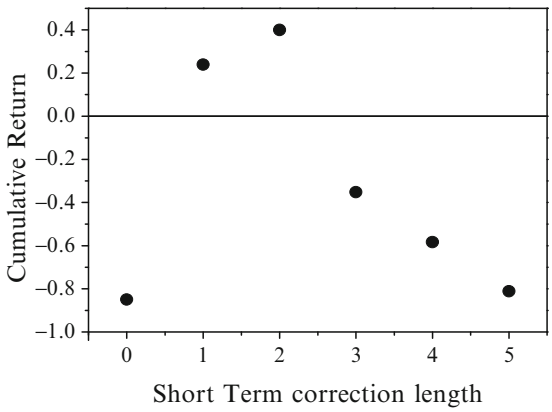


Fig. 2 Cumulative return as a function of short term correction length

chen.shuheng@gmail.com

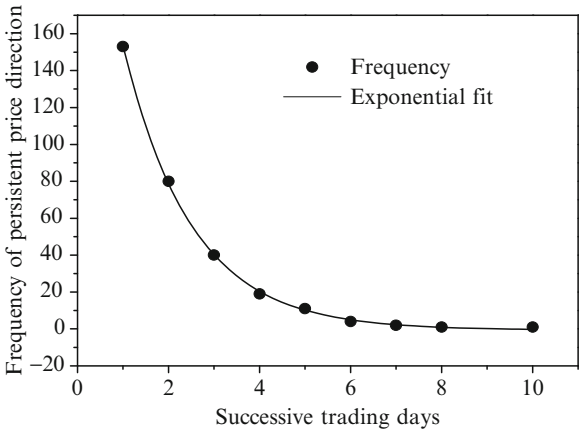


Fig. 3 Histogram of the successive trading days (n) with persistent stock price direction

We may understand the dependence of the cumulative return on the length of short term correction by means of a histogram of the successive trading days with persistent stock price direction as shown in Fig. 3 in solid circles. Here, the meaning of persistent stock price direction of n days refers to the consistency in the stock price trend over the past n days. For example, if the stock price are up in the past 2 days, then the number of occurrence of $n(=2)$ days persistent stock price direction will increase by one.

In Fig. 3, the horizontal axis represents the number of successive trading days with persistent averaged stock price direction, and vertical axis represents the frequency of occurrence of different successive trading days with persistent stock price direction. For most of the patterns of stock price in successive trading days, the averaged stock price change its direction from either upward to downward or downward to upward after 1 or 2 days. This means the stock price is auto-correlated within 1 or 2 days. By fitting the histogram curve with exponential decay as shown in Fig. 3 in solid line, we get the characteristic time interval to be 1.5 days. This means that the averaged stock price changes its direction around 1.5 days in most cases. As a result, the cumulative return is largest when we use the appropriate short term correction length. From this analysis of the duration of short term correction, we obtain an optimal short term correction to the modified Sharpe ratio. This argument also demonstrates the way we can introduce short term correction. For our present data analysis for the Hang Seng stocks in Table 1, the short term correction is optimally set to 2 days, a consequence of the highly fluctuating property of the stock price. This method of extracting the best "n" value for usage in (10) for the definition of short time scale is generally applicable to other stock markets. Once this n is defined, we can impose the proper short term correction, while the "long term" analysis is taken to be the 50 days simple moving average. Fine tuning with various moving average may improve the actual results in application, but here we focus on the general theory of connecting the short time scale correction to the long time scale mean variance analysis.

chen.shuheng@gmail.com

6 Conclusion

We proposed a rather conservative strategy of investment using the time dependent mean-variance analysis on a two-stock portfolio. The time dependence covers both the long term aspect of the pair of stocks, as well as the return of the stocks over the past 2 days. The long term aspect is covered by computing the "Sharpe ratio" at time t, while the short term stock price return provides a correction factor. The combined effect on the performance of the two-stock portfolio is a modified time dependent "Sharpe ratio". By setting a critical value for the trading threshold, we can avoid loss by placing emphasis on cash and take profit only when there is a clear indication of a pair of stocks with low risk and high return. Numerical simulation of this trading strategy with real data on a set of blue chips in the Hang Seng Index indicates good return on bull market and small loss on bear market. The overall performance of our strategy beats the performance of the chosen set of stocks as well as the Hang Seng Index. This strategy should be suitable for conservative investors.

Acknowledgements K.Y. Szeto acknowledges the support of the grant CERG 602506 and 602507.

References

1. Markowitz H (1952) J Finance 7:77
2. Horasanl M, Fidan N (2007) Portfolio selection by using time varying covariance matrices. J Econ Soc Res 9(2):1
3. Hakansson NH (1971) J Finance 26:857
4. Maccheroni F (2009) Portfolio selection with monotone mean-variance preferences. Math Finance 19:487
5. Campbell J, Viceira L (2002) Strategic asset allocation-portfolio choice for long-term investors Clarendon Lectures in Economics. Oxford University Press, Oxford
6. Schweizer M (1995) Variance optimal hedging in discrete time. Math Oper Res 20:1
7. Blanchet-Scalliet C, El Karoui N, Jeanblanc M, Martellini L (2008) J Math Econ 44:1100
8. Fukunaga K (1990) Introduction to statistical pattern recognition. Academic, San Diego
9. Zemke S (1999) Nonlinear index prediction Physica A269:177–183
10. Szeto KY, Cheung KH (1998) Multiple time series prediction using genetic algorithms optimizer. In: Proceedings of the international symposium on intelligent data engineering and learning, Hong Kong, IDEAL'98, pp 127–133
11. Szeto KY, Cheung KH (1997) Annealed genetic algorithm for multiple time series prediction. In: Proceedings of the world multiconference on systemic, cybernetics and informatics, Caracas, vol 3, pp 390–396
12. Szeto KY, Chan KO, Cheung KH (1997) Application of genetic algorithms in stock market prediction. In: Weigend AS, Abu-Mostafa Y, Refenes APN (eds) Proceedings of the fourth international conference on neural networks in the capital markets: progress in neural processing, decision technologies for financial engineering, NNCM-96, pp 95–103. World Scientific
13. Froehlinghaus T, Szeto KY (1996) Time series prediction with hierarchical radial basis function. In: Proceedings of the international conference on neural information processing, Hong Kong, 1996, ICONIP'96, vol 2, pp 799–802. Springer
14. Szeto KY, Fong LY (2000) How adaptive agents in stock market perform in the presence of random news: a genetic algorithm approach. In: Leung KS et al (eds) LNCS/LNAI, 2000, IDEAL 2000, vol 1983, pp 505–510. Springer, Heidelberg

chen.shuheng@gmail.com

15. Fong ALY, Szeto KY (2001) Rule extraction in short memory time series using genetic algorithms. Eur Phys J B 20:569–572
16. Chen C, Tang R, Szeto KY (2008) Optimized trading agents in a two-stock portfolio using mean-variance analysis. In: Proceeding of 2008 winter WEHIA & CIEF 2008, pp. 61–67

chen.shuheng@gmail.com

Exchange Rate Forecasting with Hybrid Genetic Algorithms

Jui-Fang Chang

Abstract In recent years, Artificial Intelligence (AI) methods have proven to be successful tools for forecasting in the sectors of business, finance, medical science and engineering. In this study, we employ a Genetic Algorithm (GA) to select the optimal variable weights in order to predict exchange rates; subsequently, Genetic Algorithms, Particle Swam Optimization (PSO) and Back Propagation Network (BPN) are utilized to construct three models: GA_GA, GA_PSO, GA_BPN to compare results with a traditional regression model. Fundamentally, we expect enhanced variable selection to provide improved forecasting performance. The results of our experiments indicate that the GA_GA model achieves the best forecasting performance and is highly consistent with the actual data.

Keywords Genetic algorithm (GA) · Particle Swam Optimization (PSO) · Back Propagation Network (BPN)

1 Introduction

Managers and investors rely on exchange rates as indexes of international market forces, as well as indicators of monetary policies of individual countries. However, accurate forecasting is difficult; moreover, many nations have experienced dramatic changes in the exchange values of their currencies over the past decade or so [1].

Nelly and Weller [2] adopted exchange rates such as US/DM and US/JPY with GA, GARCH and Risk Metrics model to predict the volatility of foreign currency markets. The results show that GA has better performance than GARCH and Risk Metrics, based on standard criteria of performance evaluation, mean square error (MSE), mean absolute error (MAE) and R2. Brandl et al. [3] used the genetic

J.-F. Chang (✉)
Department of International Business, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan
e-mail: rose@cc.kuas.edu.tw

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_4, © Springer 2011

chen.shuheng@gmail.com

algorithm methodology to select explanatory variables used in dynamic regression models to generate the exchange rate forecasts. Using genetic algorithms in the variable selection process is new and overcomes typical problems associated with traditional variable selection procedures. GA methods do not require assumptions about the kind of explanatory variables, the number of explanatory variables, and the length of time series.

Recently, the artificial neural network (ANN) has been successfully applied to forecasting time series in many areas such as business, finance, medical science, engineering, etc. ANN works by imitating biological neural networks; moreover, the Back Propagation Network (BPN) is one of the most representative ANNs. In 1997, Rauscher [4] used the long-term equilibrium function and macroeconomic factors of the Monetary School as input variables of BPN and produced better experimental results than the error correction model (ECM) in predicting exchange rates. Ashok and Amit [5] state that a solution obtained through standard ANN algorithms, such as BPN, suffers from serious drawbacks regarding neuron selection. A hybrid artificial intelligence method was therefore adopted and combined with the GA approach. The results indicate superior performance of the proposed method compared to the traditional non-linear time series techniques and also fixed-geometry ANN models.

Natarajan et al. [6] utilized PSO and BPN to train a neural network, and compared PSO with BPN. The results showed that PSO outperforms BPN. A model proposed by Geethanjali et al. [7] monitors and discriminates among the different operating conditions of power transformers. The PSO technique is used to train the multi-layered feed forward neural networks to discriminate among different operating conditions. These two ANNs were trained using back propagation neural network algorithm (BPN) and the PSO technique; next, the simulated results were compared. A comparison of the simulated results of the above two cases indicates that training the neural network by PSO technique gives more accurate (in terms of sum square error) and also faster (in terms of number of iterations and simulation time) results than BPN. The PSO trained ANN-based differential protection scheme provides faster, more accurate, more secure and reliable results for power transformers.

Some research has taken account of the macroeconomic factors to forecast exchange rates by certain models, including the Purchasing Power Parity [8], Monetary model [9], Interest Rate Parity, Balance of Payment, Portfolio Balance Model [1, 10].

In this paper, hybrid models GA_GA, GA_PSO, and GA_BPN have been separately utilized to forecast monthly foreign NT dollar/dollar exchange rates. The GA model is first adopted to provide the optimal variable weights. After selecting the optimal set of variables as the input factors, we use a GA, PSO and BPN model to forecast the future (t period) NT dollar/dollar exchange rate. The detailed comparisons of the proposed method with the GA_GA, GA_PSO, GA_BPN and regression model are also comprised, with criteria of MSE, MAE, and root mean square error (RMSE).

chen.shuheng@gmail.com

2 Methodology and Data Description

In this paper, the combinations of GA, PSO, BPN were used to forecast the NT dollar/US dollar exchange rate. In the first step, the GA started with a random initial population; likewise, the chromosomal representation was established. The GA gives an optimal set of beta values when achieving the stopping condition. We determined the top ten variables by sorting the beta, and achieved similar or superior forecast performance when the *beta* became an input factor.

2.1 Genetic Algorithms

Genetic Algorithms (GA) [11] work on a population of potential solutions in the form of chromosomes, attempting to locate the best solution through the process of artificial evolution. It consists of the following repeated artificial genetic operations: evaluation, selection, crossover, and mutation.

We use GA to formulate operations in the basic GA algorithm, as follows:

1. Encoding procedure: the parameters of the solution are denoted as

$$\beta = (\beta_1, \beta_2, \ldots, \beta_n), lb_j \leq \beta_j \leq ub_j, j = 1 \ldots .n \tag{1}$$

   where *lb* and *ub* indicate the lower and upper bounds for each parameter, this allows each parameter β_j to be transformed by function $L(x)$ to be an integer y_j satisfying its boundaries, which can be encoded as n*D bits of binary code.

$$y_j = L(x_j),\ c_j = u_1^j u_2^j \ldots u_{D_j}^j,\ j = 1 \ldots D,\ u_D^j \in \{0,1\} \tag{2}$$

   A chromosome is defined as a binary string combining binary codes of all parameters. Specifically, there are 27*D bits in a chromosome.
2. Random crossover and mutation: In this study, we set the mutation rate to be 0.05, and the crossover rate to be 0.5. The mutation and cross-over operations occur in random order. They exchange the parts of chromosome to inverse the selected one bit of chromosome randomly. This procedure would then produce the fittest chromosomes from the population that are selected based on the objective function value, and so the chromosomes are then allowed to reproduce.
3. Selection: The Elitist selection is used in this research study. It will sort the fitness function, and then keep the better population as the selection range. It is a selection strategy where a limited number of individuals with the best fitness values are chosen to pass to the next generation, and the crossover and mutation operators are avoided. Elitism prevents the random destruction by crossover or mutation operators of individuals with good genetics. The number of elite individuals should not be too high, otherwise the population will tend to degenerate.

chen.shuheng@gmail.com

4. Parameters selection of GA

- *Population size*: 200
- *Crossover rate*: 0.5
- *Stopping condition*: 0.02
- *Total generation*: 2,000
- *Mutation rate*: 0.05

2.2 Particle Swarm Optimization

PSO is an evolutional algorithm that is used to find optimal solutions [12]. The primary forecasting model of PSO is shown in (3).

$$\overset{\wedge}{St} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_{D-1} x_{D-1} \tag{3}$$

The PSO algorithm explores the optimal target through local history and global communication in populations that consist of potential solutions, defined as particles. For achieving a high efficiency of searching the optimal solution in high dimensions, PSO drives the particle swarm to move toward the higher-object-value region based on the best experiences of each particle and the entire population. Thus, PSO converges to the optimal point significantly faster than evolutionary optimization. Note that each particle indicates a candidate solution; if we have D-dimensional parameters in solution, the i-th particle for the t-th iteration can be represented as

$$B_i^t = \left(\beta_{i,0}^t, \beta_{i,1}^t, \beta_{i,2}^t, \beta_{i,3}^t, \beta_{i,4}^t \ldots, \beta_{i,D-1}^t\right) \tag{4}$$

Assume that the best previous position of the i-th particle at the t-th iteration is represented as

$$P_i^t = (p_{i,1}^t, p_{i,2}^t, \ldots, p_{i,D}^t) \tag{5}$$

Then, the values of cost function $MAPE(.)$ are expected to decrease by iterations.

$$Min \quad MAPE\,(.) = \frac{\sum_1^n \left| (\overset{\wedge}{St} - St)/St \right|}{n} \tag{6}$$

Where $\varepsilon\,(.)$ is the $\left|(\overset{\wedge}{St} - St)\right|$, to minimize the forecasting errors and is adopted as the fitness function.

$$LB \le B_i^t \le UB \tag{7}$$

$$\varepsilon\,(P_i^t) \le \varepsilon\left(P_i^{t-1}\right) \le \ldots \le \varepsilon\left(P_i^1\right) \tag{8}$$

chen.shuheng@gmail.com

The velocity of the i-th particle at the t-th iteration, V_i^t can be represented as

$$V_i^t = \left(v_{i,1}^t, v_{i,2}^t, \ldots, v_{i,D}^t\right) \tag{9}$$

The best position among all particles is G^t, until the t-th iteration,

$$G^t = \arg\min_{p_i} \left(\varepsilon\left(p_{i,j}^t\right)\right) \tag{10}$$

which is denoted as $G^t = (g_1^t, g_2^t, \ldots, g_D^t)$.

The modified version of PSO algorithms has been applied by an inertia weight W^t multiplying the velocity term in numerous other applications. This version can be expressed as:

$$\begin{aligned} V_i^{t+1} &= W^t \cdot V_i^t + c_1 \cdot r_1 \cdot \left(P_i^t - B_i^t\right) + c_2 \cdot r_2 \cdot \left(G^t - B_i^t\right) \\ r_1, r_2 &\sim U(0,1) \end{aligned} \tag{11}$$

$$B_i^{t+1} = B_i^t + V_i^{t+1}, i = 0, 1, \ldots, N-1 \tag{12}$$

where the updating velocity is bounded, and r1 and r2 are random variables with uniform distributions. Here, c1 and c2 are factors used to control any influences from local and global velocity terms. The parameters are: Population size: 200; Iteration: 2,000; Vmax: 0.2; c1: 2; c2: 2; Initial weight: 0.9; Final weight: 0.4.

After training by PSO algorithm, we obtain the optimal solution shown below:

$$B_i^t = \left(\beta_{i,0}^t, \beta_{i,1}^t, \beta_{i,2}^t, \beta_{i,3}^t, \beta_{i,4}^t \ldots, \beta_{i,D-1}^t\right) \tag{13}$$

2.3 Back Propagation Network

The feed-forward neural network with back propagation learning is the most conventional model of ANN that updates weight and bias values. The proposed BPN has ten input layer neurons; in this study, we propose an appropriate transfer multiple-hidden-layer (at most two hidden layers), with a Hyperbolic tangent transfer function and normalized data (between -1 and 1). We also select the gradient descent with momentum back propagation as the learning function; hence, the parameters of BPN are:

- *Input layer neurons: n* (*selected from GA*)
- *Hidden layer neurons*: (5, 5)
- *Learning rule: traingdm*
- *Transfers function: TanH*
- *Hidden layer*: 2
- *Stopcriteria: error* $<$ 1*e*-5

chen.shuheng@gmail.com

Table 1 Macroeconomic forecasting variables

Variables	Description
Constant	1: Beta 0
CPI	2–4: CPI index (t_{-1})–(t_{-3}) CPI index as an independent variable
M1 and M1B	5–7: US M1 (t_{-1})–(t_{-3}) 8–10:TW M1B (t_{-1})–(t_{-3})
Commercial paper rate	11–13: Commercial paper rate (t_{-1})–(t_{-3})
Federal funds rate	14–16: Federal fund rate (t_{-1})–(t_{-3})
Balance of trade	17–19: Balance of trade (t_{-1})–(t_{-3})
Exchange rate: USD/NTD	20–22: USD/NTD (t_{-1})–(t_{-3})
Stock return	23–25: Stock return (t_{-1})–(t_{-3})
Foreign investment	26–28: Foreign investment (t_{-1})–(t_{-3})

2.4 Variables and Data Normalization

The experimental data in this paper is to collect the historic exchange rate NT dollar/US dollar, and eight macroeconomic factors. The period is from January 1997 to December 2007 and a total of 132 monthly data were accumulated. We collected three lagged times from each factor, thereby providing 27 factors for 132 months. We divide the whole data into six periods of our sliding window. Through this process, the research can obtain more regressions so that the prediction result will be more accurate.

Recently, some research has considered the macroeconomic factors as variables when forecasting exchange rates; a linear model is formed to describe the fluctuation of exchange rates. The main variables are displayed in Table 1. Each variable will be divided into three data periods, specifically $t-3$ period, $t-2$ period, and $t-1$ period, to forecast the exchange rate by considering the lagged period. Normalization is necessary because the variables have different units. In (1), data is normalized by function premnmx in Matlab with a range of -1 to 1.

$$[pn, Min\,p, Max\,p] = premnmx(p)$$
$$pn = \frac{2 \cdot (p - Min\,p)}{Max\,p - Min\,p} - 1 \tag{14}$$

Where p is the original data, and pn is the normalized data.

Since macroeconomic factors could affect exchange rates, the variables which we propose are summarized as follows:

3 Experiment Design

In this paper, we apply the GA method to select the optimal variables. This method also replaces the traditional trial and error method that determines the input factors, layer number or input neurons.

chen.shuheng@gmail.com

Step 1 Macroeconomic forecasting variables selection

We sort the weight obtained from GA, and test the top 5, top 10, top 15, top 20, and top 25 variables from the 27 variables as input factors. The smallest value obtained after repeating the experiment three times, and the optimum forecast performance of the set with ten variables have been found. The results illustrate that the set with ten variables have the best forecasting performance as shown in Table 2.

Step 2 Input variables selection

In the second stage of the experiment, the input factors have been chosen. According to (15) and (16), $f(x)$ are composed by 27 variables $(x_1, x_2, \ldots\ldots, x_{27})$ with weights of $(\beta_0, \beta_1, \beta_2, \ldots, \beta_n)$. $f(x)$ is the linear model in which we formulate the exchange rate (i.e. $f(x) = \hat{S}_t$). $\hat{S}_t$ is t period for forecasting exchange rates.

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots. + \beta_n x \tag{15}$$

$$B = \{\, \beta_0, \beta_1, \beta_2, \ldots\ldots, \beta_n \,\} \quad (n = 27, here) \tag{16}$$

$$S_t = B \cdot X + \varepsilon \tag{17}$$

where β_0 is the constant, X_n is the weight of each variables, and is the set of constant and variables, β_n is the set of constant and variables, S_t is the real exchange rate, B is the set of weights, and is the error between real and forecast exchange rate, ε is the error between real and forecast exchange rate.

Among 27 variables, we selected the top ten variables of smallest values in each data set. The final top ten variables were chosen to be input factors of our model, as shown in Table 3.

Step 3 Forecast criteria

In consideration of several earlier empirical studies, we will evaluate the forecasting results based on three criteria: MSE (18), MAE (19), and RMSE (20).

Table 2 Comparison of different variables

Variables	RMSE	MSE	MAE
5	0.7936	0.6298	0.7121
10	0.6382	0.4073	0.4946
15	0.6844	0.4683	0.6690
20	1.0971	1.2037	0.5458
25	1.2915	1.6680	0.6882

Table 3 The optimal ten variables

Variables	Variables
Exchange rate (t_{-1})	Foreign investment (t_{-1})
Exchange rate (t_{-2})	Commercial paper rate (t_{-2})
Stock return (t_{-1})	M1 (t_{-3})
Federal fund rate (t_{-1})	M1B (t_{-1})
Stock return (t_{-2})	Stock return (t_{-3})

chen.shuheng@gmail.com

1. Mean square error,

$$MSE = \frac{\sum \left(\hat{S}_t - S_t \right)^2}{n} \tag{18}$$

2. Mean absolute error,

$$MAE = \frac{\sum \left| \hat{S}_t - S_t \right|}{n} \tag{19}$$

3. Root mean square error,

$$RMSE = \sqrt{\frac{\sum \left(\hat{S}_t - S_t \right)^2}{n}} \tag{20}$$

where S_t is real exchange rate, and $\hat{S}_t$ is the forecast exchange rate. GA, PSO, BPN and Regression model are used to forecast exchange rates.

4 Experiment Results

4.1 Comparing the Forecasting Results by Three Criteria

To test which selected variables will have better forecast performance, we randomly select ten variables for GA to select. The out-of-sample forecasting performance of the GA_GA model, the GA_PSO model, GA_BPN and the regression model are illustrated in Tables 4–7, respectively.

As shown in Table 4, the average forecasting performance of the GA_GA model, which is measured with the average RMSE, is 0.5147. The average of MSE is 0.2991, and the average MAE is around 0.4040.

As shown in Table 5, the average forecasting performance of the GA_PSO model, which is measured with the average of RMSE, is 2.1188. The average of MSE is 1.0423, and the average of MAE is around 0.9690. The results (in all RMSE,

Table 4 Forecasting results of GA_GA model

Sliding window 1–6	RMSE	MSE	MAE
GA_GA 1	0.8187	0.6702	0.6818
GA_GA 2	0.6720	0.4515	0.5268
GA_GA 3	0.3870	0.1498	0.2789
GA_GA 4	0.2526	0.0638	0.1607
GA_GA 5	0.4636	0.2150	0.3403
GA_GA 6	0.4941	0.2441	0.4352
Average	0.5147	0.2991	0.4040

chen.shuheng@gmail.com

Table 5 Forecasting results of GA_PSO model

Sliding window 1–6	RMSE	MSE	MAE
GA_PSO 1	0.9711	0.9430	0.8606
GA_PSO 2	0.5542	0.3071	0.4599
GA_PSO 3	0.3148	0.0991	0.2689
GA_PSO 4	0.3765	0.1417	0.3221
GA_PSO 5	0.7795	0.6076	0.6713
GA_PSO 6	3.2580	10.6145	3.2314
Average	1.0423	2.1188	0.9690

Table 6 Forecasting results of GA_BPN model

Sliding window 1–6	RMSE	MSE	MAE
GA_BPN 1	0.5503	0.3028	0.3858
GA_BPN 2	0.8918	0.7952	0.7372
GA_BPN 3	0.6978	0.4869	0.6505
GA_BPN 4	0.6363	0.4049	0.5469
GA_BPN 5	0.9744	0.9494	0.9248
GA_BPN 6	0.8460	0.7156	0.7940
Average	0.7661	0.6092	0.6732

Table 7 Forecasting results of Regression model

Sliding window 1–6	RMSE	MSE	MAE
Regression 1	0.5338	0.4179	0.5029
Regression 2	0.7158	0.4997	0.6383
Regression 3	0.5299	0.4169	0.5026
Regression 4	0.5882	0.4354	0.5592
Regression 5	0.5508	0.4227	0.5081
Regression 6	0.9617	0.7155	0.8733
Average	0.6467	0.4847	0.5974

MSE, MAE) of the GA_GA model are lower than the GA_PSO model with no outlier values. This indicates that the GA_GA model offers better performance than the GA_PSO model in forecasting exchange rates.

As presented in Table 6, the forecasting performance of GA_BPN, which is measured with the average of RMSE is 0.7661, and the average of MSE is 0.6092. The average of MAE is 0.6732. The results (in all RMSE, MSE, MAE) of GA_GA model are lower than the GA_BPN model, with no outlier values. This indicates that the GA_GA model has better performance than the GA_BPN model in forecasting exchange rates.

As presented in Table 7, the forecasting performance of the Regression model, which is measured with the average of RMSE is 0.6467, and the average of MSE is 0.4847. Moreover, the average of MAE is 0.5974. The results (in all RMSE, MSE, MAE) of the GA_GA model are lower than the Regression model with no outlier values. This indicates that the GA_GA model has better performance than the Regression model in forecasting exchange rates.

chen.shuheng@gmail.com

4.2 Comparison of the Forecast Results by Monthly Exchange Rate

There are 60-months of training data and 12-months of testing data in this study. The testing data will have a 6-month overlap with the training data. For instance, the training data start from January 1997, which is the 1st data to the 60th data of December 2001; and the testing data start from the 55th collected in data June 2001 to the 66th data collected in May 2002. The monthly forecast results are shown in Table 8.

In Fig. 1, the dotted line with the open square represents the forecast exchange rate by GA_GA, and the dotted line with the cross represents the forecast exchange rate by GA_PSO. The line with the triangle represents the forecast exchange rate by GA_BPN. The line with the star represents the forecast exchange rate by regression model, and the line with diamond represents the real exchange rate. It is evident

Table 8 Comparison of the forecasting results by monthly exchange rate

Item	Real exchange rate	GA_GA	GA_BPN	GA_PSO	Regression model
Month1	0.0208	−0.1330	−0.2029	−0.5135	−0.2831
Month2	0.0581	0.0919	0.0515	−0.4627	−0.1064
Month3	0.1357	0.1029	−0.0788	−0.4621	−0.1460
Month4	0.1564	0.2700	−0.0614	−0.4705	−0.0873
Month5	0.1030	0.0161	−0.1769	−0.5575	−0.2394
Month6	0.1146	0.2387	−0.3657	−0.3703	−0.1658
Month7	0.1846	0.0140	−0.1678	−0.3297	−0.1612
Month8	0.0448	0.1999	−0.1744	−0.2269	−0.0671
Month9	0.0134	0.4179	0.0076	−0.3275	0.0327
Month10	0.0603	0.0644	0.0299	−0.3582	−0.0880
Month11	−0.0152	0.0918	−0.2719	−0.4708	−0.2170
Month12	−0.1329	−0.0573	−0.3098	−0.5307	−0.2993

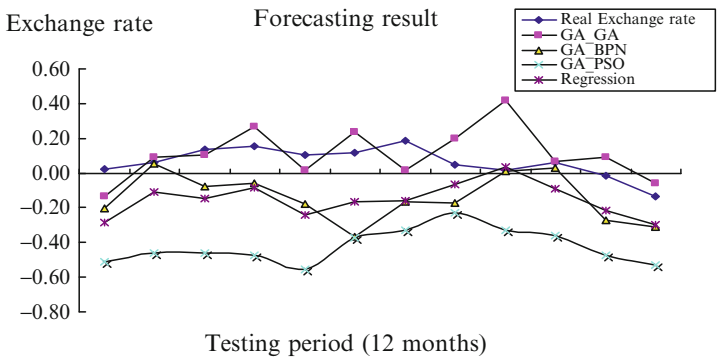


Fig. 1 Comparison of the forecasting results by monthly exchange rate

chen.shuheng@gmail.com

that the forecast exchange rates with GA_GA model have only a slight disparity with the real exchange rate, with the exception of Months 2, 7 and 9. Based on these figures, we observe that the forecasting exchange rate of GA_GA has the least error in relation to the actual exchange rate. Therefore, the GA_GA model has the best forecast performance of these three models, followed by the Regression model, then the GA_BPN model, and the results of the GA_PSO model produced the least accurate results.

5 Conclusion and Suggestions

In this study, we proposed the GA_GA, GA_PSO, GA_BPN models and regression model, which are applied to the forecasting of the NT dollar/US dollar exchange rate. In addition to the measurement by RMSE, MSE, and MAE, we also compared the forecasting results based on monthly exchange rates. The results show that GA_GA has superior forecasting ability. Moreover, we found that its forecast of NT dollar/USD dollar exchange rates is virtually identical to the real exchange rates in this study's experiments. For instance, the GA_GA (represented by a square sign) is matched to the trend of real exchange rates (diamond sign). Therefore, regarding the forecast performance of these three models, the GA_GA model has superior forecast performance, followed by the Regression model, then the GA_BPN model, and the results of the GA_PSO model were last. Additionally, future studies could adopt the nonlinear model in GA; likewise, future researchers may compare the results by applying other optimal methods, (i.e. Ant algorithm, or Cat algorithm, etc). Also, we suggest applying these models to other currencies over a longer period of time.

References

1. Branson WH (1997) Assets markets and relative prices in exchange rate determination. Sozialassenschaftliche Ann 68–89
2. Nelly CJ, Weller PA (2002) Predicting exchange rate volatility: genetic programming versus garch and risk metrics. Fed Reserv Bank St. Louis 84(3):43–54
3. Brandl B, Keber C, Schuster MG (2006) An automated econometric decision support system: forecasts for foreign exchange trades. Cent Eur J Oper Res 14(4):401–415
4. Rauscher FA (1997) Exchange rates forecasting: a neural VEC approach to non-linear time series analysis. J Time Ser Anal 461–471
5. Nag AK, Mitra A (2002) Forecasting daily foreign exchange rates using genetically optimized neural networks. J Forecast 21(7):501–511
6. Natarajan U, Periasamy VM, Saravana R (2006) Application of particle swarm optimization in artificial neural network for the prediction of tool life. Int J Adv Manuf Technol 28:1084–1088
7. Geethanjali M, Slochanal SM, Bhavani R (2008) PSO trained ANN-based differential protection scheme for power transformers. Neurocomputing 71(4–6):904–918
8. Frenkel JA (1981) Flexible exchange rates, prices, and the role of "news": lessons from the 1970s. J Polit Econ 665–703
9. MacDonald R, Taylor MP (1994) The monetary model of the exchange rate: long-run relationships, short-run dynamics and how to beat a random walk. J Int Money Finance 276–290

chen.shuheng@gmail.com

10. Bisignano J, Hoover K (1982) Some suggested improvements to a simple protfolio balance model of exchange determination with special reference to the U.S. dollar/Canadian dollar rate. Weltwirtsch Arch 19–37
11. Holland JH (1975) Adaptation in natural and artificial systems. The university of Michigan Press, Ann Arbor, MI
12. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: Sixth international symposium on micro machine and human science, pp. 39–43

chen.shuheng@gmail.com

Part III
Cognitive Modeling of Agents

chen.shuheng@gmail.com

Learning Backward Induction: A Neural Network Agent Approach

Leonidas Spiliopoulos

Abstract This paper addresses the question of whether neural networks (NNs), a realistic cognitive model of human information processing, can learn to backward induce in a two-stage game with a unique subgame-perfect Nash equilibrium. The NNs were found to predict the Nash equilibrium approximately 70% of the time in new games. Similarly to humans, the neural network agents are also found to suffer from subgame and truncation inconsistency, supporting the contention that they are appropriate models of general learning in humans. The agents were found to behave in a bounded rational manner as a result of the endogenous emergence of decision heuristics.

Keywords Agent based computational economics · Backward induction · Learning models · Behavioral game theory · Simulations · Complex adaptive systems · Artificial intelligence · Neural networks

1 Introduction

This paper investigates the potential of neural networks to learn the sophisticated concept of generalised backward induction. This is of particular interest as failures of backward induction reasoning in humans have been documented in the literature [1, 10]. Neural networks in conjunction with a backpropagation learning algorithm were chosen to model the learning agents as they are considered to be the most biologically plausible cognitive model of parallel information processing, approximating real neuronal adaptation in the human mind [4, 11–14, 17].

The universal approximation theorem [3, 5, 9] states that given a large enough finite number of neurons a single layer neural network (NN) is a universal function approximator i.e. it will be able to approximate any function, with any desired

L. Spiliopoulos (✉)
Hong Kong University of Science and Technology, Kowloon, Hong Kong
e-mail: leonidas.s@alumni.sydney.edu.au

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_5, © Springer 2011

chen.shuheng@gmail.com

level of accuracy. Modeling agents as neural networks evades the problem of ad hoc specification of the exact functional form of the learning model. Instead, they evolve endogenously according to the backpropagation algorithm that adjusts the neural networks' parameters in the direction of error minimization starting from an initial non-informative random state. However, as the backpropagation algorithm is a gradient descent approach it is possible for it to become mired in a local rather than global optimum, so that the NN agents will likely eventually learn to behave in a bounded-rational manner as is observed in human behavior.

The closest paper to this line of research is Sgroi and Zizzo [15] who employed neural networks in an effort to examine whether they can learn to play the Nash equilibrium in 3×3 strategy games with a unique pure strategy Nash equilibrium, and to observe how closely correlated these agents' behavior was to humans'. Spiliopoulos [16] has extended their work by examining neural network agents learning to play all possible strategic types of 3×3 games. Another paper [2] incorporating neural networks in game theory uses perceptrons, or simple NNs, as a way of modeling bounded rational behavior in repeated situations, such as a repeated prisoner's dilemma.[1]

2 Methodology

2.1 The Class of Games

The class of games studied in this paper, Fig. 1, involve two players, consist of two subgames and exhibit a unique pure strategy subgame-perfect Nash Equilibrium (SPNE). Games are sampled by randomly drawing each payoff from a uniform distribution with support from 0 to 100, with the proviso that the resultant games

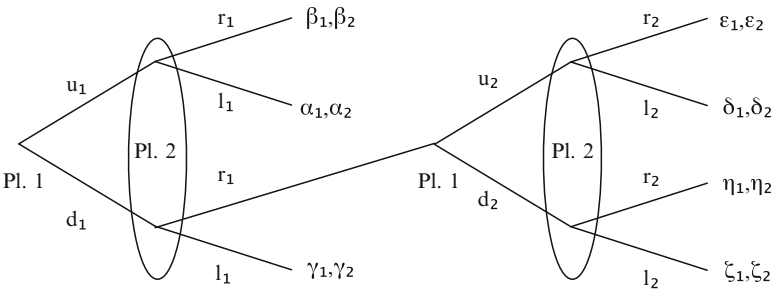


Fig. 1 Extensive form representation of the game

[1] They show that even the simplest NN, a single perceptron, is capable of implementing trigger strategies in a repeated prisoner's dilemma game and of supporting all subgame perfect payoffs. They also prove that only a slightly more complicated network is capable of supporting all equilibrium payoffs in a general 2×2 game.

chen.shuheng@gmail.com

exhibit a unique SPNE in pure strategies. There exist seven terminal nodes or outcomes overall, for ease of future exposition let $(\alpha_1, \alpha_2), (\beta_1, \beta_2), (\gamma_1, \gamma_2)$ be referred to as first stage outcomes and $(\delta_1, \delta_2), (\varepsilon_1, \varepsilon_2), (\eta_1, \eta_2), (\zeta_1, \zeta_2)$ as second stage outcomes.

The SPNE of these games is easily calculated using a generalised backward induction procedure. Identify the Nash equilibrium[2] of the final subgame, and then derive the reduced extensive form game where the final subgame is replaced with its Nash equilibrium strategies and proceed by identifying the Nash equilibrium of this game. The SPNE strategy profiles of the complete two stage game consist of the NE moves identified by this procedure at each information set i.e. the SPNE strategy profiles induce a Nash equilibrium in each subgame.

2.2 Introduction to Neural Networks

A detailed technical diagram of the topology of the feedforward neural networks employed in this study is given in Fig. 2. The first layer (leftmost in the diagram) is the input layer and each input neuron is denoted by p_r where $r = 1, \ldots, R$, and each p_r neuron inputs the payoff for a specific player from each terminal node of the game. The second layer,[3] referred to as the hidden layer, consists of S neurons, each of which is connected to all the input neurons, resulting in a total of $R \cdot S$ connections between the first and second layers. Each connection is associated with a weight, $w^{2,1}_{s,r}$, with s, r denoting a connection from the rth neuron to the sth neuron and where the superscript $2, 1$ represents that these weights are between the first and second layers of the NN. The activation of each neuron in the second layer, i^2_s, is the summation of the product of the inputs and their corresponding weights plus a constant or bias, b^2_s i.e. for each of S neurons in the second or hidden layer:

$$i^2_s = b^2_s + \sum_{r=1}^{R} w^{2,1}_{r,s} \cdot p_r \tag{1}$$

These inputs are now passed through a non-linear function, $f_1(i_s) = 2 \cdot (1 + e^{-2i^2_s})^{-1} - 1$, in this particular case the hyperbolic tangent sigmoid (or tansig) function, which maps values from $-\infty$ to $+\infty$ to the interval $(-1, 1)$. The resulting outputs, a_s, are passed to the the final or output layer which is comprised of T neurons. Each neuron in the output layer is associated with one of the terminal nodes and the output value represents the strength with which the NN associates that outcome with the subgame perfect Nash equilibrium. Again each neuron in

[2] The Nash equilibrium is the set of strategies for which neither player has an incentive to unilaterally change strategy.

[3] For simplicity, the network presented has only one hidden layer however some of the NNs in this paper will employ more than one hidden layer, each with identical structural and functional properties.

chen.shuheng@gmail.com

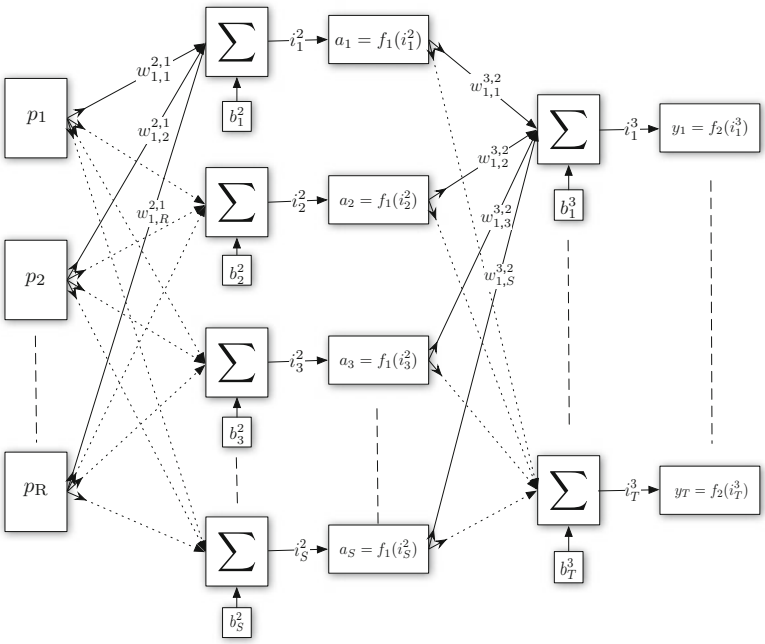


Fig. 2 Detailed structure and topology of feedforward neural networks

the output layer is connected to every neuron in the second layer with connection weights, $w^{3,2}_{s,t}$. The input to each t neuron is the summation of product of the outputs, a_s, and the corresponding weights, $w^{3,2}_{s,t}$ plus a bias b^3_t:

$$i^3_t = b^3_t + \sum_{s=1}^{S} w^{3,2}_{s,t} \cdot a_s \tag{2}$$

These inputs are transformed by the transfer function of the output layer neurons which in this case is simply a linear function f_2, such that the final outputs of the NN are defined by $y_t = f_2(i^3_t) = i^3_t$.

In processing the data, information flows forwards through the neural network, from left to right in the diagram, whereas by contrast the standard learning mechanism of a NN implicitly uses a backward flowing system, as exemplified by its name, the backpropagation rule. After making a decision the NN compares the values of the neurons at the output layer to the desired values and makes adjustments to all the connection weights in such a way that would reduce the error of the NN. The backpropagation learning rule uses the chain rule to assign the contribution of each neuron to the observed error, from which it is then possible to extract the necessary information regarding how to change each connection weight in order to reduce the overall error. Each connection weight is changed according to a

chen.shuheng@gmail.com

gradient descent method with the intent that the network successively approaches a state where the error function attains a minimum. However, this algorithm is not immune to the possibility of settling on a local rather than global minimum. A detailed mathematical presentation of the backpropagation algorithm is given in Appendix A.

2.3 Simulations, Models and Heuristics

Four different agent simulations were run using standard feedforward neural networks of different topologies, denoted by v_l where l is the number of hidden layers. For each simulation, ten different networks of the same topology were trained with different initial weights, emulating players with different initial, random weights. This will allow the examination of how important prior knowledge may be to the learning process and whether different types of players emerge in the simulation. Although the number of hidden layers of the NNs is varied across simulations, the number of neurons in each hidden layer is always ten.[4] The input layers of the NNs consist of 14 neurons, with each neuron entering each player's payoffs from each of the game's terminal nodes. The output layer consisted of seven linear transfer function neurons each one representing a terminal node and the desired output was simply a 7-tuple vector with values of one for the terminal nodes which were the SPNE and zero otherwise.[5] The terminal node chosen by the network as representing the SPNE is the one whose corresponding output neuron had the maximum value compared to the rest.

All the NN agents learned via a batch backpropagation algorithm incorporating an adaptive learning rate parameter. One generation of learning, or batch, consisted of the presentation of 1,000 different games, and the training process was terminated when the cross-validation MSE of the network's output did not improve for 100 consecutive generations. Cross validation was employed as an early stopping procedure to avoid overfitting and provide an upper bound for neural network performance. All testing was done on sets of 2,500 out of sample games in order to test the networks' ability to generalize.

To provide benchmarks against which to compare the NNs generalizing ability, six other standard classification algorithms will be estimated from the data and the performance of two simple heuristics will also be investigated. The classification algorithms are the linear, quadratic, logistic and canonical linear discriminant analyses (DA), nearest neighbor, and a classification tree, all of which were implemented in Stata 10 except for the latter which was estimated in Matlab.

[4] Simulations of neural networks with more neurons were at best not found to significantly improve performance, and often degraded performance due to overfitting.

[5] Note that this is a stricter prediction criterion than asking the neural network to simply play the NE strategy.

chen.shuheng@gmail.com

Two simple heuristics that make unique predictions for the whole set of games are simply choosing the terminal node with the maximum own payoff *ownmax*, and a heuristic that picks the terminal node with the maximum social payoff, *socialmax*.

2.4 Performance Testing

The networks' performances were documented for four different test sets as presented in Table 1. The *complete* test set is comprised of new games of exactly the same structure as the ones presented during the training session. Three other test sets that will be used are inspired by Harsanyi and Selten [8]. They prove that subgame perfection for finite games with imperfect information where each subgame has a unique Nash equilibrium is equivalent to compliance with three principles: rationality, subgame consistency and truncation consistency. Subgame consistency states that a player's behavior in a subgame does not depend on the position of the subgame within the encompassing game. Truncation consistency occurs if replacing a subgame with its equilibrium payoff does not alter players' behavior anywhere else in the game.

The *truncated* set was created by using the first stage outcomes from the games in the *complete* set, but replacing the payoffs of one of the terminal nodes in the second stage with the Nash equilibrium of the second stage, whilst setting the payoffs for all other second stage terminal nodes to zero. The *truncated/last stage* set simply moved all the non-zero payoffs from the *truncated* set games to the final subgame, whilst all payoffs in the first stage were set to zero. The *last stage* dataset is created by keeping the second stage nodes of the *complete* test set games and setting the payoffs of all the nodes in the first stage equal to zero, thereby testing whether the NN agents learned to solve for the Nash equilibrium of the final subgame, a necessary subgoal to successfully apply backward induction.

Table 1 Neural network performance in predicting the subgame-perfect NE (% correct predictions)

		Out of sample test sets				
	Training	Complete	Truncated	Truncated/last stage	Last stage	Test set mean
Feedforward v_1	80	71	76	73	74	74
Feedforward v_2	87	72	77	72	75	74
Feedforward v_3	88	74	71	73	75	73
Feedforward v_5	90	73	67	69	71	70
Linear DA	74	72	73	71	72	72
Quadratic DA	72	78	56	71	76	70
Logistic DA	74	73	70	74	75	73
Nearest neighbor	60	54	48	57	56	54
Canonical LDA	74	72	73	71	72	72
Classification tree	85	56	68	32	35	48
socialmax	–	61	70	70	66	67
ownmax	–	40	54	54	51	50

chen.shuheng@gmail.com

3 Results

3.1 Comparison of NN Topologies

Table 1 highlights some of the main results from the analyses of the various neural network agents, where the data is derived from by the performance of the ten trained neural networks in each simulation. This aggregation is justified by the finding in Sect. 3.5 that the NNs behave very similarly after training implying little impact of the initial weights.

Comparing the results between the various neural network topologies the success rate on the training set was high ranging from 80 to 90%. Also, there do not exist significant differences in the mean performance on all the test sets, as they range from 74% for the v_1 and v_2 networks to 70% for v_5. As the number of layers used in the feedforward networks increases performance on the training set also increases as expected, however performance on the tests sets is decreasing. This implies that using more than one hidden layer is not necessary and only leads to overfitting to the training sample. Given the close performance of all the neural networks in conjunction with the fact that the v_1 agents have significantly fewer parameters, all discussion henceforth will refer only to this network topology for the sake of brevity.

3.2 Comparison of NNs to Standard Classification Algorithms and Heuristics

Comparison of the v_1 agents' performance and the standard classification algorithms reveals that the NNs were the most successful in the out of sample test sets, as would be expected given their flexibility. Upon closer inspection however, four of the classification algorithms had near identical performance, the linear, quadratic, logistic and canonical discriminant analyses. The most intriguing result is that the linear DA performed very well compared to the other discriminant analyses, demonstrating that non-linearity was not important. The worst performers were the classification tree and the nearest neighbor algorithm, whose performance was significantly lower than that of the other models.

Focusing attention on the two simple heuristics, the results are particularly impressive given their simplicity. Most striking is the performance of the *socialmax* heuristic which correctly performs predicts the SPNE in 67% of all the test set games, compared to 74% for the sophisticated v_1 neural networks. The greedy *ownmax* heuristic which completely ignored opponents' payoffs still performed relatively well given its simplicity and low information requirements, with a mean test set performance of 50%. These results imply that the *socialmax* heuristic is ecologically valid for this type of problem. The intuition behind its excellent performance is that the higher the payoff of a terminal node the more likely it is to be the SPNE for a player, therefore choosing the outcome with maximum social payoffs very often

chen.shuheng@gmail.com

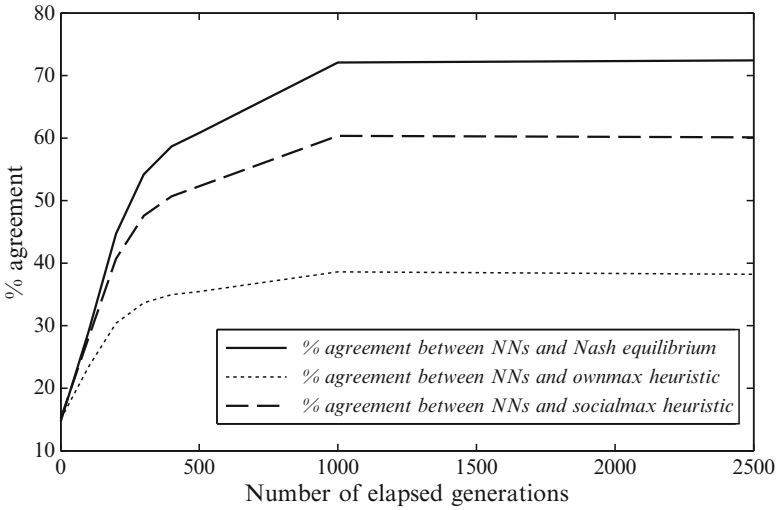


Fig. 3 Agreement between NN behavior and various other times series

coincides with the SPNE. This result agrees with the conclusions of research on fast and frugal heuristics [6,7], specifically that ecologically valid heuristics can perform as well or even better than significantly more sophisticated models of behavior.

Although the success of the *socialmax* heuristic in predicting the SPNE is high, this does not necessarily imply that this heuristic is a good approximator of the NNs' behavior. The two lower time series in Fig. 3 graph the percentage of games for which the NN agents and the two heuristics make the same predictions as a function of the elapsed generations of learning experience of the agents. The *socialmax* heuristic is a much more accurate approximator of the NN agents play as it correctly predicts their behavior at a limiting value of 60%, compared to just under 40% for the *ownmax* heuristic.

This result is of great importance because if it had arisen in an experimental study, then it would be tempting to attribute altruistic behavior to the players as they seem to care not only about their own payoffs but also those of their opponents. This conclusion however would be erroneous as there exist no altruistic incentives in the learning mechanism of the neural networks employed in this study. This highlights the danger of attributing specific types of behavior to experimental subjects only on the basis of their observable actions. In this case, it is the ecological validity of the *socialmax* heuristic which induces it to endogenously emerge in the NN agents' behavior. The possibility that the ecological validity of certain types of behavior have been misattributed to built-in incentives or motives of agents has not been addressed adequately in the literature, and is a distinction that deserves further attention.

chen.shuheng@gmail.com

3.3 Detailed Analysis of v_1 Agents' Performance

The v_1 agents predict the Nash equilibrium for 71% of the games in the complete test set. The evolution of the agents' mean performance on all the test sets is shown in Fig. 3. Learning at the beginning of the training phase is initially very fast but progressively slows down as the NNs' experience increases, culminating in a final performance above 70%. The performance for the *truncated* game set rises to 76%, as a result of directly providing the Nash equilibrium of the second stage. This implies that the NNs have not perfectly learned to solve for the NE of the final subgame, otherwise there should be no discernible change in performance. This is further validated by the performance on the *last stage* test set which directly examines the degree to which the NNs are capable of solving for the NE of the final subgame. The performance of 74% on this test set although quite good confirms that the NNs have not managed to perfectly learn the importance of backward induction.

Finally, it is interesting to note that the NNs performed almost equally well both on the *complete* test set and the *last stage* test set. Hence, the decision rules that have endogenously emerged in the NNs are equally capable of predicting the SPNE both in games with a single stage not requiring backward induction, as well as two-stage games requiring it, as long as the games exhibit a unique pure strategy SPNE.

3.4 Subgame and Truncation Consistency

Experimental tests of subgame and truncated consistency [1] reveal that human subjects often violate these principles. Subgame consistency is tested in this paper by comparing the agents' behavior in the *truncated* and *truncated/last stage* test sets. As documented in Table 2, the NNs' behavior is compatible with subgame consistency only 69% of the time, indicating that similarly to humans these agents also violate this principle. Truncation inconsistency was also found as comparing the *complete* set with the *truncated* test set led the NN agents to change their behavior in roughly 23% of the games.

It should be noted that by construct the *socialmax* and *ownmax* heuristics do not suffer from subgame inconsistency, however they are vulnerable to truncation inconsistency. Therefore despite the appeal of the *socialmax* heuristic as exhibited by its ability to approximate the NNs quite well, it does not capture the truncation inconsistency that is found both in humans and the NN agents.

In conclusion, the evidence supports that the NN agents exhibit the same inconsistencies that experimental subjects do, increasing their credibility as realistic learning models of human behavior.

Table 2 Compliance with notions of subgame and truncation consistency as % of NN agents' actions

	Subgame	Truncation
Feedforward v_1	69.24	76.84

chen.shuheng@gmail.com

Table 3 Percentage agreement between different v_1 NN agents

	Complete	Truncated	Truncated/last stage	Last stage
Average	86.76	89.53	85.96	87.15
Minimum	83.2	84.72	71.08	75.76
Maximum	90.64	95.28	92.96	93.36

3.5 Agent Heterogeneity

The ten NN agents were trained starting from different initial values, therefore an examination into the heterogeneity of their behavior after training is warranted. One possibility is that each NN may have evolved into a different type of player as a result of these different initial conditions. However the longer the training phase the less should be the dependence upon the initial conditions, unless the NNs have converged on significantly different local minima. Table 3 shows the percentage agreement of all possible pairwise comparison between the ten NNs for each of the test sets. For all the test sets the average agreement is high, ranging from 85.96 to 89.53%, and the maximum and minimum values of agreement are quite tightly distributed around this mean. In conclusion, the NN agents have evolved very similarly despite different initial conditions and show relatively little between-agents heterogeneity.

4 Conclusion

This paper has shown that applying backward induction to solve for the subgame perfect NE in two-stage games is not trivial even for the processing capabilities of a neural network. Although perfect performance is theoretically possible the NNs were found to behave in a bounded rational way in agreement with documented human behavior, likely the result of the gradient descent backpropagation method settling in a local rather than global optimum. Despite this the neural networks performed quite well, predicting the subgame perfect Nash equilibrium of new games roughly 71% of the time. Also, the NNs and experimental subjects both exhibited the same violations of backward induction, namely that of truncation and subgame consistency.

The evolved behavior of the NNs appears to have great portability and robustness as it is capable of predicting equally well in environments where backward induction is not necessary i.e. single stage games, as long as the games have a unique pure strategy Nash equilibrium.

The *socialmax* heuristic that predicted the SPNE as the outcome with the highest social payoff was found to be ecologically valid for games with a single pure strategy SPNE, as it was correct in 67% of all test games. This heuristic also explained the behavior of the NN agents quite well as it made the same prediction approximately 60% of the time.

chen.shuheng@gmail.com

The NN agents were found to perform slightly better than other classification techniques such as linear discriminant analysis in recognizing the SPNE. However, the former are preferable on two grounds. Firstly, they model the complete dynamic learning process using a plausible cognitive model of information processing of the human mind. Secondly, the exact functional form of the agents' behavior emerges endogenously as a consequence of the learning procedure rather than being postulated ad hoc.

Concluding, the neural network agents were proven to be valuable learning models as their behavior was qualitatively very similar to human behavior, whilst also being more biologically plausible than other models.

A Technical Presentation of the NN Backpropagation Algorithm

Knowledge is stored in NNs by the weights and biases of all the neurons, which is why it is referred to as distributed knowledge since it is not localized in any specific region of the NN structure. Hence, learning in a NN is accomplished through the updating of the weights and biases after presentation of each set of inputs, in this case each game's payoff matrix. In supervised learning, for each game g and each set of inputs, $P = \{p_{1,g}, \ldots, p_{R,g}\}$, there exists a set of ideal outputs, $Z = \{z_{1,g}, \ldots, z_{T,g}\}$, denoting whether the corresponding terminal node is the Nash equilibrium or not. The desired output corresponding to the terminal node which is the Nash equilibrium is encoded with a value of 1, or 0 otherwise.

In the case of batch training which is used in this paper, all the 1,000 training games are presented in a single batch to the neural network and the mean square error over all games, E, is computed according to:

$$E = \frac{1}{2} \cdot \sum_{g=1}^{G} \left(\sum_{t=1}^{T} (z_{t,g} - y_{t,g})^2 \right) \tag{3}$$

The neural network weights will then be adjusted as explained below, after which another batch (with exactly the same inputs as the previous batch) will be presented to the NN. This process is repeated until the cross validation criterion is met and training ends. The most common learning algorithm used in training NNs is the backpropagation algorithm which uses a gradient descent technique. After the presentation of each set of inputs the weights are changed according to the following equation:

$$\triangle w_{\cdot,\cdot}^{\cdot,\cdot} = -\eta \frac{\partial E}{\partial w_{\cdot,\cdot}^{\cdot,\cdot}} \tag{4}$$

This is a gradient descent technique as the updating of the weights depends on the negative of the gradient of the error function and on its magnitude, where η is

chen.shuheng@gmail.com

referred to as the step size (or learning rate) that governs the magnitude of the change in the weights. Hence, weights will be changed in the direction which reduces the error, E, and the magnitude of the change will also be related to the sensitivity of the error function to small changes in the weight. In standard backpropagation η is a constant, whereas backpropagation with an adaptive learning parameter scales η upwards if the mean square error E decreased in the previous batch presentation or scaled downwards if E increased. The necessary algebra to derive $\partial E/\partial w$ for both output layer and hidden layer neurons is presented below.

In more detail, for weights in the output layer, using the chain rule leads to the following derivation:

$$\frac{\partial E}{\partial w_{s,t}^{3,2}} = \sum_{g=1}^{G} \frac{\partial E}{\partial y_{t,g}} \frac{\partial y_{t,g}}{\partial i_t^3} \frac{\partial i_t^3}{\partial w_{s,t}^{3,2}} \tag{5}$$

However, from (2) and (3) it is clear that:

$$\frac{\partial i_t^3}{\partial w_{s,t}^{3,2}} = a_s \tag{6}$$

$$\frac{\partial E}{\partial y_{t,g}} = (y_{t,g} - z_{t,g}) \tag{7}$$

Substituting these equations into (5) and recognizing that for a linear transfer function $f_2^{'}(i_t^3) = 1$ results in:

$$\frac{\partial E}{\partial w_{s,t}^{3,2}} = \sum_{g=1}^{G} (y_{t,g} - z_{t,g}) a_s \tag{8}$$

The necessary calculations for weights in hidden layers is more involved as the desired output of such neurons is not immediately available as is the case for output layer neurons. Using the chain rule, the analog to (5) for a hidden layer neuron is:

$$\frac{\partial E}{\partial w_{s,r}^{2,1}} = \sum_{g=1}^{G} \sum_{t=1}^{T} \frac{\partial E}{\partial y_{t,g}} \frac{\partial y_{t,g}}{\partial a_s} \frac{\partial a_s}{\partial w_{s,r}^{2,1}} \tag{9}$$

This equation now has a summation of terms over t since hidden layer weights can affect the error of the NN through all the output layer neurons due to the propagation of the effect of $w_{s,r}^{2,1}$ through the interconnections between the sth neuron and all T neurons in the output layer. The derivative of the output of the sth neuron with respect to the weight under investigation is given by:

$$\frac{\partial a_s}{\partial w_{s,r}^{2,1}} = p_r f_1^{'}(i_s^2) \tag{10}$$

chen.shuheng@gmail.com

The derivative of the error function w.r.t. the output of each final layer neuron, $\partial E/\partial y_{t,g}$, is still given by (7). Finally, the derivatives of the output of each final layer neuron w.r.t. the output of each hidden layer neuron are given by:

$$\frac{\partial y_{t,g}}{\partial a_s} = w_{t,s}^{3,2} f_2^{'}(i_t^3) \tag{11}$$

In conclusion, substituting (7), (10), (11) and $f_2^{'}(i_t^3) = 1$ into (9) leads to the following equation:

$$\frac{\partial E}{\partial w_{s,r}^{2,1}} = \sum_{g=1}^{G} \sum_{t=1}^{T} p_r f_1^{'}(i_s^2)(y_{t,g} - z_{t,g}) w_{t,s}^{3,2} \tag{12}$$

References

1. Binmore K, McCarthy J, Ponti G, Samuelson L, Shaked A (2002) A backward induction experiment. J Econ Theory 104(1):48–88
2. Cho I, Sargent T (1996) Neural networks for encoding and adapting in dynamic economies. Handb Comput Econ 1:441–470
3. Cybenko G (1989) Approximation by superpositions of a sigmoidal function. Math Contr Signals Syst 2:303–314
4. Dror IE, Gallogly DP (1999) Computational analyses in cognitive neuroscience: in defense of biological implausibility. Psychon Bull Rev 6(2):173–182
5. Funahashi K (1989) On the approximate realization of continuous mappings by neural networks. Neural Netw 2:183–192
6. Gigerenzer G (2000) Adaptive thinking: rationality in the real world. Oxford University Press, New York
7. Gigerenzer G, Selten R (eds) (2001) Bounded rationality: the adaptive toolbox. MIT Press, Cambridge, MA
8. Harsanyi JC, Selten R (1988) A general theory of equilibirum selection in games. MIT Press, Cambridge, MA
9. Hornik K (1991) Approximation capabilities of multilayer feedforward networks. Neural Netw 4:251–257
10. Johnson EJ, Camerer CF, Sen S, Rymon T (2002) Detecting failures of backward induction: monitoring information search in sequential bargaining experiments. J Econ Theory 104:16–47
11. Kettner R, Marcario J, Port N (1993) A neural network model of cortical activity during reaching. J Cogn Neurosci 5:14–33
12. Lehky SR, Sejnowski TJ (1988) Network model of shape-from-shading: neural function arises from both receptive and projective fields. Nature 333:452–454
13. Mazzoni P, Andersen RA, Jordan MI (1991) A more biologically plausible learning rule than backpropagation applied to a network model of cortical area 7a. Cereb Cortex 1:293–307
14. Robinson T (2000) Biologically plausible back-propagation. Tech. rep., Victoria University of Wellington
15. Sgroi D, Zizzo D (2009) Learning to play 3×3 games: neural networks as bounded-rational players. J Econ Behav Organ 69(1):27–38
16. Spiliopoulos L (2009) Neural networks as a learning paradigm for general normal form games. SSRN eLibrary. http://ssrn.com/paper=1447968
17. Zipser D, Andersen RA (1988) A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. Nature 331:679–684

chen.shuheng@gmail.com

Cognitive-Costed Agent Model of the Microblogging Network

Mitsuhiro Nakamura and Hiroshi Deguchi

Abstract Microblogging is a new paradigm spreading in social web services that provides us a light-weight, speedy way of communication. In the microblogging system, users post short messages just as quickly as chatting. Users can easily communicate with each other. However, the system brings users huge cognitive costs since they always need to follow up their friends' posts every second. Can such cognitive costs affect the structure of the microblogging network? Here we extracted data from the most major microblogging service: *Twitter*. We find that the network structure in *Twitter* has some features: power law decay in the small degree range, link reciprocity and asymmetry between distributions of the in-degree and out-degree. To explain such characteristics, we introduce a simple stochastic agent model based on the Barabási–Albert model. With the mathematical analysis and computational experiments, we confirm that even such a simple model explains well the behavior of the observed data.

Keywords Complex network · Social web service · Microblogging · Cognitive cost

1 Introduction

Microblogging is a new paradigm spreading in social web services that provides us a light-weight, speedy way of communication. Recently, more and more services have been launching their microblogging systems, like *Twitter*,[1] *Jaiku*,[2] and so on. In these systems, users post short messages (usually less than 140 characters) as quickly as chatting via PC, mobile, or cell phone. Users can communicate and

[1] http://twitter.com.
[2] http://www.jaiku.com.

M. Nakamura (✉) and H. Deguchi
Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology,
4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan
e-mail: nakamura08@cs.dis.titech.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_6, © Springer 2011

chen.shuheng@gmail.com

share information in the blink of an eye. Furthermore, they can add the other users as their friends therefore the microblogging system is a class of social networks.

Many private companies and public organizations are making use of the microblogging service as marketing media, e.g., public relations channel and infrastructure for gathering latent customer information. For such purposes, it is important to understand the structure and function of the microblogging system. Despite that, there have been still few studies investigating microblogging systems [1, 2] (Cheng et al., 2009, unpublished report by Sysomos, inc). In this paper, we adopt a point of view of complex network and attempt to understand properties of the microblogging network. We make our analysis based on the actual data obtained from *Twitter*, which is the first and most popular microblogging service. By studying some statistical features of the data, we introduce a simple agent model which represents the growth of the network. With mathematical analysis and agent-based simulation, we show that even such a simple model can explain macroscopic properties in the observation. Finally we discuss several issues contained in the model and directions for future research.

2 Data Analysis

Microblogging is an instance of the social network and *Twitter* shows ordinary characteristics already observed in the other networks such as small-world and scale-free features [3–6]. For example, Java et al. [1] reported that *Twitter*'s social network has diameter 6 and its degree distributions follow power law with exponent $\gamma \simeq 2.4$.

Contrasting with the other social networking services such as *Mixi*,[3] *Twitter* is quite unique because of its directedness when viewed as a network. In *Twitter*'s network, like the other social networks, vertices represent users and edges represent friendships between users. Edges have directions, "following" or "followed", and thus friendships can be either one-way or two-way (see Fig. 1).

Here we analyze the data sets extracted from *Twitter*'s streaming API.[4] We monitored the data feed for two periods, during July 20–27, 2009, and during Feb 21–22, 2010. The former data set contains 1.6 million users' information (`friends_count` and `followers_count` properties) and 0.9 million users being observed in the latter one. We pay attention to the statistical properties of the network, especially degree distributions. We promise that a user's in-degree means the number of users he/she is "followed" by and out-degree indicates the number of users he/she is "following".

Figure 2 shows strong correlation between the in-degree and out-degree. Both data sets exhibit that most users have the same order of the in-degree and out-degree. This feature is caused by link reciprocity, i.e., tendency of vertex pairs to form mutual connections between each other [7]. In many social networking

[3] http://mixi.jp.

[4] http://apiwiki.twitter.com/Streaming-API-Documentation.

chen.shuheng@gmail.com

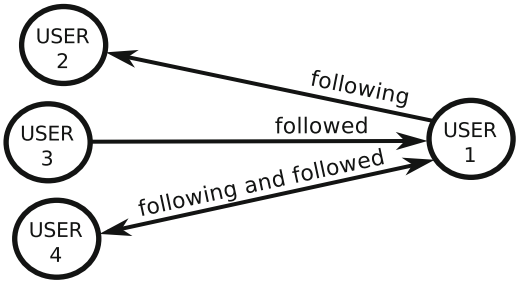


Fig. 1 Following/followed friendships in *Twitter*. To add a user as a friend is called "follow". In the figure, user 1 "follows" user 2 and user 4, while is "followed" by user 3 and user 4. Only user 1 and user 4 have a mutual friendship and the others are all one-sided. When a user follows another user, the former one follows up the latter one's updating of posts in his/her "timeline"

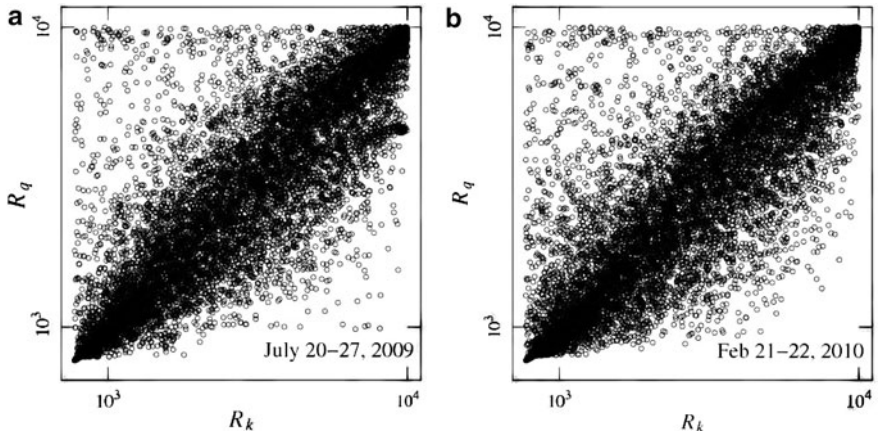


Fig. 2 All plots consist of 10^4 users chosen randomly from each data sets. (**a**) Rank scatter plot between in-degree and out-degree for the July data set. (**b**) Rank scatter plot between in-degree and out-degree for the July data set. In both of them, vertices with huge degrees have high ranks

services, creation of a link from one user to another tend to cause creation of the reverse link to be established [8–10]. Such a behavior of users is called "reciprocation".

Figure 3 displays the in-degree and out-degree distributions that follow power law (scale-free) statistics. We observed that the cumulative in-degree distributions behave $P_>(k) \propto k^{-1.2}$ in Fig. 3a, c, that is, the in-degree distributions follow $P(k) \propto k^{-2.2}$. The cumulative out-degree distributions decay discontinuously around $q \simeq 2{,}000$ because of the artificial limitation on making more than 2,000 outgoing edges [11]. While the degree distributions show scale-free behavior in the large degree range, they show power law decay in the small-degree range (around $k, q < 10^2$). We observed that $P(q)$ decays more exceedingly than $P(k)$ in the small degree range (around $k < 10^2$) in Fig. 3b, d.

chen.shuheng@gmail.com

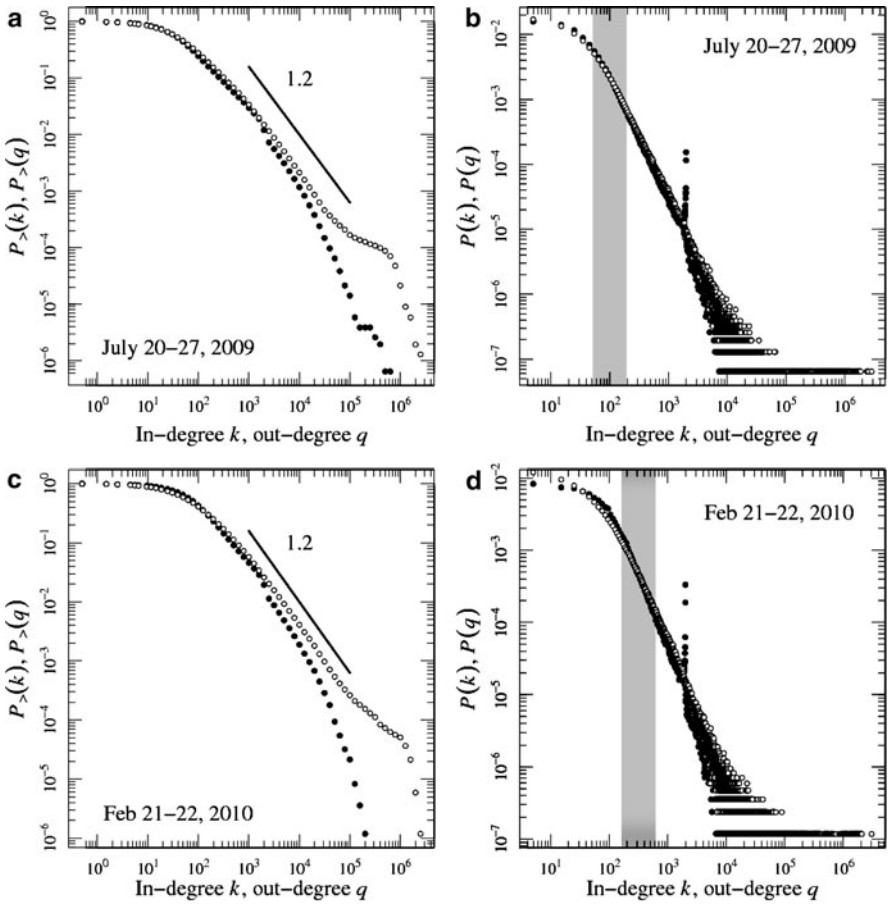


Fig. 3 Plots of the in-degree and out-degree distributions. *Blank points* show the in-degree and *filled points* display the out-degree. (**a**) Cumulative in-degree and out-degree distributions for the July data set. (**b**) In-degree and out-degree distributions for the July data set. (**c**) Cumulative in-degree and out-degree distributions for the February data set. (**d**) In-degree and out-degree distributions for the February data set. In (**b**) and (**d**), peculiar peak around in-degree $k = 2{,}000$ in the out-degree distribution is human-made [11]. For (**a**) and (**c**), the *straight solid line* shows $P_>(k) \propto k^{-1.2}$ and $P_>(q) \propto q^{-1.2}$

$P(q)$ is greater than $P(k)$ in the small degree range($k < 10^2$ in Fig. 3b and $k < 10^{2.5}$ in Fig. 3d), whereas $P(q)$ is lesser than the $P(k)$ in the large degree range ($k > 10^2$ in Fig. 3b and $k < 10^{2.5}$ in Fig. 3d). $P(k)$ and $P(q)$ closs each other in the gray area ($k \simeq 10^2$) shown in Fig. 3b. The same clossing occurs in the gray area ($k \simeq 10^{2.5}$) in Fig. 3d. What mechanism causes such phenomena?

chen.shuheng@gmail.com

3 Model

We now try to get further understanding about the phenomena reported above. To explain observed statistical behavior, basically we apply a rate equation approach to the extended Barabási–Albert model [12]. We first focus on a feature observed in the actual network: reciprocation. We also pay attention to the cognitive cost, i.e., a bias against users' active link creation brought by the system.

3.1 Simple Reciprocal Model

A growth model of the directed network which considers vertices' reciprocation was studied by Zlatić and Štefančić [13]. Here we briefly introduce such a stochastic agent model.

We start with one vertex in the network. In each time step, a new vertex i is added to the network and create an edge to an existing vertex j with probability $\Pi(k_j)$, which is proposal to j's in-degree k_j (preferential attachment). Then, the vertex j creates a return edge to the vertex i with probability ρ_j, where the constant ρ_j is determined with each vertex.

Applying continuous approximation, we get the following time evolution about the average in-degree.

$$\frac{\partial \bar{k}(s,t)}{\partial t} = \Pi(\bar{k}(s,t)) = \frac{\bar{k}(s,t)}{\int_0^t du\, \bar{k}(u,t)} = \frac{1}{1+\langle\rho\rangle}\frac{\bar{k}(s,t)}{t}\ ,\ \bar{k}(s,s) = \langle\rho\rangle\ , \quad (1)$$

where $\bar{k}(s,t)$ means the average in-degree of a vertex at any time t which is added at the time s and $\langle\rho\rangle$ denotes the average over time of ρ_i.

The solution of (1) is,

$$\bar{k}(s,t) = \langle\rho\rangle \left(\frac{s}{t}\right)^{-\frac{1}{1+\langle\rho\rangle}}\ . \quad (2)$$

Hence, at the equilibrium time, the cumulative in-degree distribution becomes,

$$P_<(k) = \Pr\left[\bar{k}(s,t) < k\right] = \Pr\left[\frac{s}{t} > \left(\frac{\langle\rho\rangle}{k}\right)^{1+\langle\rho\rangle}\right] = 1 - \left(\frac{k}{\langle\rho\rangle}\right)^{-(1+\langle\rho\rangle)}\ . \quad (3)$$

We note that $P_<(k)$ represents the probability that the in-degree is smaller than k.

Then we obtain the in-degree distribution,

$$P(k) = \frac{\partial}{\partial k} P_<(k) \propto k^{-(2+\langle\rho\rangle)}\ . \quad (4)$$

chen.shuheng@gmail.com

We also get the out-degree distribution $P(q) \propto q^{-(2+\langle\rho\rangle)}$ from the boundary condition $\bar{q}(s,s) = 1$ and the following relation,

$$\frac{\partial \bar{q}(s,t)}{\partial t} = \langle\rho\rangle \frac{\partial \bar{k}(s,t)}{\partial t} \ . \tag{5}$$

The out-degree distribution is thinner than the in-degree distribution on account of the multiplicative factor $\langle\rho\rangle$ in (5). This result is statistically consistent with the observation in Fig. 3.

3.2 Reciprocal Model with Cognitive Cost

We extend the reciprocal model described above with considering the effect of cognitive cost.

In *Twitter*, posted messages from one's "following" friends appear on his/her "timeline", which keeps updating every second (see Fig. 4). Consequently, the more the number of following users increases, the faster and the faster the timeline flows. Thus large amounts of following users result in difficulty following up friends' posts. We regard it as that the cognitive cost increases with the out-degree.

We assume that the marginal cognitive cost is diminishing. The functional form of the cognitive cost should be,

$$C(q) \propto q^{\alpha} \ , \tag{6}$$

where q is the out-degree and $\alpha \in (0,1)$ is a constant.

Consider the following situation. In each time step, a new vertex i is added to the network with probability p, then the vertex i creates a new edge to an old vertex j with probability $\Pi(k_j)$. The vertex j reciprocates the vertex i with probability ρ_j. On the other hand, with probability $\bar{p} \equiv 1 - p$, an old vertex i is chosen with

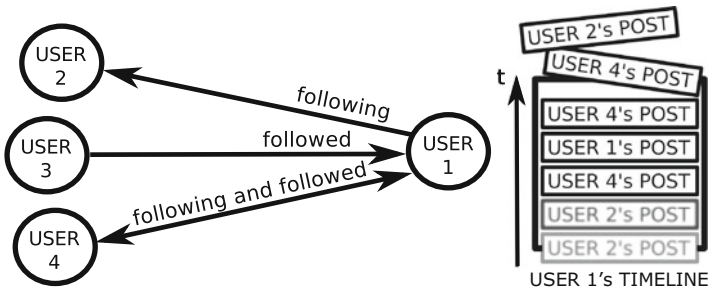


Fig. 4 The concept of "timeline". Posted messages from user 1's following users (user 2, user 4) keep real-time updating on the timeline of user 1. Updates from not-following users of user 1 (user 3) will not appear on the timeline of user 1

chen.shuheng@gmail.com

probability Ψ_i, then the vertex i creates an edge to another old vertex j in accordance with the preferential attachment rule and the vertex j reciprocates in the same manner. Let Ψ_i depend on the cognitive cost $C(q_i)$. We simply assume that $\Psi_i \propto C(q_i)^{-\beta}$ where $\beta > 0$, since vertices with smaller cost should create more out-going edges. Therefore $\Psi_i = \Psi(q_i) \propto q_i^{-\kappa}$ where $\kappa \equiv \alpha\beta > 0$.

Here we derives degree distributions in the same way in Sect. 3.1. The time evolution of the average in-degree consists of the following three cases:

1. A new vertex is added with probability p, then one gets a new edge from the new vertex with probability $\Pi(\bar{k})$.
2. One creates a new edge with probability $\bar{p}\Psi(\bar{q})$, then receives reciprocation with probability $\langle\rho\rangle$.
3. One gets a new edge from an old vertex with probability $\bar{p}\,(1-\Psi(\bar{q}))\,\Pi(\bar{k})$.

Thus,

$$\begin{aligned}\frac{\partial \bar{k}}{\partial t} &= p\Pi(\bar{k}) + \bar{p}\left[\Psi(\bar{q})\langle\rho\rangle + (1-\Psi(\bar{q}))\,\Pi(\bar{k})\right] \\ &= \Pi(\bar{k}) + \bar{p}\langle\rho\rangle\Psi(\bar{q}) - \bar{p}\Pi(\bar{k})\Psi(\bar{q})\ . \end{aligned} \tag{7}$$

In just the same way, the time evolution of the average out-degree is the summation of the following three cases:

1. A new vertex is added with probability p, then one gets a new edge from the new vertex with probability $\Pi(\bar{k})$, after that the one reciprocates with probability $\langle\rho\rangle$.
2. One creates a new edge with probability $\bar{p}\Psi(\bar{q})$.
3. One gets a new edge from an old vertex with probability $\bar{p}\,(1-\Psi(\bar{q}))\,\Pi(\bar{k})$, then reciprocates with probability $\langle\rho\rangle$.

Thus,

$$\begin{aligned}\frac{\partial \bar{q}}{\partial t} &= p\Pi(\bar{k})\langle\rho\rangle + \bar{p}\left[\Psi(\bar{q}) + (1-\Psi(\bar{q}))\,\Pi(\bar{k})\langle\rho\rangle\right] \\ &= \langle\rho\rangle\Pi(\bar{k}) + \bar{p}\Psi(\bar{q}) - \bar{p}\langle\rho\rangle\Pi(\bar{k})\Psi(\bar{q})\ . \end{aligned} \tag{8}$$

For (7) and (8), obtaining,

$$\frac{\partial \bar{q}}{\partial t} = \langle\rho\rangle\frac{\partial \bar{k}}{\partial t} + \bar{p}\left(1-\langle\rho\rangle^2\right)\Psi(\bar{q})\ . \tag{9}$$

As a consequence of (9), large enough $\langle\rho\rangle$ results in that the average in-degree and out-degree have strong correlation. Assuming such a condition, we can ignore $\Psi(\bar{q})$ in the large-degree range and thus (7) and (8) become,

$$\begin{cases} \dfrac{\partial \bar{k}}{\partial t} = & \Pi(\bar{k})\ , & \bar{k}(s,s) = \langle\rho\rangle \\ \dfrac{\partial \bar{q}}{\partial t} = & \langle\rho\rangle\Pi(\bar{k})\ , & \bar{q}(s,s) = 1 \end{cases}\ . \tag{10}$$

As the same as Sect. 3.1, we get $P(k) \propto k^{-(2+\langle\rho\rangle)}$ and $P(q) \propto q^{-(2+\langle\rho\rangle)}$.

chen.shuheng@gmail.com

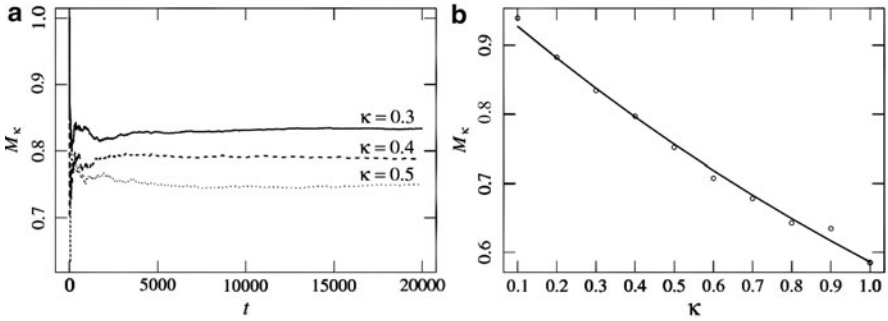


Fig. 5 (**a**) Evaluation of M_κ. At large enough time, M_κ becomes stable. (**b**) κ-dependency of M_κ. M_κ is inferred to be a form of exponential function. The *solid line* shows an exponential fit

In the small-degree range, we can ignore $\Pi(\bar{k})$. Therefore,

$$\begin{cases} \dfrac{\partial \bar{k}}{\partial t} = \bar{p}\langle\rho\rangle\Psi(\bar{q}) \,, & \bar{k}(s,s) = \langle\rho\rangle \\ \dfrac{\partial \bar{q}}{\partial t} = \bar{p}\Psi(\bar{q}) \,, & \bar{q}(s,s) = 1 \end{cases} . \tag{11}$$

$\Psi(\bar{q})$ has $\sum_j q_j^{-\kappa}$ at its denominator, whose behavior is unknown. We adopt an approximation and rewrite it as $M_\kappa pt$, where $M_\kappa \equiv \left\langle q_j^{-\kappa} \right\rangle_j$. This is supported by a numerical evaluation (see Fig. 5).

Applying this approximation to (11), we obtain,

$$P(q) = M_\kappa \frac{p}{\bar{p}} q^\kappa \exp\left[-M_\kappa \frac{p}{\bar{p}} \frac{q^{\kappa+1}-1}{\kappa+1} \right] , \tag{12}$$

$$P(k) = \frac{M_\kappa}{\langle\rho\rangle^{\kappa+1}} \frac{p}{\bar{p}} k^\kappa \exp\left[-\frac{M_\kappa}{\langle\rho\rangle^{\kappa+1}} \frac{p}{\bar{p}} \frac{k^{\kappa+1}-1}{\kappa+1} \right] . \tag{13}$$

Equations (12) and (13) do not follow the power law. Replacing $M_\kappa p/\bar{p}$ in (12) and $M_\kappa p/\langle\rho\rangle^{\kappa+1}\bar{p}$ in (13) as Θ, let us consider the behavior of the following function.

$$f(x) = \Theta x^k \exp\left[-\Theta \frac{x^{\kappa+1}-1}{\kappa+1} \right] . \tag{14}$$

Equation (14) behave as shown in Fig. 6. Small Θ (namely, small p and large κ) makes those distributions have gentle slopes. Moreover, the factor $\langle\rho\rangle^{\kappa+1}$ in (13) affects the asymmetry between the in-degree and out-degree distributions. Smaller $\langle\rho\rangle$ makes the in-degree distribution sharper.

chen.shuheng@gmail.com

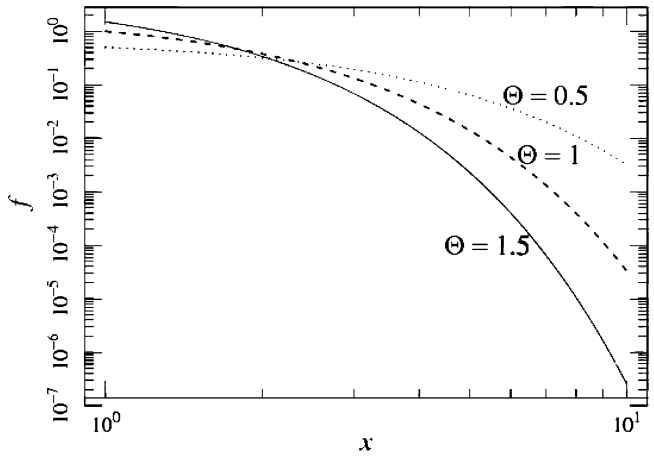


Fig. 6 The functional form of Eq. (14)

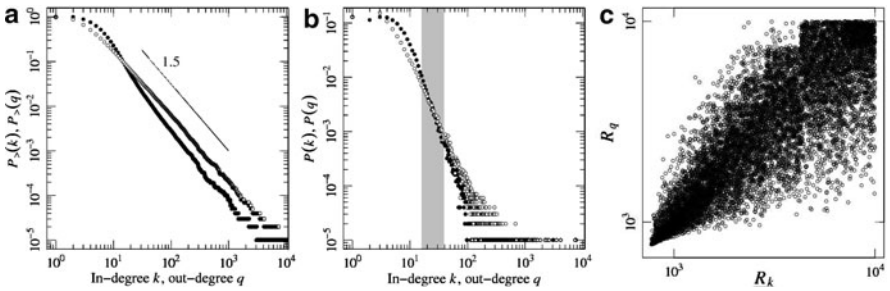


Fig. 7 Results of a numerical simulation, where $p = 0.4$, $\kappa = 0.5$, and ρ_i follows a uniform distribution in $[0, 1]$. The number of vertices is 10^5. (**a**) The cumulative in-degree and out-degree distributions. The *straight solid line* shows $P_>(k) \propto k^{-1.5}$. (**b**) The in-degree and out-degree distributions. *Blank points* represent the in-degree and *filled points* displays the out-degree in both of (**a**) and (**b**). (**c**) Rank correlation of the in/out-degree distributions

Figure 7 shows results of a numerical simulation. As the actual observation, the out-degree distribution decays more extremely than the in-degree distribution in the small degree range. Therefore the in-degree and the out-degree are asymmetric. Crossing between both degree distributions appears in Fig. 7b. In Fig. 7c, we also observe strong rank correlation between in/out-degrees.

4 Discussion and Conclusions

In summary, we have taken a look at the actual microblogging network and introduced a simple stochastic agent model containing two features: users' reciprocation and the effect of cognitive cost. Despite its simplicity, the model succeeds to

chen.shuheng@gmail.com

reproduce structural properties observed in the actual data and thus it should capture some key characteristics of the behavior of users in the real microblogging systems.

On the other hand, some assumptions in our model failed to catch the reality in the actual system. For example, we assumed that the probability of reciprocation ρ_i is constant, while it is suggested that ρ_i depends on the many factors, e.g., type of the account, out-degree, aging of activity of the user, in the real system. In future, We will enhance our model for such inconsistencies.

Microblogging is a new social web service which will be having significant importance in the next generation. Toward application in marketing, we will continue further research for understanding of such a social web system.

Acknowledgements We thank *Twitter Inc.* for providing an API and allowing us to fetch data sets from their amazing service. We also thank *Leave a nest Co., Ltd* for a research grant.

References

1. Java A, Song X, Finin T, Tseng B (2007) Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, San Jose, California, pp 56–65. ACM
2. Huberman BA, Romero DM, Wu F (2008) Social networks that matter: twitter under the microscope. SSRN eLibrary
3. Albert R, Barabási A (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47
4. Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. Adv Phys 51(4):1079–1187
5. Newman MEJ (2003) The structure and function of complex networks. SIAM Review 45(2): 167–256
6. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D (2006) Complex networks: structure and dynamics. Phys Rep 424(4–5):175–308
7. Garlaschelli D, Loffredo MI (2004) Patterns of link reciprocity in directed networks. Phys Rev Lett 93(26):268701
8. Chun H, Kwak H, Eom Y, Ahn Y, Moon S, Jeong H (2008) Comparison of online social relations in volume vs interaction: a case study of cyworld. In: Proceedings of the 8th ACM SIGCOMM conference on Internet measurement, Vouliagmeni, Greece, pp 57–70. ACM
9. Kumar R, Novak J, Tomkins A (2006) Structure and evolution of online social networks. In: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining, Philadelphia, PA, USA, pp 611–617. ACM
10. Mislove A, Koppula HS, Gummadi KP, Druschel P, Bhattacharjee B (2008) Growth of the flickr social network. In: Proceedings of the first workshop on Online social networks, Seattle, WA, USA, pp 25–30. ACM
11. Schonfeld E (2008) Twitter's 2000-follow limit raises a ruckus. But how many people can you seriously watch anyway? http://www.techcrunch.com/2008/08/12/twitters-2000-follow-limit-raises-a-ruckus-but-how-many-people-can-you-seriously-keep-track-of-anyway/
12. Barabási A, Albert R (1999) Emergence of scaling in random networks. Science 286(5439): 509–512
13. Zlatić V, Štefančić H (2009) Influence of reciprocal arcs on the degree distribution and degree correlations. Phys. Rev. E80, 016117 (arXiv: 0902.3542)

chen.shuheng@gmail.com

Part IV
Complexity and Policy Analysis

chen.shuheng@gmail.com

Landscape Analysis of Possible Outcomes

Yusuke Goto and Shingo Takahashi

Abstract The behavior of a complex social system is unpredictable because both the uncertainties and the complex interactions in the system affect its future behavior. Existing scenario analysis methods focus on the effects of complex interactions of the system upon the system's behavior, rather than the uncertainties in the system. The purpose of this paper is to develop a novel scenario analysis method that mainly focuses on evaluating a range of possible outcomes in a system based on selected uncertainties. We validate this method by applying it to a case example in which the configuration of an evaluation system for a sales division is examined.

Keywords Scenario analysis · Agent-based social simulation · Decision support technique · Visualization technique

1 Introduction

The behavior of a complex social system is, after all unpredictable, because both uncertainties and complex interactions in the system affect its future behavior. Therefore, Agent-Based Social Simulation (ABSS) is used to acquire deeper knowledge about a complex social system rather than seeking to precisely predict its future behavior. ABSS therefore facilitates better-informed decisions by increasing decision-makers' knowledge about the system [1]. ABSS analysis provides the following two types of knowledge about a system.

Y. Goto (✉)
Iwate Prefectural University, 152-52 Sugo, Takizawa, Iwate, Japan
e-mail: y-goto@iwate-pu.ac.jp

S. Takahashi
Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo, Japan
e-mail: shingo@waseda.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_7, © Springer 2011

chen.shuheng@gmail.com

- Type 1: Knowledge about possible outcomes that result from the simulation of a policy alternative in a given situation.
- Type 2: Knowledge about a mechanism that results in a notable outcome, which results from the simulation of a policy alternative in a given situation.

The ABSS approaches that provide such knowledge are called "scenario analyses" [2], and several studies have been conducted in this regard. Deguchi [3] emphasized the importance of referring to a landscape that demonstrates a rough overall image of possible outcomes after a simulation of a large number of policy alternatives. Such a landscape is represented by a large number of dots on a two-dimensional plane that is defined by a vertical performance axis and a horizontal policy axis. Such a landscape can be helpful in elucidating the features of policy alternatives that satisfy a given performance criterion.

Yang et al. [4] use an inverse simulation technique to search in a large parameter space for a set of parameter values of a complex social system that would result in a notable outcome. This technique would be helpful when a desired outcome has already been identified. Both Deguchi [3] and Yang et al. [4], on the other hand, focused on the effect of complex interactions of the system upon the system's behavior, rather than the uncertainties of the system.

These uncertainties, however, do affect the behavior of a complex social system. We define these uncertainties as being those concepts for which modelers of the system have insufficient information or knowledge regarding their elements and/or their interactions in the system. These uncertainties would generally be faced by ABSS users. For example, we can easily imagine a situation in which the modelers of a system assume a distribution of the parameter values but lack a real set of parameter values. In this situation, they generate a tentative set of parameter values that are consistent with the assumed distribution.

The results observed after every run of a simulation in which both the same and varying sets of parameter values are used can vary considerably. The variation obtained using the same sets comes from the complex interactions in a system, while that obtained using varying sets comes from both the complex interactions and the variations in the parameter-value sets themselves. As described above, we assume variant sets of parameter values if we consider the uncertainty of the parameters. By reducing the uncertainty in the system, for which the results may have implications that vary considerably, we can enhance our Type 1 knowledge.

The primary purpose of this paper is to develop a novel scenario analysis method that mainly focuses on evaluating a range of possible outcomes in a system in the light of a given set of uncertainties. This paper is organized as follows. Section 2 introduces the proposed scenario analysis method. Section 3 presents a case example and the results obtained when this scenario analysis method was applied to it. Section 4 discusses the results obtained in Sect. 3 and evaluates the scenario analysis method. In Sect. 5, we summarize this study and indicate future avenues of research.

chen.shuheng@gmail.com

2 Landscape Analysis of Possible Outcomes

Figure 1 shows the configuration of the developed scenario-analysis method. Step 1 precedes the other three steps. Steps 2, 3, and 4 are independent of each other. Below, we describe each step in detail.

Step 1: Draw a landscape

In this step, the user draws a landscape of the possible outcomes of the considered uncertainties. This landscape illustrates the possible outcomes that would follow the implementation of each policy alternative. This step is further divided into the following two sub-steps:

- Step 1-1: Select the alternatives, a performance index, and a point in time for analysis.
- Step 1-2: Run ABSS and plot the performance index values based on the selected alternatives at the selected point in time.

In Step 1-1, it is important for the user to define a performance index that effectively reflects the system's behavior. In Step 1-2, the user summarizes a simulation log and visualizes both the distributions of the performance index values and their averages for the policy alternatives being considered. For each policy alternative, the user repeatedly runs ABSS a given number of times under the same condition. The user records the performance index values for each policy alternative at the target point in time and plots them on a two-dimensional plane defined by a vertical performance axis and a horizontal policy-alternative axis.

Step 2: Landscape analysis

In this step, the user analyzes the landscape drawn in Step 1 and understands the outcomes that are possible after the implementation of the given policy alternatives. The user then finds a feature of the policy effect with the considered uncertainties. This step consists of the following two sub-steps:

- Step 2-1: Observe a range of possible outcomes of a policy alternative.
- Step 2-2: Review the ranges of these policy alternative outcomes.

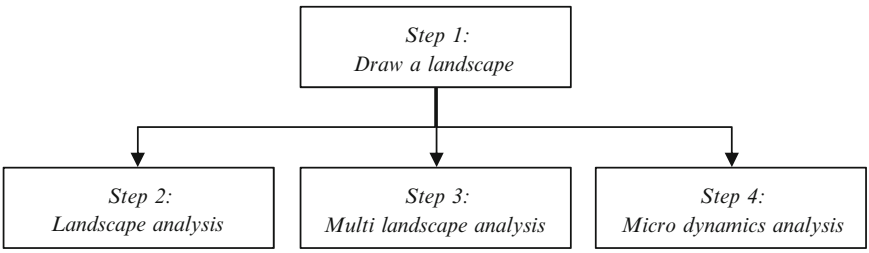


Fig. 1 Configuration of landscape analysis of possible outcomes

chen.shuheng@gmail.com

In Step 2-1, the user focuses on a policy alternative and studies its results, such as the highest or lowest performance of the possible outcomes, the mode or average performance of the outcomes, or the ranges of the outcomes. In Step 2-2, the user compares the ranges of the possible outcomes of different policy alternatives on the basis of the observations made in Step 2-1. Through this comparison, the user infers which policy alternative is potentially impacted by the considered uncertainties.

Step 3: Multi landscape analysis

In this step, the user compares the landscape that includes the considered uncertainties with that in which the uncertainties have been reduced. If there exist a difference between these two landscapes, this naturally implies that the uncertainties of the parameter affect the performance of the policy alternatives. The following two sub-steps are defined:

- Step 3-1: Draw a new landscape in which the uncertainties of a parameter are reduced.
- Step 3-2: Review the difference between the new landscape and the original one.

In Step 3-1, the user assumes that the considered uncertainties are reduced. Therefore, the user fixes the set of the parameter values that varies if there exist the uncertainties, and runs ABSS using this fixed set. The user draws a new landscape from the outcomes of the ABSS.

Statistical tests of both average (avg.) and variance (var.) are recommended for a comparison of the new distribution of possible outcomes with the original one in Step 3-2. The user infers the different possible outcomes from the statistical test if the user finds a significant difference between the two landscapes.

Step 4: Micro dynamics analysis

An agent-based system consists of agents. The agents make a decision referring their value, norms, or attributes. Agents' behavior is determined by the decision. The behavior of the system at the macro level is formed by agents' behavior at the micro level. Meanwhile, a policy alternative and the system behavior have an impact on agents' behavior. These bottom-up and top-down interactions are called "micro-macro links" of agent-based systems.

In this step, the user investigates when and how a notable outcome is generated. Micro dynamics analysis reveals a mechanism for the outcome at the macro level. The micro dynamics analysis focused on time-series log analyses and statistical analyses of agents or agents' behavior at the micro level, but not limited only them. This step is divided into the following three sub-steps:

- Step 4-1: Select a policy alternative of interest and a notable outcome of the policy alternative.
- Step 4-2: Path analysis of the outcome.
- Step 4-3: Understand the mechanism of the outcome.

chen.shuheng@gmail.com

In Step 4-2, a comparison is made between a notable outcome and an ordinary outcome in order to elucidate the differences between the two. Step 4, however, is a highly exploratory process. Therefore, we need to tentatively explore some viewpoints on the analysis.

3 Case Example

3.1 Model

We selected the model of Goto et al. [5], which describes the phenomenon of a change in a salesperson's behavior due to changes in their attitude as a result of organizational learning. A sales manager evaluates salespersons using an evaluation system in the sales division. The salespersons have their own attitudes toward their sales activities, and their actions are based on these attitudes. The attitude of each salesperson is different. Basically, while some salespersons learn new attitudes to improve their evaluation by the evaluation system, others do not because of a lack of interest in their own evaluations. The salespersons learn by exchanging information about their evaluations.

Our model is classified as a middle-range model, not as a facsimile model whose intention is to reproduce some specific target phenomenon [7]. Four types of sales activities are introduced in our model. We do not, however, claim that these include all possible types of sales activities. We sift through the various activities while trying to analyze the possible effects of the evaluation system. The evaluation indices in practical use reflect various intentions. For example, some indices try to facilitate a selfish/altruistic activity and others focus on short-/long-term performance.In this paper, we define these four activities to analyze the effect of these evaluation indices.

Whether the salespersons adjust their attitude or not will have a critical impact on the effect of an evaluation system. We therefore introduce the following two learning factors in order to analyze this impact: (1) which salesperson learns and (2) when salespersons are satisfied with their evaluation.

3.1.1 Sales Division and Salespersons

Consider a sales division that has GN groups that have AN salespersons. The salespersons are required to sell goods, and all groups have the same types of goods to sell. Group $i(=1,2,\cdots,GN)$ initially has m_i customers and can sell up to m_i units of goods (one unit for each customer) in a sales period. Let m'_i be the number of customers who have not yet purchased a good from group i in the current period.At the beginning of the sales period, we set $m'_i = m_i$. Both m_i and m'_i will increase or decrease depending on the activities of the salespersons.

chen.shuheng@gmail.com

Salespersons have the following four attributes: (1) sales capability cp $(0 \leq cp \leq 1)$; (2) three types of sales attitude: aggressiveness ag, cooperativeness co, and innovativeness $in\,(ag, co, in \in \{0, 1, \cdots, 7\})$; (3) a learning discriminator $ld \in \{0, 1\}$; and (4) a learning threshold $th\,(> 0)$. Higher cp values result in higher probabilities of sales success. Aggressiveness ag is connected to the frequency of market cultivation during a sales period; cooperativeness co, to the frequency of instances of educating teammates; and innovativeness in, to the frequency of taking training initiatives. A salesperson's cp increases or decreases with the sales activity of the salesperson, with that of others, and with time. Organizational learning of an attitude can lead to a change in ag, co, and in. Both ld and th are defined initially and are fixed.

3.1.2 Organizational Behavior

Salespersons take part in the following four activities during a sales period: (1) sales, (2) market cultivation, (3) education of teammates, and (4) training. Sales refers to the selling of goods to a customer, if $m_i' > 0$. The probability of sales success is equal to a salesperson's cp. Market cultivation refers to the process of seeking new customers for a good, and it succeeds at a probability of $1 - {}^{m_i}/_{(AN \times T)}$, where T denotes the number of time units in a sales period. If the salesperson's market cultivation is successful, then both m_i and m_i' increase by one. The education of teammates leads to an increase in the cp of all teammates by $E_o (0 \leq E_o \leq 1)$. As a result of the training, the salesperson's cp increases by $E_s (0 \leq E_s \leq 1)$.

Both cp and m_i diminish over time. The salespersons' cp decreases by $E_d (0 \leq E_d \leq 1)$ with each time unit. The consumers of group i m_i decrease by one at a probability of $P_d (0 \leq P_d \leq 1)$ for each time unit.

3.1.3 Evaluation System

The evaluation system consists of a number of evaluation indices and their weights. The salesperson's evaluation value ev is defined as the weighted sum of n evaluation indices $(ev_1, \ldots, ev_n) : ev = w_1 \cdot ev_1 + \cdots + w_j \cdot ev_j \cdots + w_n \cdot ev_n$, where w_j is a weight of the jth index $(0 < w_1, \ldots, w_n \leq 1, w_1 + \cdots + w_n = 1)$. According to Otomasa [7] and interviews we conducted in a number of sales divisions, 40 evaluation indices were found and used in this case example.

3.1.4 Organizational Learning

Salespersons learn their attitude toward sales activity (ag, co, and in) until their evaluation value ev meets their learning threshold th. The attitude of salespersons whose ld is 1 and whose ev is below their th gradually changes to attitudes with which they achieve higher evaluation values. Moreover, the attitudes of all salespersons change randomly at a very low rate irrespective of their ld.

chen.shuheng@gmail.com

3.2 Simulation

3.2.1 Verification and Validation

Accurate verification and validation (V & V) is critical to the development and use of the ABSS model [1, 8]. Gilbert indicates that verification refers to confirming whether an implemented model matches its conceptual specification [6]. In this paper, our method of verification is to trace all intermediate outputs from ABSS and to check if these outputs are consistent with those calculated manually by an analyst.

An adequate validation of an ABSS model depends on the model's aim. As described above, our model is classified as a middle-range model. In this paper, therefore, we confirm the validation in terms of consistency with the stylized facts of the considered area. To do this, we performed model-to-model analysis [9] and parameter-sweeping on our model.

Table 1 shows a list of parameters and their validated values. The behavior of our model, as specified by the parameter values in Table 1, matches the stylized facts of the concerning area. We have not presented the V&V test results here due to space limitations.

3.2.2 Experimental Design

The sales manager, who is a user of our scenario-analysis method, analyzes the configuration of the evaluation system for the sales division. The manager is aware of the distributions of both a salesperson's capability and attitude toward sales activity, and he/she also knows how many persons actually learn their attitude. However, the manager does not know the real set of these parameter values or which person actually learns and which does not. The manager examines the effect of such uncertainties. Table 2 shows the uncertainties involved in salesperson's characteristics.

In the experiment, we make the uncertainties operational. Every run starts with a unique set of parameter values which are consistent with the distribution, as described in Table 2, if the parameter has uncertainties. On the other hand, every run starts with the same set if the parameter uncertainties are reduced. This operationalization is realized by managing random seeds of our ABSS program.

Table 1 Parameter setting

Parameter		Value
Number of groups	GN	10
Salespersons/group	AN	10
Number of customers of group i	m_i	100
Rate of customer decrease	P_d	0.25
Improvement by training	E_s	0.015
Improvement by education	E_o	0.005
Decrease over time	E_d	0.005
Number of cycles	CN	60
Number of possible actions in a sales period	T	50
Mutation rate	P_m	0.001

chen.shuheng@gmail.com

Table 2 Experimental design

Parameter		Value
Capability		$cp \sim N(0.4,\ 0.01)$
Aggressiveness		$ag \sim N(2.0,\ 4.0)$
Cooperativeness		$co \sim N(2.0,\ 4.0)$
Innovativeness		$in \sim N(5.0,\ 4.0)$
Threshold		$th \sim N(1.1,\ 0.01)$
Number of sales periods	*CN*	36
Number of learning persons	*LN*	80

Table 3 Effect of uncertainties

	(1) → (2) Learning member specified	(1) → (3) Capability, attitude, and threshold specified	(1) → (4) All characteristics specified
Avg.	7.5% (3/40)	55% (22/40)	55% (22/40)
Var.	10% (4/40)	37.5% (15/40)	47.5% (19/40)

3.2.3 Result

The experimental result of ABSS is shown in Figs. 2 and 3 and summarized in Table 3. Figure 2 illustrates the four landscapes under different conditions of uncertainties. In Step 2 of our method, the manager analyzes the drawn landscape and understands the possible outcomes that would follow the implementation of the considered policy alternatives. For example, in Fig. 2(1), the manager gains the following insights into the 1st policy alternative: the best outcome for average sales per member is 29.2, the worst outcome is 8.9, and the average outcome is 19.1. The manager also learns that the 38th alternative achieves the best average outcome (30.8), the 32nd alternative, the best outcome (34.5), and the 17th alternative, the minimum variance of the outcomes (0.134).

In Step 3 of our method, the manager compares a landscape that includes the considered uncertainties with one in which the uncertainties have been reduced. Table 3 shows the percentages of the occurrence of changes in avg. or var. between the two landscapes. This change under a policy alternative is counted when the possible outcomes based on an uncertainty setting are statistically different from those based on another uncertainty setting. The denominator 40 represents the number of considered policy alternatives. The result indicates that all three changes in uncertainties make the landscape statistically different.

Figure 3 illustrates an interesting phenomenon in the 32nd policy alternative, in which a bifurcation arising in the salespersons' performance causes a notable and very-low-performance outcome; the outcome of this policy is emphasized in Fig. 2(1). By performing the micro dynamics analysis in Step 4 of our method, the manager learns that this phenomenon reflects a specific structure that depends on the characteristics of the salespersons.

chen.shuheng@gmail.com

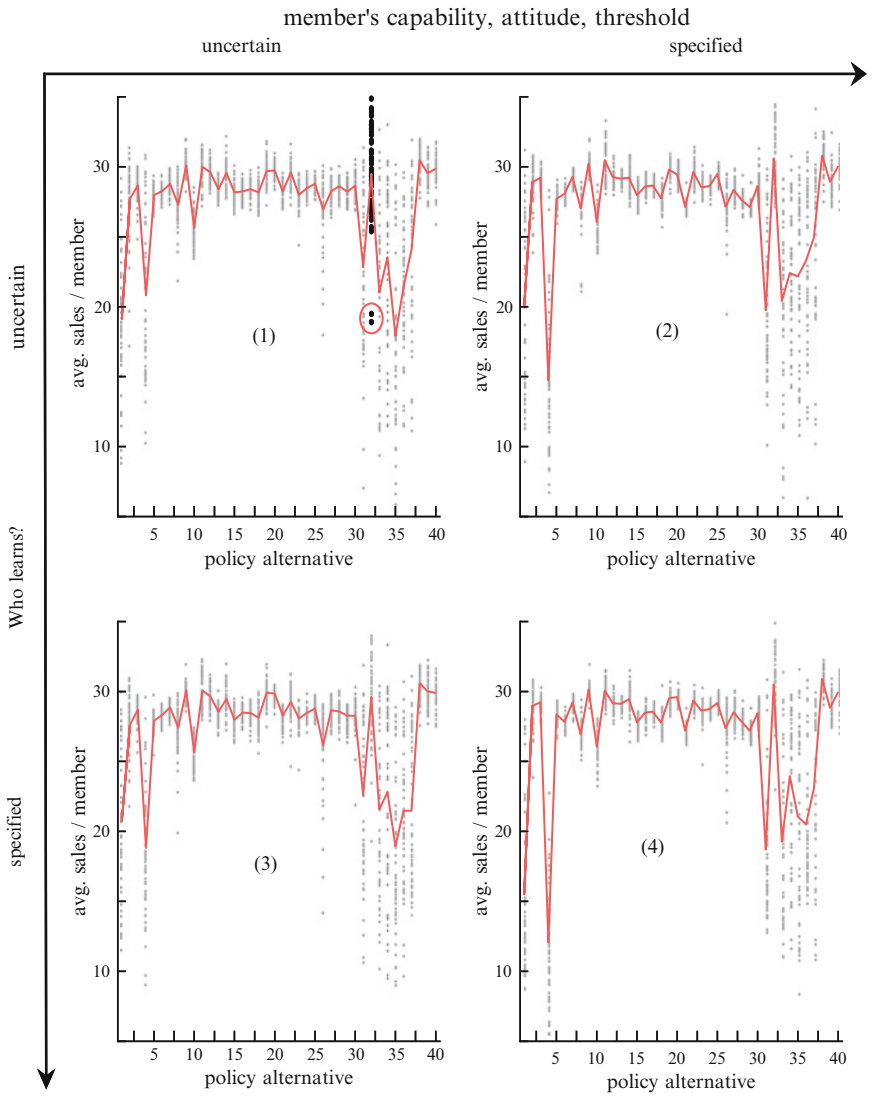


Fig. 2 Landscapes of possible outcomes

4 Discussion

4.1 Analysis of Uncertainties

The simulation result for the case example shows that we can see the difference in both the avg. and var. of the possible outcomes provided by uncertainties in the characteristics of salespersons. Micro dynamics analysis also shows when and how the

chen.shuheng@gmail.com

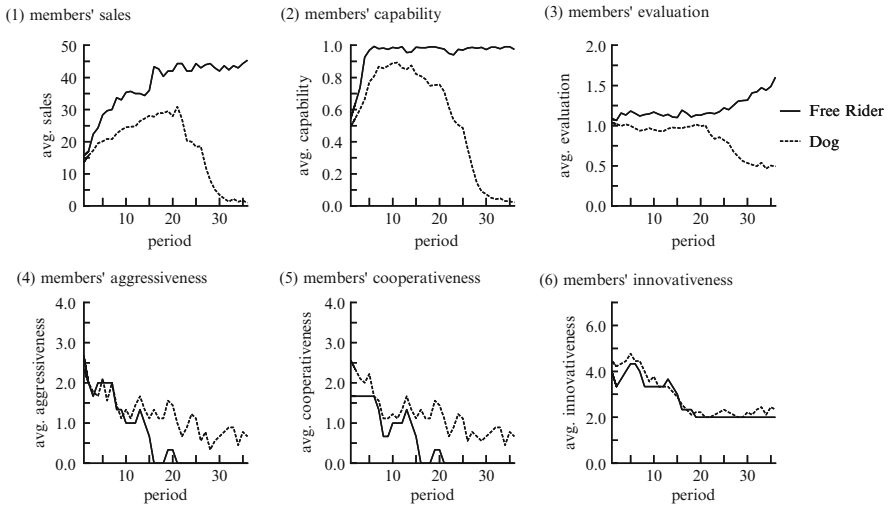


Fig. 3 Bifurcation phenomenon in a notable outcome

structure of salesperson's characteristics produces notable outcomes. It thus appears reasonable to conclude that the uncertainties have an impact on the effectiveness of policy alternatives.

In the case example, the reduction of the uncertainties allows for a more precise evaluation of the possible outcomes of the considered policy alternatives. From the viewpoint of managers, our result implies that the reduction of uncertainties is effective.

A number of studies have been conducted about effects of an evaluation system. However, there are few credible insights. Propositions claimed in a study frequently conflict with those claimed in another study. Siggelkow and Rivkin have claimed that such a misalignment arises from differences in implicit assumptions about the model and the unrecognized experimental conditions [10]. The result in this paper also supports this claim.

4.2 Evaluation and Scope

We applied our scenario analysis method to the case example in Sect. 3, and the result supports our claim that our method enables us to acquire the two types of knowledge described in Sect. 1. Although the effectiveness of our method has been demonstrated in just one example, it was nonetheless clearly effective. Our scenario analysis method does not include a concrete, rigid procedure that is restricted to a specific domain of research. It is therefore natural to consider it a domain-free method.

chen.shuheng@gmail.com

Additionally, our scenario analysis method and other such methods are more compatible than competitive with one another. The landscape proposed by Deguchi [3] provides a rough overall image of possible outcomes after a simulation of a large number of policy alternatives. Following this analysis, we can focus on a few policy alternatives rather than carrying out further analyses for all the alternatives. Our method provides knowledge about the possible outcomes of a system with given uncertainties after the implementation of the focused policy alternatives, and about the mechanism involved in obtaining a notable outcome. An inverse simulation technique [4] may allow for a detailed study of the mechanism specified by our method. This comprehensive process of scenario analysis with ABSS is expected to help managers make better-informed decisions.

5 Summary and Future Study

We have developed a scenario analysis method that mainly focuses on evaluating a range of possible outcomes of a system with considered uncertainties. We applied this method to the example of decision-making in the configuration of the evaluation systems for a sales division. As validation of our method, we confirmed the existence of effects that are dependent on the uncertainties in salespersons' characteristics, and we found a mechanism in which the specific structure of salesperson's characteristics causes a notable outcome of the system.

However, our scenario analysis method captures an outcome of a system at a specific point in time. Nakada et al. have proposed a scenario analysis method that captures an outcome as a dynamic time-series behaviour [11]. The future study will consider an outcome of the system from a dynamical point of view. This study will contribute toward further advancing applicable fields of ABSS and strengthen the power of scenario analyses.

Acknowledgement This work was supported in part by Grants-in-Aid for Scientific Research No. 21310097 and No. 22730312.

References

1. North M, Macal CM (2007) Managing business complexity: discovering strategic solutions with agent-based modeling and simulation.Oxford University Press, New York
2. Takahashi S (2008) Organization design and social simulation. Commun Oper Res Soc Jpn 53(12):686–691 (in Japanese)
3. Deguchi H (2009) Dawn of agent-based social systems sciences. In: Deguchi H, Kijima K (eds) Manifesto: agent-based social systems sciences. Keiso-Shobo, Tokyo (in Japanese)
4. Yang C, Kurahashi S, Kurahashi K, Ono I, Terano T (2009) Agent-based simulation on women's role in a family line on civil service examination in Chinese history. J Artif Soc Soc Simul 12(2):5 http://jasss.soc.surrey.ac.uk/12/2/5.html

chen.shuheng@gmail.com

5. Goto Y, Takahashi S, Senoue Y (2009) Analysis of performance measurement system for knowledge sharing under intraorganizational competition. J Jpn Soc Manage Inf 18(1):15–49 (in Japanese)
6. Gilbert N (2007) Agent-based models. Sage, London
7. Otomasa S (2003) On use of performance measurement indices in Japanese companies. Rokkodai-Ronshu Manage Ser 49(4):29–54 (in Japanese)
8. Richiardi M, Leombruni R, Saam N, Sonnessa M (2006) A common protocol for agent-based social simulation. J Artif Soc Soc Simul 9(1):15 http://jasss.soc.surrey.ac.uk/9/1/15.html
9. Hales D, Rouchier J, Edmonds B (2003) Model-to-model analysis. J Artif Soc Soc Simul 6(4):5 http://jasss.soc.surrey.ac.uk/6/4/5.html
10. Siggelkow N, Rivkin JW (2005) Speed and search: designing organizations for turbulence and complexity. Organ Sci 16(2):101–122
11. Nakada T, Takadama K, Watanabe S (2009) Analysis method depending on Bayes' theorem for agent-based simulations. In: Proceedings of the 6th international workshop on agent-based approaches in economic and social complex systems, pp 306–317

chen.shuheng@gmail.com

The Flow of Information Through People's Network and Its Effect on Japanese Public Pension System

Masatoshi Murakami and Noriyuki Tanida

Abstract In this article, we would like to verify a positive or negative policy impact that comes from decreasing or increasing people's distrust in Japanese public pension system. For the sake of tackling these issues, firstly, we pick up some network models that fit well in real people's network. Secondly, we put the information, which is concerning about Japanese public pension system, on agent-to-agent network. Finally, we ascertain the effect of releasing the information and its expansion on Japanese public pension system. Consequently, it is revealed that releasing information over again have a profound effect on reducing distrust in public pension system and on pension premium fund. With releasing information to a limited extent and to a limited number of agents, there is limited effect on reducing agent's distrust.

Keywords Agent based simulation · Network structure · Public pension system · Policy impact

1 Introduction

Japanese public pension system is burdened by many problems. That is, rapid demographic aging, slow economic growth, the change of household structure and the change of employment situation. It is said that one of the biggest problems is the hollowing out of national pension. The term, "the hollowing out of national pension," means that many people refuse to join or not to pay pension premiums.

M. Murakami (✉)
The Research Institute for Socionetwork Strategies, Kansai University,
3-3-35, Yamate, Suita, Osaka 564-8680, Japan
e-mail: murakami@rcss.kansai-u.ac.jp; a094056@kansai-u.ac.jp

N. Tanida
Faculty of Economics, Kansai University, 3-3-35, Yamate, Suita, Osaka 564-8680, Japan
e-mail: tanida@kansai-u.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_8, © Springer 2011

chen.shuheng@gmail.com

Although the hollowing out of national pension has a variety of causes, people's distrust in Japanese public pension system and its propagation is a potential source of it. In spite of the situation, except our simulation, many simulations for Japanese public pension system do not deal with the cause or change of the hollowing out of national pension. Therefore, in this article, we will deal with its problem using agent-based simulation.

In our previous study [1], we explored the relationship between people's network structure and instability of the pension system. We clarified how spreading people's distrust in Japanese public pension system through their network would have effect on pension premium fund. In that study, we created our agent based simulation model with care and attention for economical, sociological and network theoretical points. Consequently, we concluded that people's distrust in public pension system and the propagation of distrust has a decisive influence on the pension premium fund. Especially, it is clarified that not hub-agents but agents whom the value of betweeness centrality is high have significant role to the future pension premium fund.

However, some significant issues still need to be tackled in our previous study. For example, using our network models, we should verify a policy impacts that comes from increasing or decreasing people's distrust in public pension system. Which is acceptable to increase or decrease the entire people's distrust in public pension system or to increase or decrease some people's one? For the sake of tackling these issues, firstly, we will pick up some network models that fit well in real people's network. Secondly, we will put the information, which is concerning about Japanese public pension system, on agent-to-agent network. Nonetheless, we will adopt three ways of releasing the information. The first is to put the information on the entire network. The second is to release it on hub agents. Moreover, the third is to release it on agents whom the value of betweeness centrality is high. Finally, we will ascertain the effect of releasing the information and its expansion on Japanese public pension system.

2 Basics

2.1 The Function for Agent to Agent Interactions

First of all, we will begin with explain each agent's interaction inside our network model. Based on some agent's networks that described above, agents interact with each other according to the function as described below.

In Japan, many people might form their own opinion about Japanese public pension system. On the ground of their opinion about Japanese public pension system, they should determine whether they should pay pension premium. However, recently in Japan, the negative opinions for Japanese public pension system become common and many people do not come to pay pension premium. Why does the number of people non-paying pension premium increase? It is partly because some people's negative opinion in public pension system penetrates into many others.

chen.shuheng@gmail.com

It is not difficult to understand that the propagation of people's distrust in public pension system is the conformation of public opinion, in a way. Therefore, we treat this subject as the process of forming public opinion.

Many studies have been made on the process of forming public opinion in the area of sociology, especially, in the area of Dynamic Social Impact Theory and the simulation by using it. We denote Dynamic Social Impact Theory as *DSIT*, hereafter.

In [1], we have carefully examined a series of works in terms of *DSIT*. *DSIT* and its simulation capture how a collective phenomenon is emerged as a result of the interpersonal interaction in terms of "Consolidation", "Clustering", "Continuing Diversity" and "Correlation". According to [2], in the area of *DSIT*, many researchers used "individual-level theory of social impact, which is explained in [3], to explain and predict the emergence of group-level phenomena as consequence of social influence among interacting people."

In [3], Latané clarified the gap between the theory of [4, 5] and practice. In [4] and [5], Asch and Abelson concluded that individuals easily defer to the opinion of the majority. As a corollary to these conclusions, all people should have same opinion. However, in reality, people have different opinions in a variety of context. To clarify this subject, Latané developed SIT and *DSIT* in [3].

Although, in this area, Latané made pioneering study in [3], many literatures also have proposed improvements. The latest research developments in Japan were made by Morio in [6].

In [6], agent's attitude score that shows the change of agent's attitude for a matter as the discrete value (0 or 1) does not change rapidly, since the function that he adopts is continuous function. An agent's attitude score changes non-linearly and gradually. Also, in [6], the attitude function can have various style depending on the value p of linearity coefficient, for example, step function, non-linear function, and linear function for $p = 0,\ 0.5,\ 1$, respectively. In addition to this, according to [7], "The assumption that attitudes have a flip-flop character is at odds with the long tradition of attitude measurement in social psychology that seems to show that attitudes are continuous." Moreover, according to [7], "As recent developments have shown, however, the dividing line between the dynamics of models lies not between continuous and discrete models, but between linear and nonlinear ones." These clearly show that non-linear and continuous function is better function than step function for describing the dynamics of agent's attitude more precisely.

For these reasons mentioned above, we implement Morio's model into our new model to reveal the transition process of peoples opinion concerning about public pension system.

Let us explain Morio's model in more detail. In his model, the influence function $y = f(x)$ that the decision function of getting feedback from the other is defined as follows:

$$f(x) = \begin{cases} x^p M^{p-1} & \text{if } x \geq 0 \\ -(-x)^p M^{p-1} & \text{if } x < 0 \end{cases}, \tag{1}$$

where x is the input value, and p is the linearity coefficient and M is the maximum value of agent's attitude score.

chen.shuheng@gmail.com

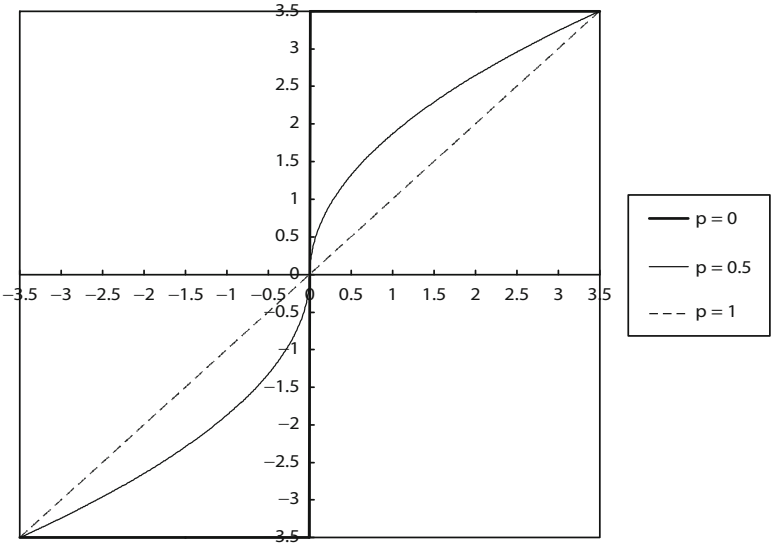


Fig. 1 The decision function divided by p, source [6], p. 161, Fig. 1

Given that p is 0, 0.5 and 1, the shapes of the function (1) can be drawn as in Fig. 1. As in Fig. 1, the shape of the function (1) is changed with an increase in p. If p is 0, the function (1) get into step function. On the other hand, if p is 0.5, the function (1) gets into non-linear function. In addition, if p is 1, the function (1) gets into linear function. It means that the function (1) become non-linearity function if p is greater than 0 and lesser than 1.

In addition to this, the aggregate sum of influence from the other agents I is defined as follows:

$$I = \frac{1}{N} \sum \frac{X_j}{d^m}, \tag{2}$$

where X_j is j th agent's attitude score, and m is distance coefficient, and N is total number of agent, and d is the distance between ith agent and j th agent.

With these definitions described above in mind, we will now take a look at function I in more detail. In [6], the value of X_{ij+1} which is ith agent's attitude score in next step, is defined as follows:

$$\begin{aligned} X_{ij+1} &= f(q) \\ q &= \left| wX_{ij} + (1-w)I_j \right|, \end{aligned} \tag{3}$$

where w is the weight of "ego-involvement." $1 - w$ is the weight of the influence from the other agents. I_j is total sum of influence from the other agents to ith agent, X_{ij} is ith agent's attitude score in jth step, and q is the absolute value of the weighted average which is calculated between the attitude score of ith agent and the other agents.

chen.shuheng@gmail.com

We implement the functions as stated above into our model. However, it is important to notice that our model differ substantially from [6]. Our model is not concerned with the closed set of agent. In his model, the number of agent is predetermined and it is not change. By contrast to this, in our model, each agent gets older and dies in his (or her) bed.[1] At the beginning of year, new agents (20 years old) are created. In these regards, our model is different from [6] and these points are important for making our new model fit well in real world.

2.2 Three Basic Network Models

In this section, we will explain about three basic network models that adopted in this article. That is Random Network Model (*RNM*, hereafter), Barabasi–Albert Model (*BAM*, hereafter) and Threshold Model (*TRM*, hereafter).

RNM has a feature that edges between vertices are stochastically-generated. It is obtained by adding edges between vertices at random. It is said that *RNM* is the most simplified network.

BAM has a feature that (a) network expands continuously by the addition of new vertices and (b) new vertices attach preferentially to sites that are already well connected.[2]

In *TRM*, all of vertices has weight $w_i (i = 0, 1, \ldots, n)$ and they are distributed according to probability density function $f(x)$ defined as follows:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x \leq 0) \end{cases}, \tag{4}$$

where λ is a positive constant. That is, the weight of each vertex is exponentially distributed in *TRM*. If the sum of w_1 and w_2 exceed or even equal threshold θ, vertices link with each other. According to the condition of parameter, *TRM* also have small world phenomenon and meet scale freeness.[3]

3 Choice of Network Model that Fit Well in Real People's Network

In our previous study, using some network indices (e.g. the average of shortest path, the clustering coefficient and degree distribution), we selected some network models that fit well in real people's network. In general, it is said that real people's

[1] In our model, we define one simulation step as 1 month.

[2] See [8], p. 509.

[3] Refer to [9] about *TRM*.

chen.shuheng@gmail.com

network has some features, which meet small world phenomenon and scale freeness. Therefore, they are key factors in case where we select some network models.

Unlike three basic network models, that is, *RNM*, *BAM* and *TRM*, in our model, agents repeatedly remake their network in their life. An agent who is hub agent at one time gets into less linked agent at any other time (simulation step). This is the special and distinguished points of our simulation model.[4]

In our previous study, we implement and test seven models based on *RNM*, *BAM* and *TRM*. These models differ in the formation of agent's network, in the timing of breaking off agent-to-agent relationships and in the linked agent's age groups. Table 1 shows the feature of each model. In addition, agents interact with each other in each model. Each agent has attitude score that vary with the interaction with the other agent.[5]

As a consequence, Model 4 and Model 6 $(\lambda = 0.5, \theta = 3)$ is selected because they fit in with the reality of people's network. Specifically, their averages of shortest path are small, their clustering coefficients are relatively high and their degree distributions have good fit with power law. That is, they have small world phenomenon and meet scale freeness.

Table 1 The feature of each model

Model	Group	Linked agent	The timing of breaking off agent-to-agent relationships	Remarks column
0	RNM	Same age groups	Every 13 steps	20% of agents break their link
1	BAM	Same age groups	Every 13 steps	20% of agents break their link
2	RNM	Different age groups	Every 13 steps	20% of agents break their link
3	BAM	Different age groups	Every 13 steps	20% of agents break their link
4	RNM	Different age groups	Every 179 steps	Assume that agents drift away from each other
5	BAM	Different age groups	Every 179 steps	Assume that agents drift away from each other
6	TRM	Same age groups	Every 13 steps	#20% of agents break their link #Combination of threshold and the exponentially-distributed parameters

[4] In general, the number of agents in many other models is fixed or not decreased.

[5] In this article, we define agent's attitude score as the degree of trust in public pension system. If an agent's attitude score is negative, we identify the agent has distrust in public pension system. If it is positive, we identify the agent does not have distrust in public pension system.

chen.shuheng@gmail.com

4 Put the Information on the Agent-to-Agent Network

We will release the information, which is concerning about Japanese public pension system, on agent-to-agent network. The information is composed by the positive and negative one. The positive information produces great improvement of Japanese public pension system. For example, you can imagine the prediction of the increase in pension benefit that every pensioner receives after retirement.

On the contrary, the negative information aggravates Japanese public pension system. For example, you can imagine the prediction of the decrease in pension benefit that every pensioner receives after retirement. People's reaction to two types of information should be different. In general, people react to the positive information with near-silence, but they overreact to the negative information. However, because of the limitation on the number of pages that we can write, we will deal especially with positive information.

Nonetheless, we will adopt three ways of releasing the information. The first is to release the information on the entire network. The second is to release it on hub agents. The third is to release it on agents whom the value of betweeness centrality is high. Each way is shown in Fig. 2.

As shown in Fig. 2, in the first way, the information that decreases or increases agent distrust in public pension system is released to the entire agent. In this case, it does not matter whether the agent is hub agent or not. By contrast, in the second way,

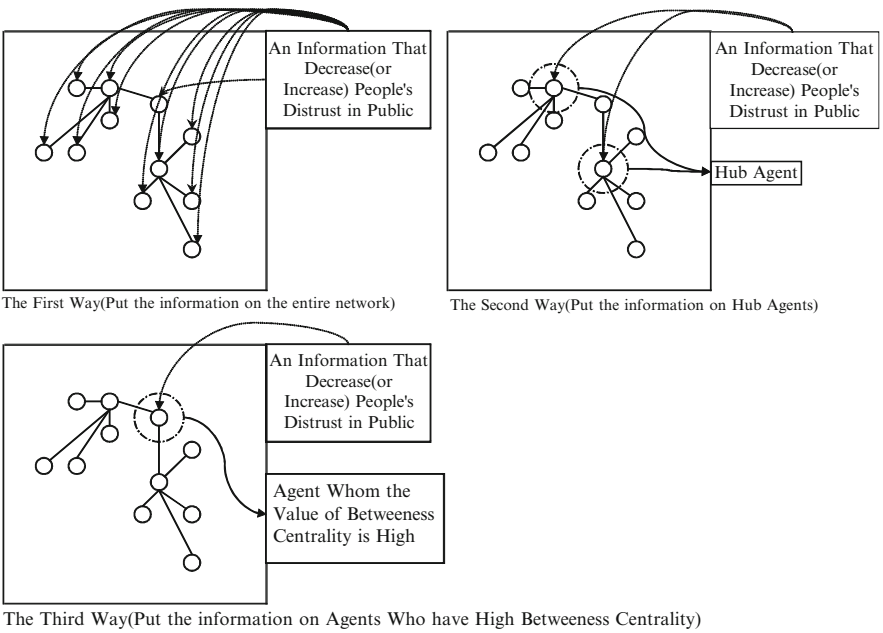


Fig. 2 Three ways of releasing positive information

chen.shuheng@gmail.com

the information that decreases or increases agent distrust in public pension system is released to hub agent. Hub agents link many other agents. Therefore, it is expected that the information diffuse rapidly in almost all agent. As well, in the third way, agents whom the value of betweeness centrality is high may be linked hub agents. In this case, the positive information may also diffuse rapidly in almost all agents.

5 The Effect of Releasing the Information and Its Expansion on Japanese Public Pension System

5.1 The Execution of Simulation

Based on the models and ways as described before, we will ascertain the effect of releasing the information and its expansion on Japanese public pension system. In the verification process, it is important to consider all the various factors together. How agents are affected by the communication exchanged between them?

In our simulation, positive information is released at the 384th step. It is intermediate step. Even though, in our simulation model, the number of agent who receives information change at will, in the second and third way, information is released to the top 5, 20, 50, and 100 agents who have many links or whom the value of betweeness centrality is high. Although how many agents would be receiving information is a moot point, we hasten to decide 10% of agents receive information. The models and each ways are shown in Table 2.

Table 2 The models and the methods, types, timings of releasing information

	Model	The method of releasing information	The type of information	The timing of releasing information
Case 1	Model 4	To all agents	Positive information	384th step
Case 2		To hub agents (Top 5, 20, 50, 100)	Positive information	384th step
Case 3		To agents who have high betweeness centrality (Top 5, 20, 50, 100)	Positive information	384th step
Case 4	Model 6 ($\lambda = 0.5,\ \theta = 3$)	To all agents	Positive information	384th step
Case 5		To hub agents (Top 5, 20, 50, 100)	Positive information	384th step
Case 6		To agents who have high betweeness centrality (Top 5, 20, 50, 100)	Positive information	384th Step

chen.shuheng@gmail.com

As shown in Table 2, we select two network models that fit well in real people's network. That is Model 4 and Model 6 ($\lambda = 0.5, \theta = 3$). Three method of releasing information are adopted. That is, we released positive information to all agents, to top 5, 20, 50, and 100 hub agents who have many links and to top 5, 20, 50 and 100 agents who have high betweeness centrality. The results of simulation are discussed in more detail below.

5.2 Simulation Results

5.2.1 The Altered Distribution in Agent's Attitude Score

Now, we will explain about the results of our simulation. Especially, we will explain about the results of simulation in which the positive information is released to agents. First, we checked the change in the number of agent who has distrust in public pension system. The results of simulation are indicated in Table 3.

Except for the case that information is released to the entire agent, the trends are split down the middle. That is, after releasing information, the number of agent who have distrust in public pension system continues to mount in Model 6, but it does not so in Model 4. However, except for Case 6, it can be said that releasing the information to top 50–100 agents who have many links or high betweeness centrality have more effective than that to top 5–20 agents who have many links or high betweeness centrality. Especially, there really is difference between information is released to top 20 agents and that to top 50 agents. In addition, it can be said that releasing the information to agents who have many links is more effective than that to agents whom betweeness centrality is high. In model 4, it is obtained by comparison between Case 2 and Case 3. In addition, in model 6, it is obtained by comparison between Case 5 and Case 6.

5.2.2 Investigating the Reasons for the Differences from "Agent's Links"

So, what caused these differences? We will investigate the possibility of linked agents, the number of them and their attitude score as a cause. Especially, we will investigate the change of them before and after step when information is released. Figure 3 indicates the number of link that each agent has and their attitude score at 354th step and 413th step in Case 2. The result from simulation that information is released to top 20 hub agents and top 50 hub agents is shown in Fig. 3.

As shown in Fig. 3, the difference in the number of agents who receive information is obvious. It is just conceivable that agents who receive information affect the other agents who do not receive information. In the case that the number of hub agents who receive information is 50, agents who are supposed to have negative attitude score have positive attitude score at 413th step. Consequently, this is attributed the fact that many other agents' attitude score go positive after releasing

chen.shuheng@gmail.com

chen.shuheng@gmail.com

Table 3 The change in the number of agent who has distrust in public pension system

			The number of steps							
			1st	100th	200th	300th	400th	500th	600th	700th
Model 4	Case 1	Information is released to all the agent	112	107	125	148	5	16	11	11
	Case 2	Information is not released	112	107	125	148	163	187	156	166
		Information is released to top 5 hub agents	112	107	125	148	161	186	155	166
		Information is released to top 20 hub agents	112	107	125	148	149	180	154	162
		Information is released to top 50 hub agents	112	107	125	148	122	142	111	118
		Information is released to top 100 hub agents	112	107	125	148	58	49	19	12
	Case 3	Information is not released	112	107	125	148	163	187	156	166
		Information is released to top 5 betweeness centrality agents	112	107	125	148	158	176	153	166
		Information is released to top 20 betweeness centrality agents	112	107	125	148	145	152	146	159
		Information is released to top 50 betweeness centrality agents	112	107	125	148	111	116	131	152
		Information is released to top 100 betweeness centrality agents	112	107	125	148	51	45	21	12
Model 6	Case 4	Information is released to all the agent	52	85	91	138	62	49	46	21

Case 5	Information is not released	52	85	91	138	180	227	268	286
	Information is released to top 5 hub agents	52	85	91	138	173	225	268	286
	Information is released to top 20 hub agents	52	85	91	138	155	209	246	275
	Information is released to top 50 hub agents	52	85	91	138	123	169	220	259
	Information is released to top 100 hub agents	52	85	91	138	88	124	146	156
Case 6	Information is not released	52	85	91	138	180	227	268	286
	Information is released to top 5 betweeness centrality agents	52	85	91	138	172	226	268	286
	Information is released to top 20 betweeness centrality agents	52	85	91	138	148	204	251	276
	Information is released to top 50 betweeness centrality agents	52	85	91	138	117	160	218	256
	Information is released to top 100 betweeness centrality agents	52	85	91	138	104	167	226	257

chen.shuheng@gmail.com

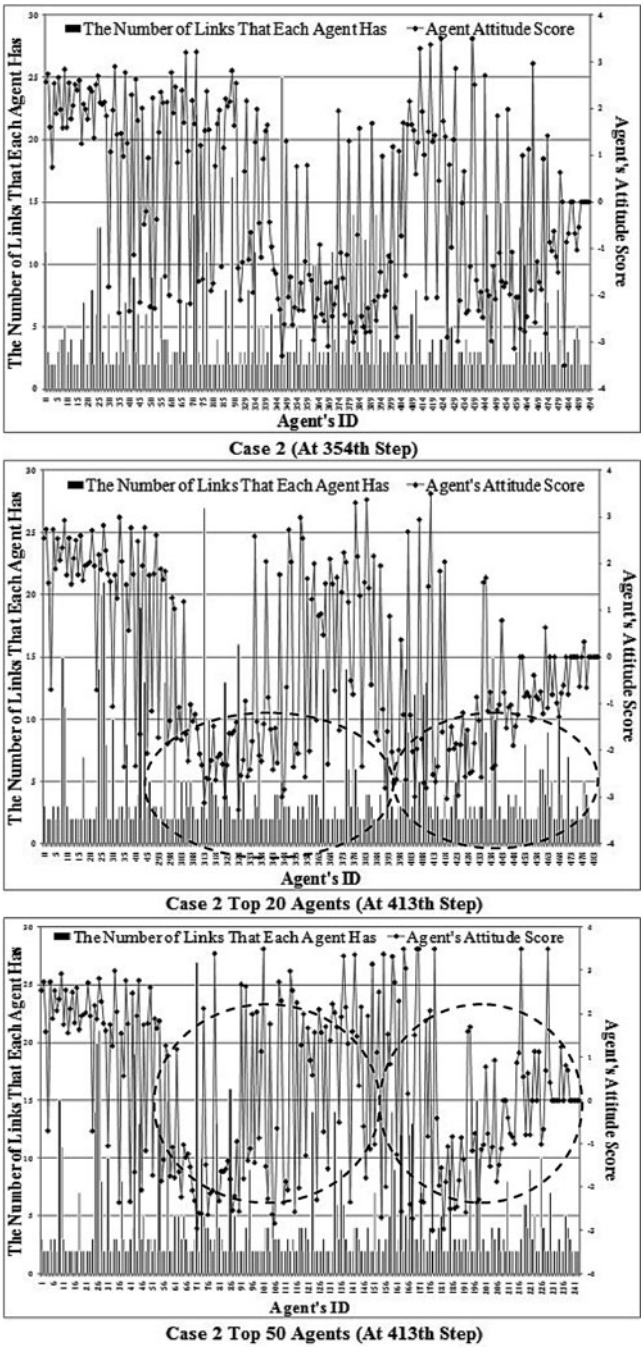
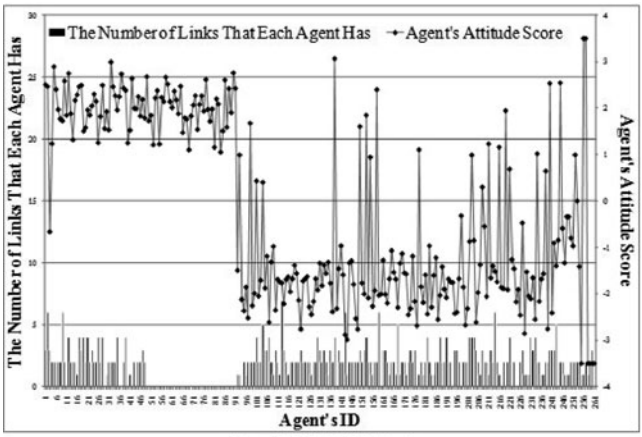
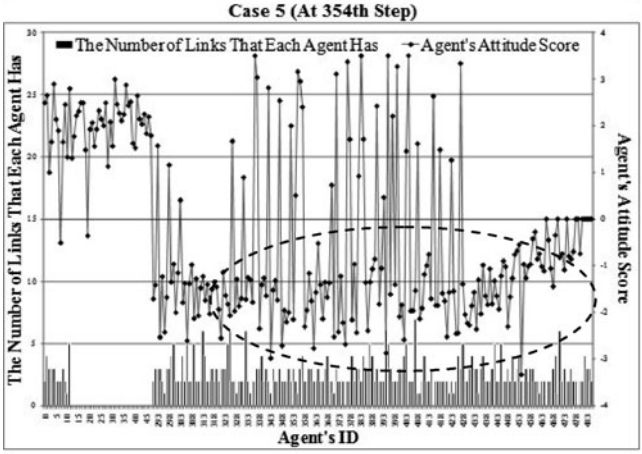


Fig. 3 The difference in the number of hub agents who are released information (Case 2)

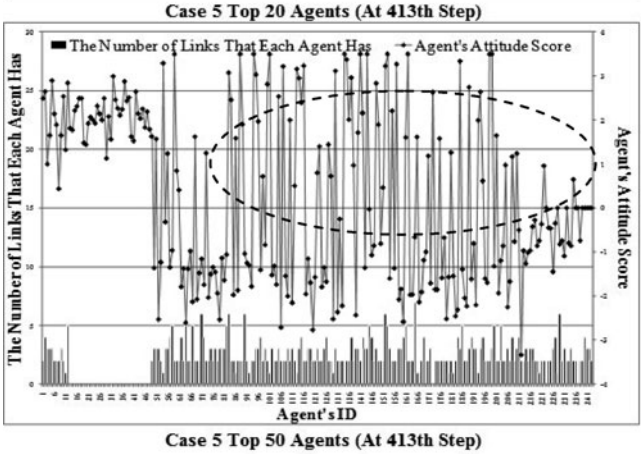
chen.shuheng@gmail.com



Case 5 (At 354th Step)



Case 5 Top 20 Agents (At 413th Step)



Case 5 Top 50 Agents (At 413th Step)

Fig. 4 The difference in the number of hub agents who are released information (Case 5)

chen.shuheng@gmail.com

information. At the same time, we would like to indicate the number of link that each agent has and their attitude score at 354th step and 413th step in Case 5. Figure 4 indicates the number of link that each agent has and their attitude score at 354th step and 413th step in Case 5. The result from simulation that information is released to top 20 hub agents and top 50 hub agents is shown in Fig. 4.

As shown in Fig. 4, in Case 5, the difference in the number of agents who receive information is obvious but the number of agents who have negative attitude score decrease very gradually.

Surprisingly, the difference in the number of agents who receive information is less ambiguous than that in Case 2. So why with such a sharp contrast, in Case 5, do the number of agents who have negative attitude score decrease very gradually? It is because network structure is changed frequently, every 13 steps, in Case 5. We might say that the difference between Case 2 and Case 5 stem from the difference in frequency of the change in the conformation of network. In Case 5, an agent does not continue to be hub agent. Therefore, released information does not expand.

5.2.3 Investigating the Reasons for the Differences from Agent's "Betweeness Centrality"

Then, we would like to examine Case 3 and Case 6. In these models, information is released to agents whom the value of betweeness centrality is high. In addition, as stated before, to release the information to agents who have many links have more effective than that to agents whom the value of betweeness centrality is high in decreasing the number of agent who have distrust in public pension system. Then, the result from simulation that information is released to top 20 betweeness centrality high agents and top 50 betweeness centrality high agents in Case 3 is shown in Fig. 5.

We examine Case 3 and Case 6 using three indices. That is, each agent's attitude scores, the number of links that each agent has and their betweeness centrality. Also, we will investigate the change of them before and after step when information is released. In Case 3, in the case that the number of agents who receive information is 20, many agents still have negative attitude score at 413th step. On another front, in the case that the number of agents who receive information is 50, many agents, almost all agents who have high betweeness centrality have positive attitude score at 413th step. It is showed that information is expanded steadily.

Then, we examine Case 6. As shown in Fig. 6, in Case 6, there is not much difference between the case that information is released to top 20 agents who have high betweeness centrality and that to top 50 agents. Even in this situation, many agents who have high betweeness centrality have positive attitude score at 413th step. However, in Case 6, the number of links that each agent has is smaller than that in Case 3, on the whole. In addition, as stated before, network structure is changed frequently in Case 6. We might say that the difference between Case 3 and Case 6 stem from the difference in frequency of the change in the conformation of network.

chen.shuheng@gmail.com

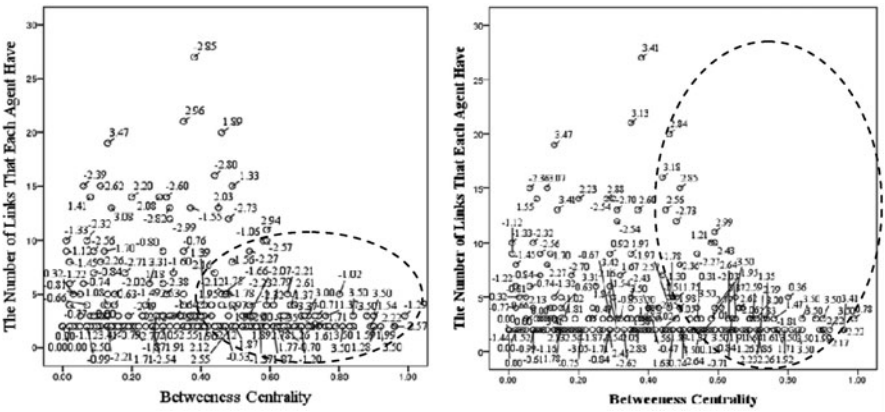


Fig. 5 The difference in distribution of agents (Case 3, At 413th Step); Top 20 agents (*Left*), Top 50 agents (*Right*)

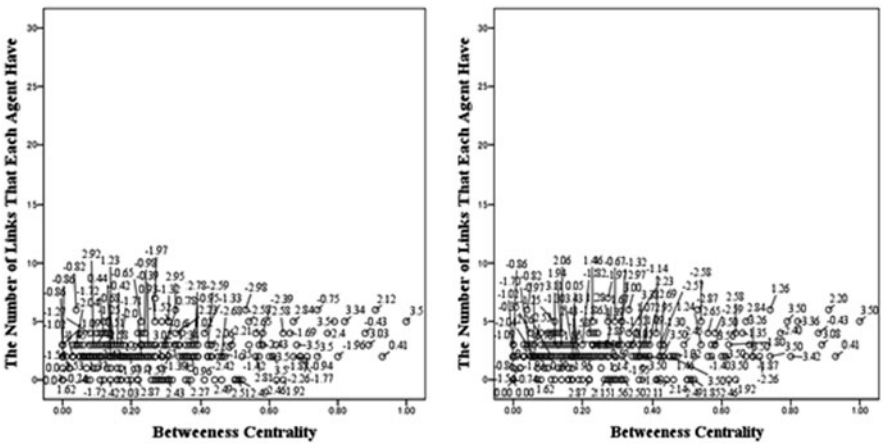


Fig. 6 The difference in distribution of agents (Case 6, At 413th Step); Top 20 agents (*Left*), Top 50 agents (*Right*)

5.2.4 The Relationships Between Information Flow and Pension Premium Fund

Now, we would like to examine the relationships between information flow and pension premium fund. It is indicated in Fig. 7. In our model, pension premium fund (P) is simply defined as below:

$$P = TAP - TPB, \tag{5}$$

where *TAP* is total amount of pension premiums paid by each agent aged 20 or over but under 60 and *TPB* is total amount of pension benefit received by agents aged 65

chen.shuheng@gmail.com

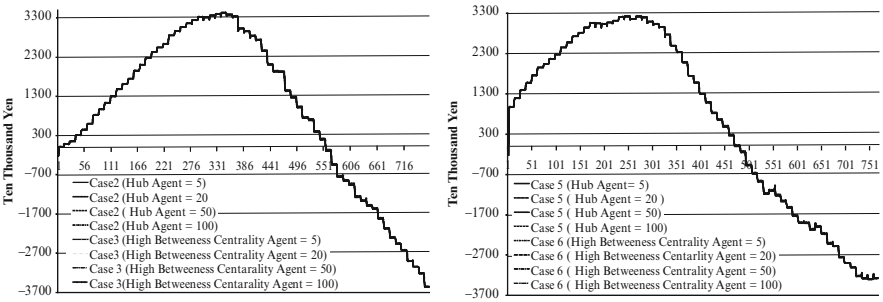


Fig. 7 Pension premium fund difference between Model 4 (*Left*) and Model 6 (*Right*)

or over. Of course, the amount of pension benefit received by each agent is changed in response to the total amount of pension premium that they have paid.[6]

There is a big difference between Model 4 and Model 6. Surprisingly, we cannot determine the distinction that is arisen by the difference in the number of agents who receive information. There are slight differences. The most notable differences between Model 4 (Case 2, Case 3) and Model 6 (Case 5, Case 6) are because network structure is changed frequently in Model 6. As mentioned before, in Model 6, many agent change link agent, frequently. Therefore, many agents' roles in the network are also changed frequently. In addition, no matter how much positive information is released, it is not propagated thorough network.

On another front, in Model 4, many agents' network is fixed to a certain extent. Therefore, many agents' roles in the network are also fixed. If an agent is hub agent at one time, he (or she) remains a hub agent at the other time. Then, information is propagated through the network. In fact, the step when pension premium fund go negative is about 555 in Model 4 and about 480 in Model 6. We can see big difference, which is depend on agent's network structure.

5.2.5 Additional Simulation

In previous section, in each model, we cannot determine the distinction that is arisen by the difference in the number of agents who receive information. It is because Japanese public pension system is composed by category I insured people, category II insured people and category III insured people. Category I insured people is composed by all registered residents of Japan aged 20 or over but under 60, except for the Category II or III insured people. The people include in the Category are mainly, self-employed people, freelance workers and students as well as their

[6] For example, in 2009, the amount of pension benefit received each people aged over 65 in each month is defined as follows: 792,000 yen (full benefit amount) $\times$ {(premium paid period) + (period of half exemption from premium) $\times 2/3+$ (period of total exemption from premium) $\times 1/3\}/40$ (the maximum participation year) $\times 12$.

chen.shuheng@gmail.com

spouses. By contrast, Category II insured people are composed by people enrolled in the Employees' Pension Insurance system or Mutual Aid Association (except for people aged 65 or over but under 70 who are eligible to receive pension due to old age or retirement). Their pension premium is automatically deducted from his (or her) wages. Category III insured people are composed by Category II insured people's dependent spouses aged 20 or over but under 60.[7]

In our model, many young agents are not hub agents. Information is released just one time. As time passes, the number of link that they have is increased. Therefore, there is a possibility that their attitude score does not changed by positive

Table 4 The difference in the number of agent who has distrust (Case 2)

	The number of steps							
	1st	100th	200th	300th	400th	500th	600th	700th
Information is not released	112	107	125	148	163	187	156	166
Information is released to top 20 hub agents	112	107	125	148	149	180	154	162
Information is released to top 20 "young" hub agents	112	107	125	148	134	152	106	89
Information is periodically released	112	98	63	68	37	23	17	15

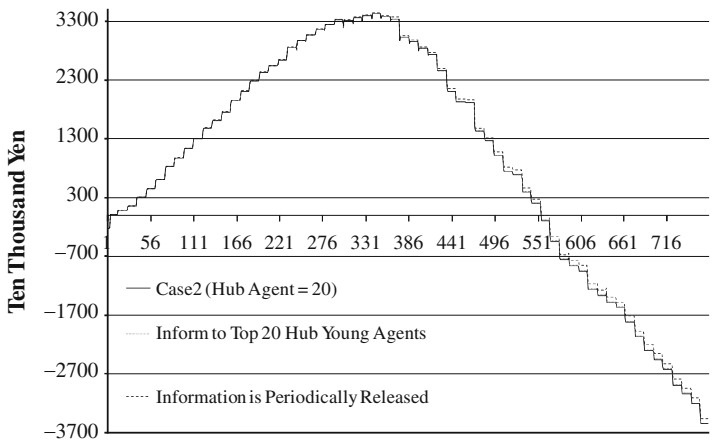


Fig. 8 The difference in releasing information and in pension premium fund (Case 2)

[7] For more detail on Japanese public pension system, see [10]. As described before, in our model, we defined one step as 1 month. So, agents get older every 12 steps. And they have life duration based on abridged life table. In Japan, the abridged life table is released at fixed intervals by Ministry of Health, Labour and Welfare. In addition, we set the start year of simulation at 2004.

chen.shuheng@gmail.com

information. For these reason, we will release information to Young (20–24 years old) Top 20 Hub Agents. In addition, we will periodically release Information. The results are shown in Table 4 and Fig. 8.

As shown in Table 4 and Fig. 8, releasing information to young agents have great effect in reducing agents that have distrust. However, the difference in pension fund is not large. The difference in pension premium fund is large thorough releasing information frequently.

6 Conclusion and Future Works

In this article, we verified a positive policy impacts that comes from decreasing people's distrust in public pension system. Especially, the difference in agent's network structure is key factor in the difference in pension premium fund. Moreover, the difference in the number of agent who receives information does not result in the difference in pension premium fund. We can see only slight differences. With that, we release information to young agents and repeat it. As a result, it is revealed that releasing Information over again have a profound effect on reducing distrust in public pension system and on pension premium fund. With releasing information to a limited extent and to a limited number of agents, there is limited effect on reducing agent's distrust. It implies that government should continuously improve pension system.

From now on, a continuous examination of the mechanism of people's network, e.g. the introduction of the other network indices, and its effect on public pension system would strengthen this proposition stated in this article. In addition, in this article, we do not verify negative policy impact because of the limitation on the number of pages that we can write. So, using our network models, we will verify a negative policy impacts that comes from increasing people's distrust in public pension system. It will differ in agent's behavior to positive information and negative one. So, we should study people's reaction to positive and negative information.

Acknowledgement We appreciate KOZO KEIKAKU ENGINEERING Inc. for providing Multi Agent Simulator (named "artisoc 1.0"). First author is partially supported by Grant-in-Aid for Young Scientists (B), MEXT (No. 21730179).

References

1. Murakami M, Tanida N How will the difference in people's network structure have impact on Japanese public pension system? In: 15th International conference on computing in economics and finance. http://editorialexpress.com/conference/CEF2009/program/CEF2009.html
2. Latané B, Morio H (2000) Maintaining diversity: simulating the role of nonlinearity and discreteness in dynamic social impact. In: Suleiman R, Troitzsch KG, Gilbert N (eds) Tools and technique for social science simulation, pp 196–217. Physica-Verlag, Heidelberg
3. Latané B (1981) The psychology of social impact. Am Psychol 36:343–365

chen.shuheng@gmail.com

4. Asch SE (1951) Effect of group pressure upon the modification and distortion of judgment. In: Guetzknow H (eds) Groups, leadership and men, pp 1–43. Carnegie, Pittsburgh
5. Abelson RP (1964) Mathematical models of the distribution of attitudes under controversy. In: Frederiksen N, Gulliksen H (eds) Contributions to mathematical psychology, pp 147–160. Holt, Rinehart & Winston, New York
6. Morio H (2003) Multi-agent DSIT simulation: reexamination of robustness of prediction. Simulation & Gaming 13(2):159–168 (in Japanese)
7. Latané B, Nowak A (1997) Self-organizing social systems: necessary and sufficient conditions for the emergence of consolidation, clustering, and continuing diversity. In: Barnett G, Boster F (eds) Progress in communication science persuasion, vol 13, pp 43–74. Ablex, Norwood, NJ
8. Barabasi A-L, Albert R (1999) Emergence of scaling random network. Science 286:509–512
9. Konno N, Machida T (2008) How-nual Visual Guide Book Yokuwakaru Fukuzatsu Network. Shuwa System Co., Ltd (in Japanese)
10. Social Insurance Agency. http://www.sia.go.jp/e/np.html

chen.shuheng@gmail.com

Identification of Voting with Individual's Feet Through Agent-Based Modeling

Rio Nishida, Takashi Yamada, Atsushi Yoshikawa, and Takao Terano

Abstract This paper describes an agent-based simulation model to analyze migration behaviors of individual in several political regions. The model was originally discussed by Tiebout in 1956 as "Voting with Feet," however, the validity of the model has not been examined very carefully. In the proposed agent simulation model, plural political decisions in each region are made by the corresponding regional government, and the inhabitants will vote the decisions based on their preferences. Both governments and individuals are modeled as decision making agents. The intensive simulation studies have revealed the emergence of decision groups and how the decisions have been made.

Keywords Agent-based modeling · Voting with the feet · Social systems · Multi-layer modeling

1 Introduction

Agent-based modeling (ABM) is a new approach in the field of study of social sciences. Using recent advanced computing powers, ABM generates an artificial society, and executes the evolutionary processes with several parameters in order to clarify the underlying principles of social, economical, political organizations and systems.

Among the pioneer works, Axelrod [1] has stated the ABM approach as follows: "*Although agent-based modeling employs simulation, it does not aim to provide an accurate representation of a particular empirical application. Instead, the goal of agent-based modeling is to enrich our understanding of fundamental processes that may appear in a variety of applications*." This requires adhering to the KISS principle, which stands for the army slogan "keep it simple, stupid".

R. Nishida (✉), T. Yamada, A. Yoshikawa, and T. Terano
Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Tokyo, Japan
e-mail: nishida@dis.titech.ac.jp; tyamada@trn.dis.titech.ac.jp; at_suhi_bar@dis.titech.ac.jp; terano@dis.titech.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_9, © Springer 2011

chen.shuheng@gmail.com

However, generally, social systems are more complicated and it is difficult for us to represent the social evolution process with a simple model. In the recent literature, Terano points out that it is necessary to build the model and employs simulation beyond KISS principle [1].

In case of modeling the process of the migration between regions as related work described in Sect. 2, it is necessary to consider both the interactions among spatial regions caused by the spatial selection and agents' interactions among different two or more spaces where the agent can move mutually.

Based on the related work, in Sects. 3 and 4, we investigate the properties of VWF through agent-based modeling techniques: (1) we have implemented a simulation system, and (2) Using the simulator, we have studied how the difference of policies of each local governments affect the migration of the inhabitants. In Sect. 5, then, we propose a new agent simulation model: Multi-Layer Modeling (MLM), in which the agents are able to move among each heterogeneous world in the simulation space. Using MLM, we will focus on the migrations between regions. The paper describes the concepts of Multiple-Layer Modeling, which will reveal the integrated process of the migration. In Sect. 6, we give some concluding remarks.

2 Related Work

2.1 Social Phenomenon of Migration

In the study of theoretical framework of migration in the economical field, a lot of researches have focused on the mechanism that people make decisions based on their movements [7, 15].

Especially, these works have paid attention to the regional difference of working environments and living standards, such as income, wage and unemployment rate, and then verify their influences on people's movement. Such works are taken as typical examples of economic analysis on the migration. These ideas have shown that the difference of the environment between regions gave the important effect to the macro migration through the people's decision making of regional movement of the individual.

A Japanese report [5] has conducted survey studies to uncover the migration mechanisms. They have reported that the wealth gap between regions has a significant influence on the migration in 2005. Ohta [8] has pointed out the capability that the individual with high human capital level move from the rural to cities.

From the point of view of the macro influence on the region caused by the movement, Tachibanaki [11] describes the influence on the attractiveness of region (commercial accumulation, transportation convenience, entertainment environment, employment opportunity) and congestion degree (population, house cost, crime rate, pollution, and various prices of city).

Moreover, as relations of the local government that manages the region to the decision making, Tiebout [14] proposed the concept of "*Voting with Feet*" (VWF)

chen.shuheng@gmail.com

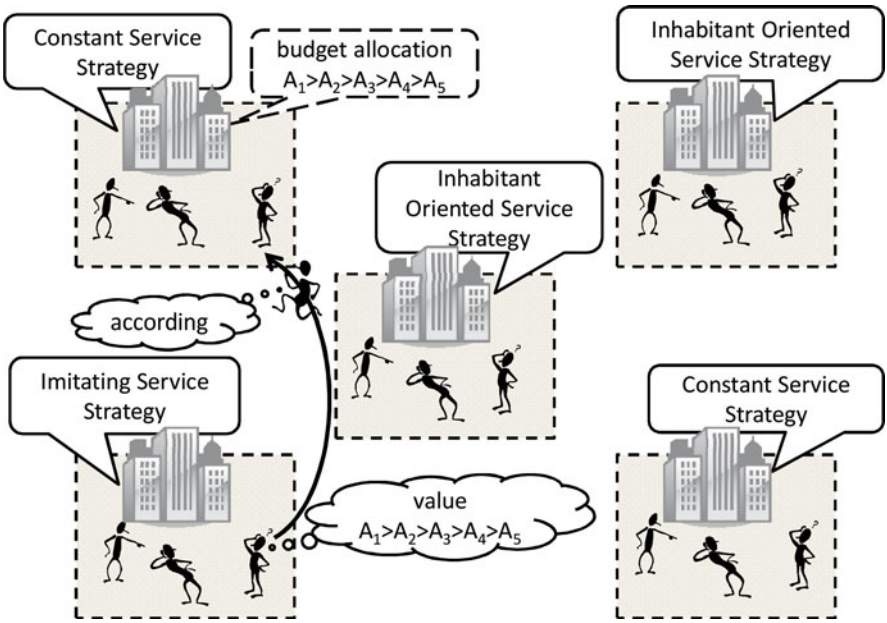


Fig. 1 Outline of the agent-based VWF model

to give a theoretical foundation of changes of local governance. VWF means such mechanisms that (1) When there are a lot of local governments in a country, (2) Each local government proposes its own tax system and political service systems as its policies, (3) Based on the proposals, each inhabitant selects his/her best local government and move to the place, (4) Local governments compete with each other to improve the policies, and (5) As a result, local governments will have the incentives for effective political services. There are a lot of experimental study where it was analyzed whether the movement based on the Tiebout hypothesis was able to happen [2, 4].

The process related to the demographic shift is very complex as enumerated it in the above-mentioned. The reason for the decision making of the municipality that manages the environment, the character, and the region in the region where people reside is that the important effect is caused in the occurrence of the movement actually as shown in Fig. 1 besides many of the movement factor, and, at present, the research that treats these processes that the demographic shift causes integrated is hardly seen.

2.2 Studies About VWF

There are a lot of researches in the literature referring to the Tiebout work. Among them, Sharp [10] has estimated that the 2.4% of the population in Kansas in the

chen.shuheng@gmail.com

United States makes migration decisions from the concept of VWF. In Lowery et al. [6] and Percy et al. [9] have conducted survey studies to uncover the migration mechanisms. They have reported that there are contradictive results between the VWF model and survey data. Although the VWF model states the migrations of inhabitants are caused by the complains about taxes they must pay and the levels of services the local governments give, their investigations have not succeeded in explaining the mechanisms.

In this paper, we will focus on the mechanisms described by the original VWF: both the policy making of local government and decisions by the inhabitants affected by the local policies. We proposed agent-based simulation model with utilities of migration of the inhabitants and policies of local governments. We will investigate the competitive conditions on plural local governments to acquire more inhabitants in the region.

3 Model Description

Our agent-based model consists of a regional system with local government agents $LG = \{lg_{i,}|i = 1.,\ldots,m\}$ and inhabitant agents $C = \{c_{ik}|k = 1.,\ldots,n\}$. Government agents provide common services with inhabitants agents. This decisions are made based on the budgets proportional with the inhabitants they have. The governments agents also open their policies to the inhabitants. In order to have more inhabitants, the governments make political decisions how to distribute the budgets. On the other hand, the inhabitants agents have their own preferences about the political decisions. Based on the preference, they decide where to move.

3.1 Government Agent Model

3.1.1 Basic Concepts of the Model

Each government agent is represented as an $lg_{i,} = \{A_i,\ T_i,\ B_i\}$. At each simulation step, a government agent $lg_{i,}$ shows its five kinds of local common service set $A_{ij}(t) = \{A_{ij}|j = 1,\ldots 5\}$ and the local tax $T_i(t)$ to inhabitant agents c_k. $A_{ij}(t)$ provided by the $lg_{i,}$ is a very abstract concept, but it is corresponding to such social capital services as medical, educational, childcare, transportation, residence, security services as police and fire stations. We further assume that the tax $T_i(t)$ is constant during the simulation.

Each government agent $lg_{i,}$ gathers T_i from a inhabitant agent c_k. The government gets the profit $B_i(t)$ and its changing rate $BS_i(t)$ determined by the following equations:

$$B_i\,(t) = \mathrm{T}_i\,(t)^{*} n(C(t)),$$

chen.shuheng@gmail.com

and

$$BS_i(t) = B_i(t) - B_i(t-1).$$

1. When the service level increases $(BS_i(t) > 0)$, then $A_{ij}(t+1) = A_{ij}(t) + \alpha_j{}^* RA_{ij}{}^* BS_i(t)$, and
2. When the service rate decreases $(BS_i(t) \leqq 0)$, then $A_{ij}(t+1) = A_{ij}(t) + RA_{ij}{}^* BS_i(t)$;

where, α_j is a constant parameter to represent the efficiency of service. In our simulation, we set α_j is equal to 0.7.

3.1.2 Budget Distribution Strategies

A government agent lg_i, makes political decisions about the budget distribution in order to have more inhabitants. We set the following three strategies:

1. Constant Service Strategy (CSS): Set the ratio for each service constant.
2. Inhabitant Oriented Service Strategy (HSS): Based on the average of the weights ω_j determined by the following equation, which represent the preference of inhabitants about each service, set the ratio for the services.
3. Imitating Service Strategy (ISS): In each several simulation steps, set the best strategy of the government, which shows the best increase of the number of inhabitants during the steps in the whole area.

3.2 Inhabitant Agent Model

An inhabitant agent is represented as $c_k = \{V_k, P_k\}$, where V_k represents the attractiveness of the region the agent c_k has, and P_k is probability of migration of c_k to the other regions. The attractiveness V_k, is determined as follows:

$$V_k = \sum_{j=1}^{5} \omega_j A_j - \omega_6 T_j$$

where, the preference weight ω_j is the importance of the inhabitant to each service, and $\Sigma_{j=1}{}^6 \omega_j = 1$.

In each simulation step, an inhabitant agent c_k calculates the attractiveness from A_j and T_i the local government $lg_{i,}$ shows, then compare the best attractive value V_{maz}, in the regions and the current attractive value V_{now}. The inhabitant agent c_k moves from the current region to the other one with the probability P_k determined by the following equation:

$$P_k{=}1/(1 + (1/P_0 - 1)exp(-r(V_{maz} - V_{now}))).$$

Changing the constant values of r and P_0, the simulator enables us to investigate the utility and easiness of the migration of inhabitants. The proposed model does not

chen.shuheng@gmail.com

include the information of the other properties such as age, sex, psychological, and economical conditions. These properties will be included into our model.

3.3 How the Agent-Simulator Works

We set initial values of inhabitant agent c_k with uniform random values ω_j, r, and P_0, where $0.000 < r,\ P_0 < 0.001$, and set the inhabitants agents randomly in the regions. The simulation step is depicted in Fig. 2 and summarized as follows:

1. A local government lg_i, gets the information the amount of budgets the inhabitant agents want, the number of inhabitants, and the number of the increase/decrease of the population.
2. The government agent lg_i, provides the inhabitant agents with $T_i(t)$ and $A_{ij}(t)$.

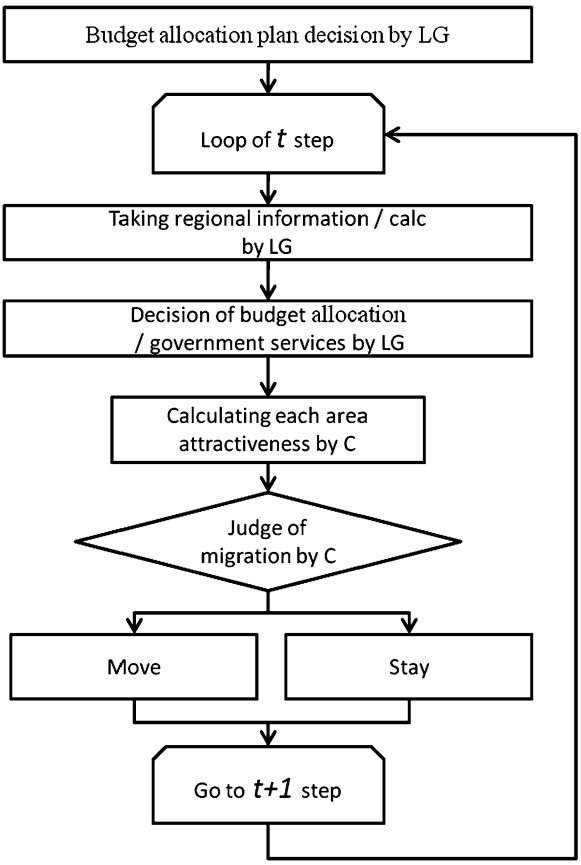


Fig. 2 Flow of the simulation model

chen.shuheng@gmail.com

3. An inhabitant agent c_k makes the decision of migration with the probability P_k based on the values of V_k and the information which the government agent lg_i, provides to c_k.
4. Based on the value of $B_i(t)$, the government agent lg_i determines $A_j(t+1)$ and $T_i(t+1)$.

4 Experiments and Discussions

Using the simulation model, we have conducted the following experiments. We set five local government agents and 5,000 inhabitant agents. At the first step, the same numbers of inhabitant agents (1,000 agents) are located in each region. We assume that one step of the simulation is corresponding to 1 year, and we will observe the changes of 30 years in each simulation. In order to evaluate the effects of different political strategies of the government agents, we set the six scenarios in Table 1 and the common parameter values in Table 2.

Table 1 Six simulation scenarios

Scenario	lg_1(A)	lg_2(B)	lg_3(C)	lg_4(D)	lg_5(E)
Case1	Fixed	Fixed	Fixed	Fixed	Fixed
Case2	Residents-oriented	Residents-oriented	Residents-oriented	Residents-oriented	Residents-oriented
Case3	Imitation	Imitation	Imitation	Imitation	Imitation
Case4	Imitation	Residents-oriented	Residents-oriented	Fixed	Fixed
Case5	Imitation	Imitation	Residents-oriented	Residents-oriented	Fixed
Case6	Imitation	Imitation	Residents-oriented	Fixed	Fixed

Table 2 Common parameter values for the experiments

Parameters	Description	Values
$\|\mathbf{LG}\|$	Number of local governments	5
N	Number of inhabitant Agents	5,000
$\|\mathbf{A}_j\|$	Number of local common service set	5
α_i, β_i	Utility and easiness of the migration of inhabitants	0.000–0.001
γ_j	Constant parameter to represent the efficiency of service	0.7

chen.shuheng@gmail.com

4.1 Experimental Results

In this subsection, we will show summarized statistical results of simulation runs and typical simulation results.

4.1.1 Statistics of the Experiments

For each scenario in Table 1, we have carried out 30 simulation runs. Figure 3 and Table 3 respectively summarize the average numbers of migrations of inhabitants and their standard deviations after 30 year simulation steps.

From Experimental results, Case 1, where all the local governments keep Fixed Service Strategy with uniformly random budget distributions at the initial step, has shown the highest number of migrations. This means that distributions of the budget causes the encouragement of individual decisions of migrations of the inhabitant agents. Figure 4 shows the changes of inhabitants. The inhabitant agents move frequently.

On the other hand, in Case 2 and Case 3, we have observed the smaller numbers of the migrations. Figures 5 and 6 show the simulation results. When the governments get the larger number of agents at earlier steps of the simulation, they tend to be only winners. In Case 2, we have set the conditions that local government

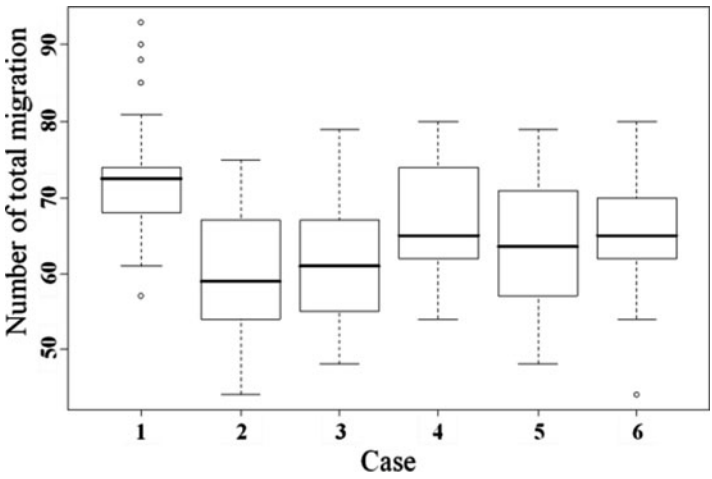


Fig. 3 Number of local migrations in each case

Table 3 The numbers of migrations and their standard deviations

Scenario	Case1	Case2	Case3	Case4	Case5	Case6
Mean	72.300	60.167	61.467	66.767	63.900	65.667
SD	8.655	7.571	8.170	7.417	8.450	7.563

chen.shuheng@gmail.com

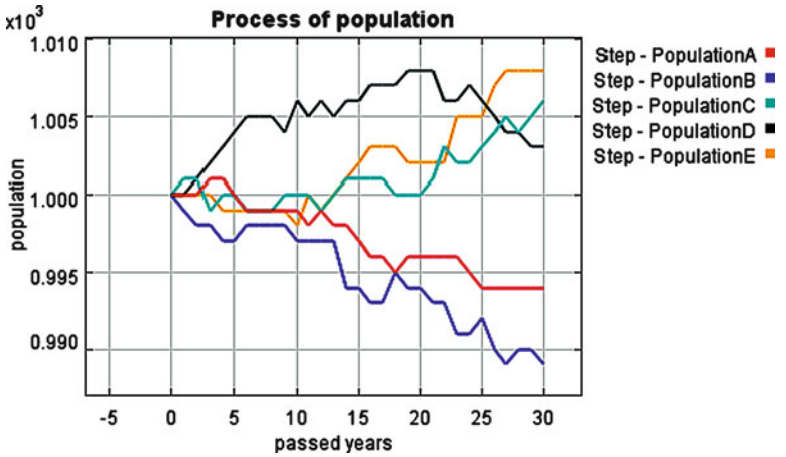


Fig. 4 Results of Case 1 (color figure online)

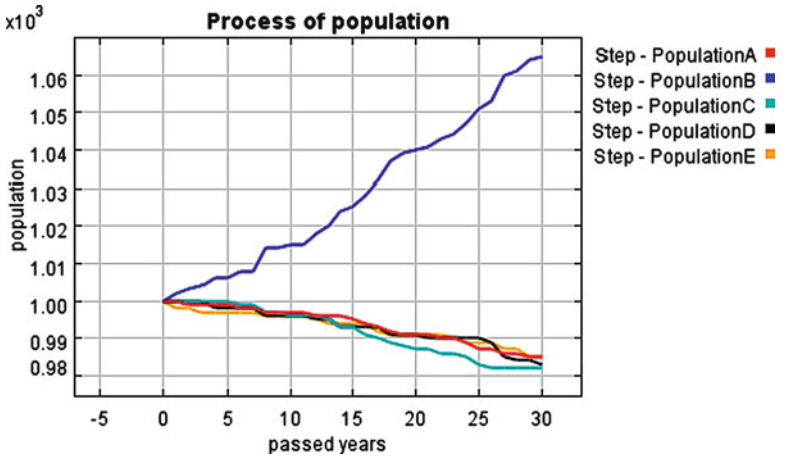


Fig. 5 Results of Case 2 (color figure online)

agents accept all the needs of the inhabitants. Therefore, the political strategies of the governments become to be similar to each other. In Case 3, because the local government agents have employed ISS, they have imitated the strategy, which wins the highest migrations of inhabitants. This has encouraged the uniform policies among the local governments.

In Cases 1, 2, and 3, the average numbers of migrations of inhabitants tend to be smaller, when the simulation steps proceed, because in such cases, all the governments will take the same strategies.

In Cases 4, 5, and 6, in order to uncover which strategies would be better to get the larger number of inhabitants, we will show the statistical results how the number of inhabitants would change. In these cases, the strategies of the government

chen.shuheng@gmail.com

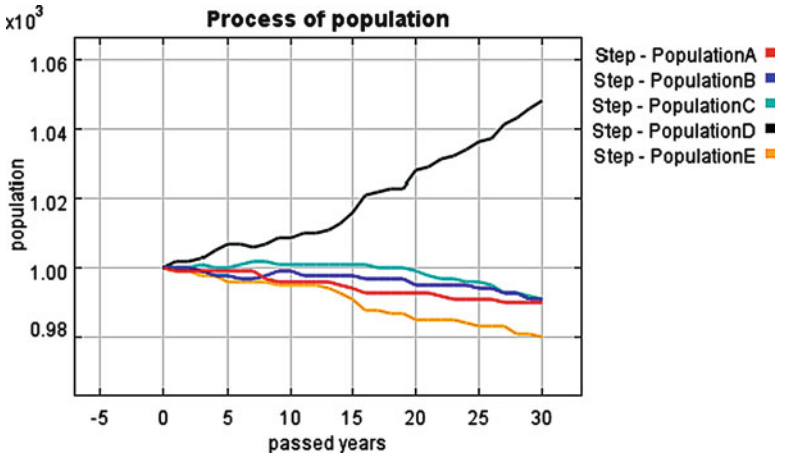


Fig. 6 Results of Case 3 (color figure online)

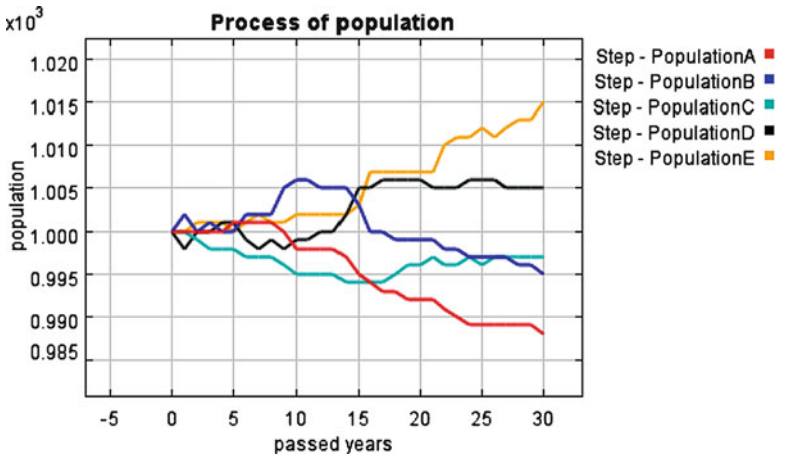


Fig. 7 Results of Case 4 (color figure online)

agents are different from each other. From Figs. 7–9 and Tables 4–6, government agents with FSS tend to have the larger numbers of inhabitants. The phenomena are able to be explained that (1) inhabitants with biased preferences will migrate to the government with the unique policies of FSS, and that (2) governments with ISS will change the strategies to the better ones, where the governments take FSS, however, they have failed the strategies because the number of habitants have been already the smaller (Tables 4–6).

chen.shuheng@gmail.com

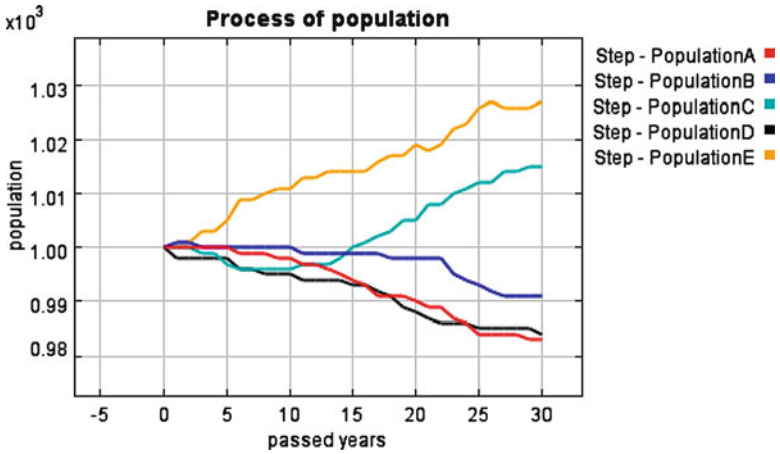


Fig. 8 Results in Case 5 (color figure online)

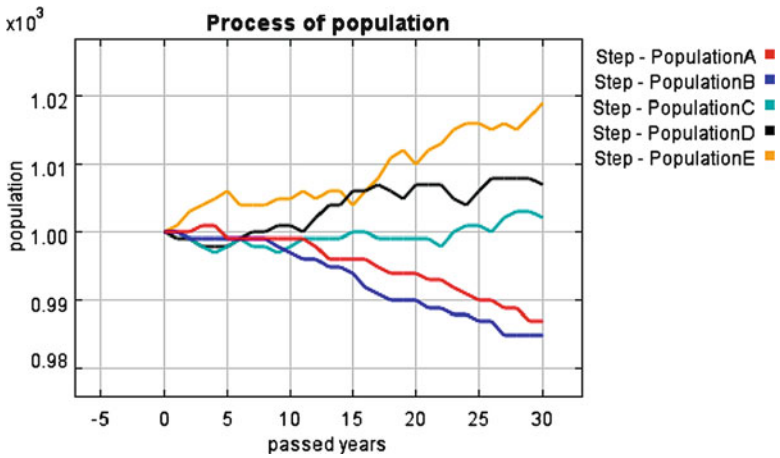


Fig. 9 Results in Case 6 (color figure online)

Table 4 Changes of inhabitants in each area of Case 4

Scenario	lg_1(A)	lg_2(B)	lg_3(C)	lg_4(D)	lg_5(E)
Mean	−14.933	−9.667	−8.467	16.867	16.200
SD	3.982	9.072	10.708	8.123	7.341

Table 5 Changes of inhabitants in each area of Case 5

Scenario	lg_1(A)	lg_2(B)	lg_3(C)	lg_4(D)	lg_5(E)
Mean	−11.700	−13.667	−3.867	2.867	26.367
SD	10.373	6.5196	15.527	15.780	7.735

chen.shuheng@gmail.com

Table 6 Changes of inhabitants in each area of Case 6

Scenario	lg_1(A)	lg_2(B)	lg_3(C)	lg_4(D)	lg_5(E)
Mean	−15.700	−15.200	−2.233	18.333	14.800
SD	3.535	5.416	7.482	8.949	9.215

5 Proposal of Multiple-Layer Modeling

5.1 Issues of the Current VWIF Model

In the current VWIF model, a region concerned is never considered about the change in an immediate regional environment by the interaction and the movement of the residents, although these residents move mutually. The VWIF model excludes the increase of the amount of the budget caused by the changes of the number of residents. Therefore, the resident's movement is too sensitive with each public service. The service levels are decided by the local governments based on the budget allotment ratio.

However, the movement, which actually occurs in the real world, is caused by residents' interactions for example, economic activity, communications, and traffics in the region. In the real societies, such municipalitical phenomena usually happen as management of the environment conditions and/or the traditional characteristics of the region. As a result of such complex activities, an integrated demographic shift will occur. To cope with the issues, we will extend the model to Multiple-Layer Modeling (MLM).

5.2 Framework of MLM

Figure 10 depicts the conceptual framework of MLM. Compared with most of conventional agent-based models so far, in the artificial social environment, the model has several layers or worlds.

Each world is the place on which the agents make some activities. The one world is corresponds to Sugarscape [3], which consists of a lattice structure similar to the cellular automata The plural worlds are connected in the forms of the network structure. The agents in the one specific world are able to move to the other world based on the network structure. In order to simulate such social phenomena related to the demographic shifts described in the previous VWF model, we only determine (1) the agents' rules to select the world based on the rules of agent's interaction in one world.

chen.shuheng@gmail.com

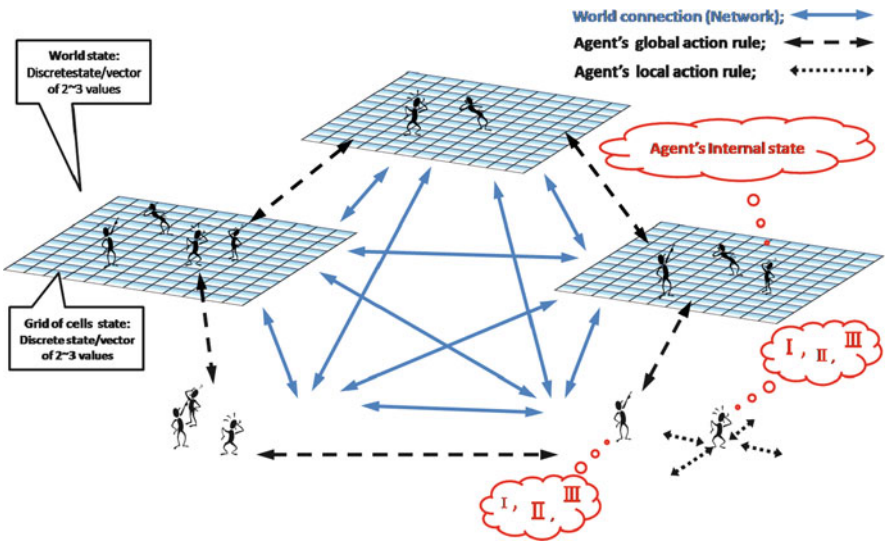


Fig. 10 Conceptual framework of multiple-layer modeling

5.3 On Agents' Action Rules in the MLM Environment

In the MLM environment, the agents make some (living) activities in each layer or a world using their decision rules similar to the original VWF model. In addition to these actions, the agents are able to move among several layers or the worlds. This means that the agents have the additional decision rules to select the place on which they act.

The features of the simulation based on the MLM agents are summarized as follows: (1) Each individual agent has an internal state in a micro viewpoint, it acts to adjust the world autonomously, and it moves to the point where the information is exchange and the problem is solved, and (2) As a result, macro-level characters of the multi layered system emerge, and (3) The layered system environment with the agents and the agents' micro-level activities form so-called micro–macro link.

6 Concluding Remarks

This paper has presented a novel agent based simulation model to investigate the effects of "Voting with Feet" mechanisms proposed by Tiebout. The model includes local government agents, which determine budget policies about political services, and inhabitant agents, which makes migration decisions based on the service levels and tax systems of the governments. The simulation results are summarized as follows: (1) unique policies of each local government will results in

chen.shuheng@gmail.com

the encouragement of migrations of inhabitants, and convergence of the policies of the governments discourages the migrations; and (2) Cases with various strategies, the imitation service strategy will cause the decreases of inhabitants. These results suggest the deeper understandings about "Vote with Feet" mechanisms and the applicability of agent-based modeling to political decision making.

Furthermore, to cope with the defects of the model, we have describes the framework of Multiple-Layer Modeling. MLM is characterized by the integrated treatment of it of the process of the demographic shift, and the methodology and the meaning were considered.

Future work related on the research includes (1) validation of parameter settings, (2) the introductions of interactions among inhabitants on the decision making of migrations, (3) landscape analyses among various government policy scenarios, and (4) grounding the simulation results to the real world phenomena. The results will be presented elsewhere.

References

1. Axelrod RM (1997) The complexity of cooperation: agent-based models of competition a collaboration. Princeton University Press, Princeton
2. Barrow L School choice through relocation: evidence from the Washington, D.C. Area. J Public Econ 86:155–189
3. Epstein J, Axtell R (1996) Growing artificial societies: social science from the bottom up. Brookings Institution Press, The MIT Press, Cambridge
4. Hoyt W, Rosentahl S (1997) Household location and Tiebout: do families sort according to preferences for locational amenities? J Urban Econ 42:159–178
5. Japan Institute for Labour Policy and Training (2006) Statistical survey on labor dispute. The Japan Labor Flash, No.71
6. Lowery D, Lyons WE (1989) The impact of jurisdictional boundaries: an individual-level test of the Tiebout model. J Polit 52(1):73–97
7. Matsushita K (1982) Attempt of theoretical approach of demographic shift. Southeast Asia Res 20(2):253–259
8. Ohta S (2006) Interregional earnings differentials and the effect of hometown on earnings in Japan. In: ESRI2006 international collaboration projects paper, pp 69–99
9. Percy SL, Hawkins BW (1992) Future tests of individual-level propositions from Tiebout model. J Polit 54(4):1149–1157
10. Sharp EB (1989) A capitalization approach to fiscal incidence at the local level. Land Econ 65(4):259–375
11. Tachinabaki T, Urakawa K (2009) Japanese regional gaps – factor of demographic shift. In: The Keizai seminar, pp 110–122
12. Terano T (2007) Exploring the vast parameter space of multi-agent based simulation. In: Antunes L, Takadama K (eds) Proc. MABS 2006, LNAI 4442, pp 1–14. Springer
13. Terano T (2008) Beyond the KISS principle for agent-based social simulation. J Socio-Informatics 1(2):175–187
14. Tiebout CM (1956) A pure theory of local expenditures. J Polit Econ 64(5):416–424
15. Todaro M (1969) A model for labor migration and urban unemployment in less developed countries. Am Econ Rev 59:138–148

chen.shuheng@gmail.com

Part V
Agent-Based Modeling of Good Societies

chen.shuheng@gmail.com

Communities, Anti-Communities, Pan-Community as Social Order

Yutaka Nakai

Abstract A society of agents who can freely attack others or not inevitably evolves into a battling society (a "war of all against all"). We investigated whether the Friend Selection Strategies based on Attribute and Reputation in Group (FSS-ARG) lead to the emergence of social order. FSS-ARG require an agent to evaluate whether others are his "friends" or "enemies", based on whether others were peaceful or hostile to his group and whether others have the same attribute as his group or not. We carried out evolutionary simulations with FSS-ARG. As a result, we found that four types of social states, what we call "battling society", "communities", "anti-communities" and "pan-community", have emerged. For example, the "communities" consist of two mutually hostile communities, in each of which all members have the sane attribute and are friendly to all other members. So, the "communities" can remind us of Max Weber's "in-group/out-group morality".

Keywords A friend and an enemy · Community · Reputation · In-group morality · Evolutionary simulation

1 Introduction

How does social order emerge among individuals acting freely? This fundamental question in sociology was named the "problem of order" by T. Parsons [14]. For example, if people, as selfish individuals, could freely choose to attack or not to attack others, the society would inevitably change into a "war of all against all" (Hobbesian state). However, the real society is not in such a state and seems to have order. As an another example, if people could freely choose whether to pay taxes or not, everyone would want to receive public services without paying taxes

Y. Nakai (✉)
Faculty of Systems Engineering and Science, Shibaura Institute of Technology,
307 Fukasaku, Minuma-ku, Saitama-City, Saitama 337-8570, Japan
e-mail: nakai@shibaura-it.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*,
Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_10, © Springer 2011

chen.shuheng@gmail.com

("free riders"), and as a result, nobody could receive public services. However, in the real society, public services are maintained by tax revenues. As is well known, a variety of solutions to this problem have been proposed. Typical examples are "central authority and social contract" [5], "institutionalization and internalization of norms" [14], "system rationality and reduction of complexity" [8], "mutual understanding based on communicative reason" [4], and "trust based on rational expectation" [3].

The Hobbesian state mentioned above has sometimes been described as a problem of generalized exchange (indirect reciprocity) under the 2-persons Prisoner's Dilemma (2PD). And, as solutions to the 2PD problem, a variety of "reputation theories" have been presented recently. These theories evaluate whether the other is good or bad on the basis of reputations about the other, and intend to establish social order by giving advantage to good persons or keeping them away from bad persons. They are interpreted as an altruistic strategy toward a selected good man, and are sometimes called "discriminator strategies" (DISC). As typical examples, we can see the following theories, "in-group altruistic strategy" [15], "imaging score strategy" [12, 13], "standing strategy" [7], and "strict discriminator strategy" [16].

Especially, Nakai and Muto [10, 11] studied an emergence of community (social order). In their study, strategies evaluating who is a "friend" and who is an "enemy" are called "Friend Selection Strategies" (FSSs). They were inspired by Carl Schmitt's "friend and enemy", but Carl Schmitt's "friend and enemy" and Nakai & Muto's are different. C. Schmitt's "friend and enemy" were defined based on the race and religion, while Nakai & Muto's were independent from them. That is, Nakai & Muto's "friend and enemy" are quite liberal concepts.

As one of FSSs, which they call "us-TFT strategy", Nakai & Muto proposed the strategy which requires agents to regard the other who attacked "us" as an enemy and the other who didn't as a friend. They pointed out that us-TFT leads to social order via the spontaneous formation of a community like a collective security system. However, the us-TFT strategy has some weak points, which are, for example, that the strategy doesn't work well in a large-scale society, and that it can't describe the dynamics between different groups. The former weak point is related to the concept of Nakai & Muto's "us". Each us-TFT agent's "us" ("me" plus friends) is mutually different and an us-TFT agent evaluates the other on the basis of the other's actions toward the agent's "us". And, an us-TFT agent has to watch all others' actions toward him and his friends. Therefore, the larger his "us" becomes, the more difficult it is to watch all his friends' experiences, because of his limited capability.

2 The Friend Selection Strategies based on Attribute and Reputation in Group (FSS-ARG)

In order to overcome the weak points, we introduced the Friend Selection Strategies based on Attribute and Reputation in Group (FSS-ARG). The strategies are defined based on who is a friend and who is an enemy for "our group". We suppose that a group member informs other members of his own experience about who attacked

chen.shuheng@gmail.com

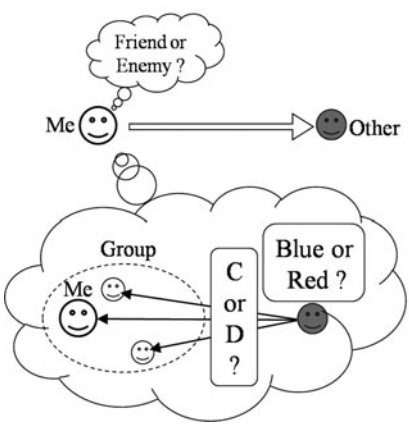
him, and participates in making reputations about others. That is, FSS-ARG don't require a member to watch a lot of persons' experience. So, FSS-ARG agents don't suffer from the us-TFT's problem in a larger society.

Following the concept of group, it seems to be natural for people, in evaluating others, to have the following viewpoints such as an attribute and an action.

(VP1) Attributes; "Is he a member of our group?" ("Does he have the same attribute as ours?")
(VP2) Actions; "Was he peaceful or hostile to our group?"
Then, from the viewpoints, we have an idea of four categories for others, as follows.
(OC1) The other who has the same attribute as ours, and attacked our group in the past (previous turn).
(OC2) The other who has the same attribute as ours, and didn't attack our group in the past.
(OC3) The other who has the different attribute from ours, and attacked our group in the past.
(OC4) The other who has the different attribute from ours, and didn't attack our group in the past.

By assigning a friend ("f") and an enemy ("e") to each category, we have 16 $(=2^4)$ types of FSS-ARG strategies, as theoretical ones. Here, we express FSS-ARG strategies by the notation of XXXX, where X is "f" or "e" and each position of four X corresponds to above $(\mathrm{OC1}) \sim (\mathrm{OC4})$ respectively (Fig. 1).

For example, a feee agent regards others who have the same attribute as his group and attacked his group as enemies, and others who the same attribute as his group and didn't attack his group as friends, while he regards anybodies who have the different attribute from his group as enemies, regardless of whether they attacked his group or not. And a fefe agent regards anybodies who attacked his group as enemies, and anybodies who didn't as friends. That is, his evaluations are independent of others' attributes.



Definition of 2^4 FSS-ARG strategies

Attribute \ Action	Attack (C)	No Attack (D)
Same	friend or enemy	friend or enemy
Different	friend or enemy	friend or enemy

Ex. feee strategy

Attribute \ Action	Attack (C)	No Attack (D)
Same	f	e
Different	e	e

Fig. 1 Friend selection strategies based on attribute and reputation in group (FSS-ARG)

chen.shuheng@gmail.com

3 Evolutionary Simulation of Peace

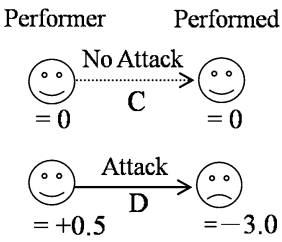
In order to examine whether FSS-ARG result in social order (peaceful state), we constructed an artificial society and carried out evolutionary simulations. In the artificial society, agents play the following "battle game".

(BG1) N agents play the game.
(BG2) Each agent has a social perception of who is a friend and who is an enemy for him.
(BG3) In one battle game, each agent (agent A) meets M other agents at random. (M stands for the "matching number" of one agent in one game.) Each agent (agent A) interacts with the other agent (agent B) one by one.
(BG4) A (as a performer) attacks or doesn't attack B (as a performed) on the basis of A's social perception of B. A attacks B if B is A's enemy, and A doesn't attack B if B is A's friend.
(BG5) If A attacks B, A obtains a payoff of 0.5 and B loses a payoff of 3.0. Conversely, if A doesn't attack B, both obtain and lose nothing (Fig. 2).

A typical instance of the battle game is a burglary. The performer obtains a payoff of 0.5 by stealing the performed's property, while the performed loses the same amount. The performed also suffers a loss of 2.5 due to physical and/or mental damages. To examine the result of the battle game, we assume two agents who interact reciprocally. Then, the payoffs in Fig. 2 result in the two-agent payoff matrix shown in Fig. 3 and, as seen easily, it is just a typical payoff matrix of the 2PD problem (the battle game can seen generalized exchanges under 2PD problem). This means that a battle game society has to fall into a battling society.

Here, we should note that the game's definition is incomplete – it doesn't include how to determine who is a friend and who is an enemy. Without this piece, an agent cannot interact with anybody at all. Then, agents adopt a friend selection strategy like FSS-ARG etc and play the battle game.

Now, we consider what strategies should be investigated in our evolutionary simulation. Here, let us remember that our essential interest is the effect of code, "a friend and an enemy", on the emergence of social order. Therefore, the code of a friend and an enemy "for me" seems the most simple and basic one, while



Action \ person	Performed Performer
No Attack (C)	0 0
Attack (D)	−3.0 +0.5

Fig. 2 Performer's and performed's payoffs

chen.shuheng@gmail.com

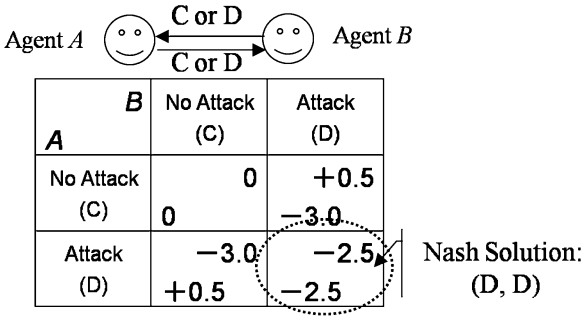


Fig. 3 Battle game's payoff matrix

FSS-ARG code is a friend and an enemy "for our group". From the viewpoint, Nakai and Muto [10] proposed what they called "My experienced-based Friend Selection Strategies" (MFSSs), as follows.

(MFSS1) ALL_D: the agent regards anybody as an enemy.
(MFSS2) me-TFT: the agent regards anybody who attacked himself as an enemy and anybody who did not as a friend.
(MFSS3) me-CWD: the agent regards anybody who attacked himself as a friend and anybody who did not as an enemy. (CWD stands for a "coward.")
(MFSS4) ALL_C: the agent regards anybody as a friend.

As seen easily, FSS-ARG are the group-based friend selection strategies and MFSSs the individual-based ones. And FSS-ARG can be regarded variants of MFSSs, because FSS-ARG are the complicated strategies made from MFSSs as basic strategies. So, we carried out the basic simulations with only MFSSs, and did also another simulation by adding FSS-ARG to MFSSs. The reason why strategies to be investigated are not limited to only FSS-ARG is that it seems strange to exclude a prior the basic strategies, MFSSs, from the simulation.

Now, the latter simulation with both FSS-ARG and MFSSs has the following merit. Even if we carry the simulation with only FSS-ARG except ffff & eeee, and found out a social order due to a formation of community, nobody will be surprised, because all these strategies are group-based ones. Therefore, the simulation with both FSS-ARG & MFSSs gives agents a chance to select strategies except group-based strategies. It assure us of the validity to point out the emergence of social order due to a community.

In sum, we adopted 18 strategies except in the latter simulations (FSS-ARG and MFSSs overlap because both include ffff (ALL_C) and eeee (ALL_D). 16 FSS-ARG + 4MFSS − 2 overlaps = 18 strategies).

The society is composed of a number of agents. Each agent has his own strategy and inherent attribute (blue or red). Each simulation run is composed of a sequence of turns, and each turn consists of five phases: reputation, perception, action, selection, and mutation (Fig. 4).

In the reputation phase, each agent informs his group of who attacked him and the group makes the reputation of who is a friend or an enemy for the group.

chen.shuheng@gmail.com

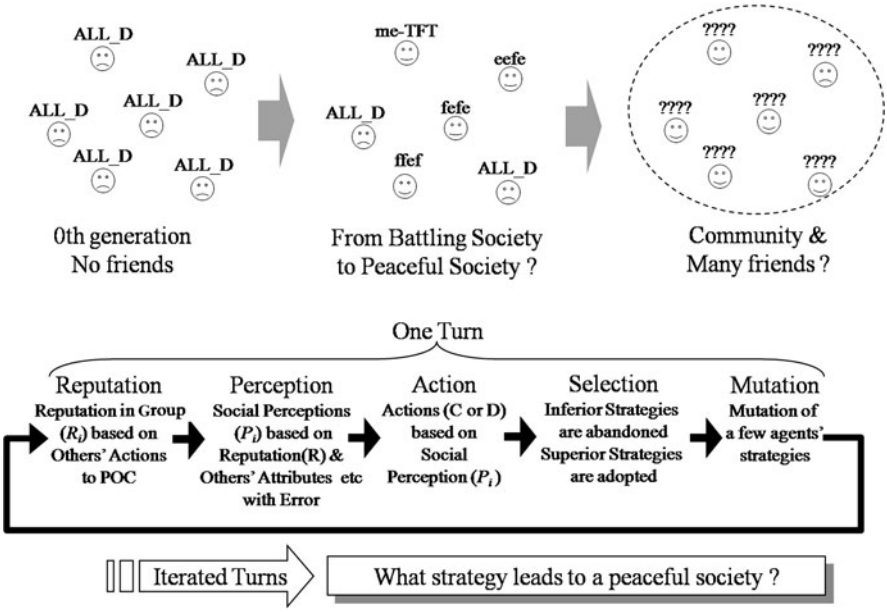


Fig. 4 Evolutionary simulation of peace

In the perception phase, each FSS-ARG agent updates his social perception of "who is a friend and who is an enemy", on the basis of his own FSS-ARG strategy and the others' reputations in his group, with an occasional error noted by the "perception error rate" (μ_p). Each MFSS agent updates his social perception, on the basis of his own MFSS strategy and others' actions toward him. In the action phase, each agent plays the battle game following his updated perception. In the selection phase, the lower $R\%$ agents in payoffs abandon their strategies and adopt that of agents whose results were highest (We call R the "reflection ratio"). In the mutation phase, a few agents are selected at random based on the "strategy's mutation rate" (μ_s). They abandon their strategies and adopt one of the 18 strategies at random.

At the 0th turn, all agents are assumed to have the ALL_D strategy and to perceive that all other agents are enemies, which expresses the state of "war of all against all." After many iterated turns, the superior strategies survive and the inferior ones fade away. We are interested in examining what strategy survives and leads to social order.

4 Emergence of Four Types of Social States

We carried out two simulations to investigate the effect of "a friend and an enemy for group" on the emergence of social order. As mentioned earlier, one examines what society evolves from only MFSSs and the other examines what society evolves from

chen.shuheng@gmail.com

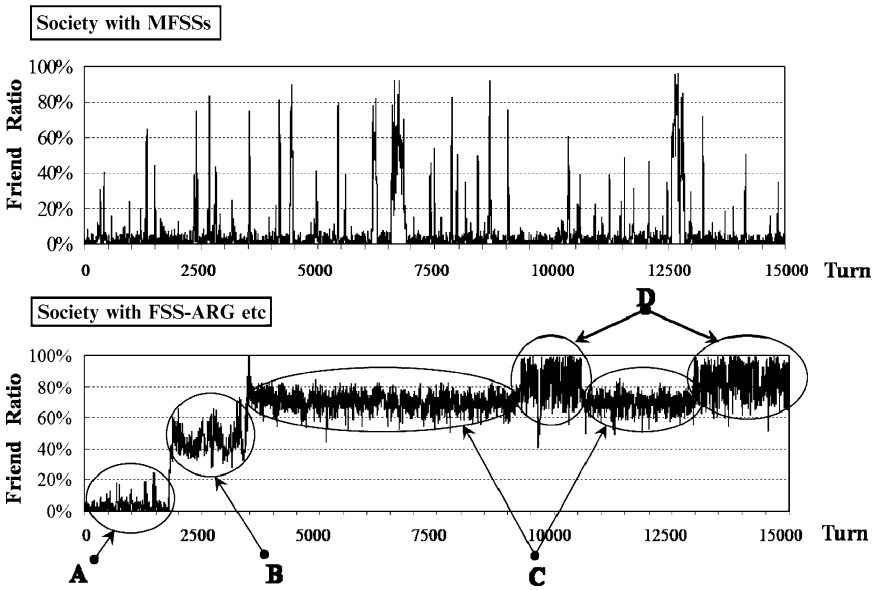


Fig. 5 Turn-series of friend ratio (*Upper*: MFSSs, *Lower*: MFSSs and FSS-ARG). * Number of Agents: $N = 40$ agents (Blue agents: 30, Red agents: 10), Matching Number of Agent in One Battle Game: $M = 20$ agents, Reflection Ratio: $R = 10\%$, Perception Error Rate: $\mu_p = 0\%$, Strategy's Mutation Rate: $\mu_s = 1\%$

both MFSSs and FSS-ARG. The conditions of the two simulations are exactly the same, except for the strategies to be examined. The number of agents is 40 ($N = 40$) which are divided into 30 blue and 10 red agents, and the perception error rate is 0% ($\mu_p = 0\%$). The matching number is 20 ($M = 20$). The reflection ratio is 10% ($R = 10\%$). The strategy's mutation rate is 1% ($\mu_s = 1\%$).

Typical results are seen in Fig. 5, which shows the turn-series of friend ratio.

The friend ratio is the average of all agents' friend ratios. The friend ratio of 1.0 corresponds to a perfectly peaceful society. The upper diagram shows the result of the MFSSs simulation and the lower shows the result of the MFSSs and FSS-ARG simulation. From the upper diagram, we can see that only MFSSs result in an instable society. And we can also find out an equilibrium corresponding to the Hobbesian state in almost all turns. In the lower diagram, in contrast, we can find out four types of social states as noted by "A", "B", "C", "D" in the diagram ("A" corresponds to the Hobbesian state). So, we can conclude that FSS-ARG have an effect on the emergence of peace to some extent.

Next, in order to examine what is going on in the equilibriums, we investigated the change in friendships among agents. For the purpose, we observed the network structure as the matrix form shown in Fig. 6. One line of a matrix is an agent's social perception, meaning who is a friend and who is an enemy. When agent i regards agent j as a friend, the element (i, j) is expressed as a white cell. A square on a diagonal is important, because it indicates the emergence of a mutually friendly community where each member sees all other members as friends.

chen.shuheng@gmail.com

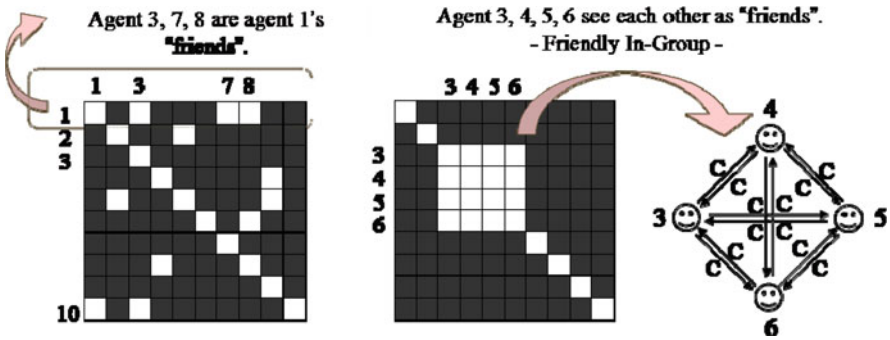


Fig. 6 Observation of relationships among agents

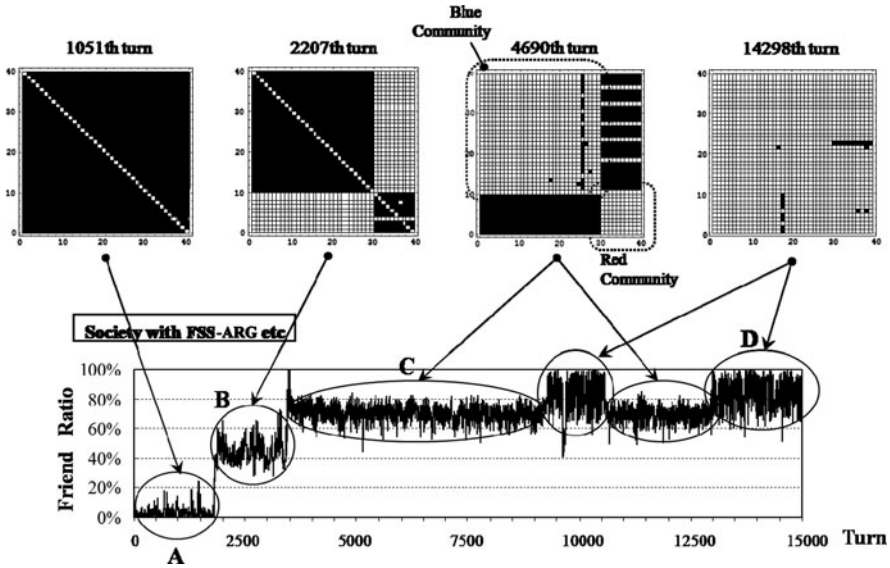


Fig. 7 Communities, anti-communities, pan-community

As a typical result, we have Fig. 7, which shows the followings.

(RLT1-1) The equilibrium "A" corresponds to one black square, which means a "war of all against all" (Hobbesian state).

(RLT1-2) The equilibrium "B" corresponds to two black squares on the diagonal. It means that all members with the same attribute are mutually hostile, while two different communities are peaceful for each other. We call the state "Anti-Communities".

(RLT1-3) The equilibrium "C" corresponds to two white squares on the diagonal. It means that all members with the same attribute are mutually friends, while two different communities are hostile to each other. So we call the

chen.shuheng@gmail.com

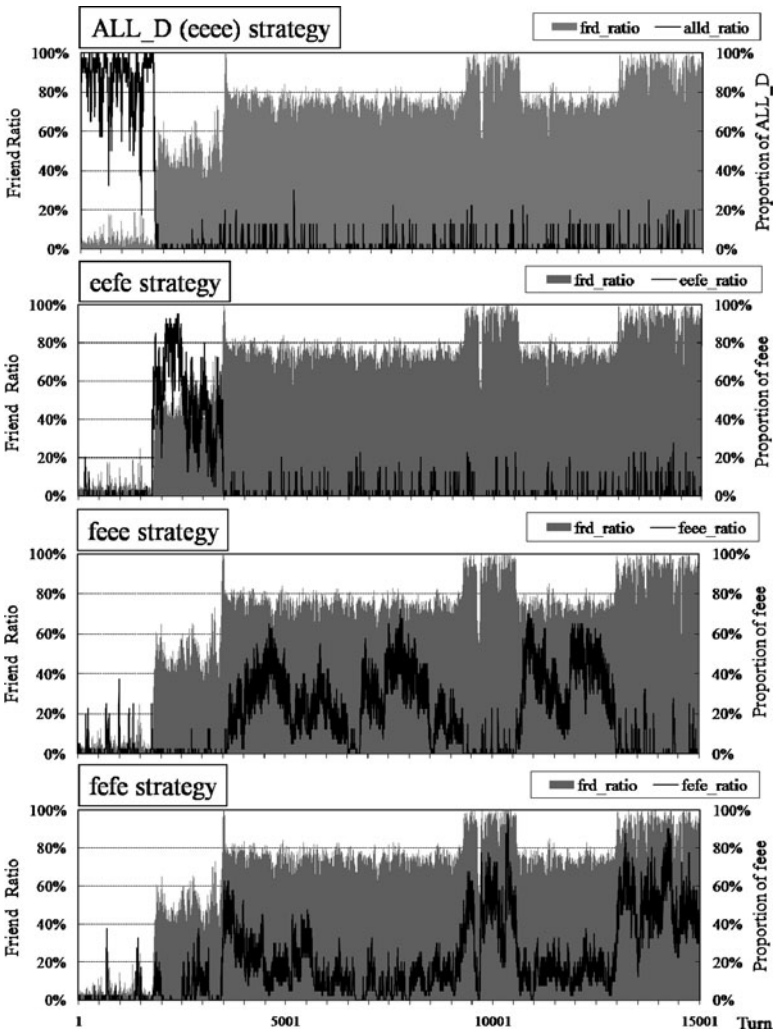


Fig. 8 Distribution ratio of prevailing strategies (*Upper*: eefe strategy, *Middle*: feee strategy, *Lower*: fefe strategy)

state "Communities". Almost all blue (red) agents form what we call a blue (red) community, where each blue (red) agent sees all other blue (red) agents as friends.

(RLT1-4) The equilibrium "D" corresponds to one white square. It means that all persons are mutually friends, regardless of their attributes. There aren't communities and anti-communities, and a whole society becomes one community. So we call the state "Pan-Community".

chen.shuheng@gmail.com

Finally, in order to understand what strategies cause the states, we observed key strategies in each state. Diagrams in Fig. 8 show the distribution ratio of prevailing strategies. The ratios were plotted together with turn-series of friend ratio. From these diagrams, we have the following findings.

(RLT2-1) ALL_D agents hold the Hobbesian state. We call it "battling society".
(RLT2-2) The feee agents hold communities. We call them "feee communities".
(RLT2-3) The eefe agents hold anti-communities. We call them "eefe anti-communities".
(RLT2-4) The fefe agents hold a pan-community. We call it "fefe pan-community".

5 Discussions

First, the characteristics of a feee community are as follows.

(FEEE1) A feee community shows the in-group reciprocity. A member who attacked his own group is always regarded as a common enemy.
(FEEE2) A feee community shows the out-group hostility. A member unconditionally attacks anybody of a different group.

Therefore, a feee community can be interpreted as the political system showing the in-group reciprocity and out-group hostility by using the group's common attribute as a symbol of community. And it seems to be interesting in the sense that the attribute comes to have the practical meaning, while it was defined as just a formal symbol.

Previous studies pointed out that group identification increases cooperation under NPD situation [2,6,9]. Especially, the minimal manipulation of group identification with abstract and vague similarities can increase cooperation rather than discussions among members [1]. These indicate that the group identification using a common symbol can foster cooperation. Although the previous studies' interests are different from ours (the former is NPD problem and the latter is 2PD problem), their findings show a degree of consistency to ours.

In addition, our results also show a similarity to Max Weber's "in-group/out-group morality". As well known, Max Weber derived it from considerations on religious communities, while feee communities don't result from religion. So it is difficult to regard Weber's finding as just a good example of a feee community. But it seems interesting that both Weber's community and a feee community have the common characteristic of unconditional hostility to aliens, which reminds us of a kind of nationalism.

Next, a fefe pan-community shows indiscriminative reciprocity, which means that a fefe agent never pay attention to others' attributes and sees others hostile to him as enemies and others peaceful for him as friends. In other words, an attribute loses a practical meaning in a fefe pan-community. In this sense, a fefe pan-community reminds us of a kind of liberalism.

chen.shuheng@gmail.com

In the end, an eefe anti-community shows the in-group hostility and out-group reciprocity, which are opposite to the feee community's case. And it seems difficult to point out good examples corresponding to an anti-community in the real world. So, an eefe anti-community can be interpreted as just a theoretical state, like an imaginary solution of a high-powered equation.

References

1. Brewer MB, Kramer RM (1986) Choice behavior in social dilemmas: effects of social identity, group size, and decision framing. J Pers Soc Psychol 50:543–549
2. Brewer MB, Schneider SK (1990) Social identity and social dilemmas: a double-edged sword. In: Abrams D, Hogg M (eds) Social identity theory: constructive and critical advances. Harvester/Wheatsheaf, New York
3. Coleman JS (1990) Foundations of social theory. Cambridge University Press, Cambridge
4. Habermas J (1981) Theorie Des Kommunikativen Handelns. Suhrkamp, Frankfurt/Main
5. Hobbes T (1651) Leviathan, printed for Andrew Crooke
6. Kerr NL (1995) Norms in social dilemmas. In: Schroeder D (ed) Social dilemmas: social psychological perspectives, pp 31–47. Pergamon, New York
7. Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. Proc R Soc Lond Ser B Biol Sci 268:743–753
8. Luhmann N (1968) Zweckbegriff und Systemrationalität über die Funktion von Zwecken in sozialen Systemen. Tubingen
9. Messick DM, Brewer MB (1983) Solving social dilemmas: a review. In: Wheeler L, Shaver P (eds) Annual review of personality and social psychology, vol 3, pp 11–44
10. Nakai Y, Muto M (2005) Evolutionary simulation of peace with altruistic strategy for selected friends. Socio-Information Stud 9(2):59–71
11. Nakai Y, Muto M (2008) Emergence and collapse of peace with friend selection strategies. J Artif Soc Soc Simul 11(3)
12. Novak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. Nature 393:573–577
13. Novak MA, Sigmund K (1998) The dynamics of indirect reciprocity. J Theor Biol 194:561–574
14. Parsons T (1937) The structure of social action. McGraw Hill, New York
15. Takagi E (1996) The generalized exchange perspective on the evolution of altruism. In: Liebrand W, Messick D (eds) Frontiers in social dilemmas research, pp 311–336. Springer, Berlin
16. Takahashi N, Mashima R (2003) The emergence of indirect reciprocity: is the standing strategy the answer?. Center for the Study of Cultural and Ecological Foundations of the Mind: Working Paper Series No.29. Hokkaido University, Japan

chen.shuheng@gmail.com

Bayesian Analysis Method of Time Series Data in Greenhouse Gas Emissions Trading Market

Tomohiro Nakada, Keiki Takadama, and Shigeyoshi Watanabe

Abstract This paper proposes the Bayesian analysis method (BAM) to classify the time series data which derives the complicated phenomena in the international greenhouse gas emissions trading. Our investigation compared the results using the method of Discrete Fourier transform (DFT) and BAM. Such comparisons have revealed the following implications: (1) BAM is superior to DFT in terms of classifying time series data by the different distances; and (2) the different distances in BAM show the importance of 1% influence of emission reduction targets.

Keywords Bayesian analysis method · Time series data · Emissions trading market · Agent-based simulation

1 Introduction

Working toward the reduction of their carbon emissions, many nations have set their greenhouse gas emissions reduction targets based on the Kyoto Protocol [1]. The Kyoto Protocol is an international agreement linked to the UNFCCC to reduce the greenhouse gas from 2008 to 2012 in the world. At present, many nations discuss the international agreement of post-Kyoto Protocol. To achieve the targets, buyers and seller of many nations transact with the emission rights in the international greenhouse gas emissions trading market. However, since it is naturally difficult for nations to achieve effective emissions trading, the compliance mechanism as rule were established by the Kyoto Protocol. Such investigations are very important

T. Nakada (✉)
Matsue National College of Technology, 14-4 Nishiikuma, Matsue, Shimane 690-8518, Japan
e-mail: t-nakada@matsue-ct.ac.jp

K. Takadama and S. Watanabe
The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan
e-mail: keiki@hc.uec.ac.jp; watanabe@ice.uec.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_11, © Springer 2011

chen.shuheng@gmail.com

because the emissions trading affects the market price of the trading, enabling participating nations decide a bid under the various uncertainties such as domestic and foreign information in greenhouse gas emissions trading. To investigate such effect of the compliance mechanism, an agent-based simulation has the great potential to analyze dynamic phenomena caused by nations participating [2]. (Hereafter, we employ the term "ABS" instead of the term "agent-based simulation"). In particular, ABS of participating nation, an analysis of the market price by the various conditions [3, 4], e.g., the time series data of the market price fluctuation show the upward and downward trend, and continue and stop the market price [5]. The analysis method uses traditional regression methods on the time series data to find the knowledge of phenomena and forecast the future price. Since a future phenomenon is uncertain, it is important to compare the time series data before investigating the relationship between a dependent variable and independent variable.

However, the time series data is difficult to completely explain a complicated phenomenon of the price fluctuation in the emissions trading. To overcome this problem, the recent research classified the time series wave data by paying attention to its' frequency, and used Discrete Fourier transform to analyze time series data calculated in ABS [6]. This analytical method using Discrete Fourier transform assumes the stationary process, i.e., some pattern of the time series data. (Hereafter, we employ the term "DFT" instead of the term "method using Discrete Fourier transform").

However, since the time series data include various properties, we cannot assume the stationary process, i.e., we should assume the non-stationary process of the time series data with uncertain pattern of the time series data. To tackle this issue directly, this paper proposes the Bayesian analysis method which classifies the multiple time series data in the greenhouse gas emissions trading as the non-stationary process, and aims at investigating the effectiveness of BAM. (Hereafter, we employ the term "BAM" instead of the term "Bayesian analysis method").

This paper is organized as follows. Section 2 explains the analytical methods of DFT and BAM, and Sect. 3 explains the ABS model of participant nations in the international greenhouse gas emissions trading market. Experiments are reported in Sect. 4. Section 5 discusses the analytical method using the simulation results. Our conclusions are given in Sect. 6.

2 Analytical Methods of Time Series Data

2.1 Overview of DFT

DFT [6] is an analytical method depending on the viewpoint of sharp fluctuation from the time series data. The method assumes the time series as pink noises or 1/f noises. This method converts the time series data from the time domain to the frequency domain by using FFT (Fast Fourier transform). Only the first few coefficients are retained to extract the low frequency as the features of the time series data. The brief sequence of DFT is summarized as follows.

chen.shuheng@gmail.com

1. DFT extracts the multiple time series data from the market price of simulation results within Hamming windows.
2. DFT converts the multiple time series into Fourier coefficients.
3. The low-pass filter is executed to get sharp fluctuation of the time series data from Fourier coefficients.
4. DFT extracts only the first few Fourier coefficients after low-pass filter, and gets distribution by the statistical calculation of the mean and the variance of them.
5. DFT classifies the multiple distribution using the procedure described in Sect. 2.3.

What should be noticed here is that the length of time series data in DFT is fixed to use Fast Fourier transform.

2.2 BAM

BAM is an analytical method depending on Bayes' theorem [7]. Bayes' theorem combines four components of the next elements, i.e., prior distribution, new data, analytical model, posterior distribution [8]. We assume the nonlinear phenomena since the market price is complicated phenomenon. BAM update posterior distribution as estimates of the evidence of the hypotheses of non-stationary process by combining the prior distribution and the data under each parameter in ABS. BAM extracts the distribution from the time series data of market price. The brief sequence of BAM is summarized as follows.

1. BAM employs the model of Gaussian distribution corresponding to the hypothesis of non-stationary process.
2. BAM sets the uninformative distribution of the market price as the prior distribution. We employ very wide normal distribution of the prior distribution (i.e., mean of zero, standard deviation of 1,000).
3. BAM gets the multiple time series data from the market price of simulation results as new data.
4. BAM calculates the posterior distribution of the market price fluctuation on Bayes' theorem.
5. BAM classifies the multiple distribution using the procedure described in Sect. 2.3.

To understand the above sequence, (1) indicates the posterior distribution $p(\theta|y)$ as follows. BAM in this paper assumes analytical model of Gaussian distribution.

$$p(\theta|y) \propto \exp\left(-\frac{1}{2}\left[\frac{1}{\tau_0^2}(\theta-\mu_0)^2+\frac{1}{\sigma^2}\sum_{i=1}^{k}(y_i-\theta)^2\right]\right) \tag{1}$$

In the above equation, $y_i = (y_1, \ldots, y_k)$ indicates the number of k from time series data of ABS. μ_0 and τ_0^2 indicates the mean and the variance of prior distribution.

chen.shuheng@gmail.com

θ and σ^2 indicates the mean and the variance of time series data of ABS such as new data. This analytical method infers posterior distribution $p(\theta|y)$ depending on Bayes' theorem from the prior distribution and new data.

2.3 Classification Method Using Distribution

The distance measure and the visualization technique were proposed as a part of the quantitative method using DFT [6]. This method visualizes the similarities among multiple distributions of the market price by calculating the distance among the distribution. This method is a common procedure after DFT and BAM. The brief sequence of classification using distribution is summarized as follows.

1. The method measures the distance between two distribution of the market price using the Mahalanobis generalized distance.
2. The method sets the distance matrix using distance between two distributions.
3. The method visualizes a geometric space of a low dimensionality using Multidimensional scaling.

To understand the above sequence, (2)–(5) indicate Mahalanobis generalized distance [9], the within-class covariance matrix, the distance matrix, and the two-dimensional space, respectively.

Mahalanobis generalized distance is shown to be a measure of distance between two distributions. It is the correlations of two distributions, and is scale-invariant. Mahalanobis generalized distance defined as follows.

$$D_M^2(m_1, m_2) = (m_1 - m_2)^t \Sigma_w^{-1} (m_1 - m_2) \tag{2}$$

In the above equation, m_1 and m_2 are mean vectors of two distributions and Σ_w refers to the within-class covariance matrix defined as follows.

$$\begin{aligned} \Sigma_w &= \sum_{i=1,2} P(\omega_i)\Sigma_i \\ &= \sum_{i=1,2} \left(P(\omega_i) \frac{1}{n_i} \sum_{x \in \omega_i} (x - m_i)(x - m_i)^t \right) \end{aligned} \tag{3}$$

In the above equation, $P(\omega_i)$, n_i, and x are a priori probability, the number of patterns of in class ω_i, and feature vector in class ω_i, respectively. If the number of samples (simulation results) in classes ω_1 and ω_2 are identical, then the relation $P(\omega_1) = P(\omega_2) = \frac{1}{2}$ holds.

This comparison method measures the distance between every two distributions by constructing a distance matrix shown in (4). Here, d_{ij} represents the distance between distributions i and j of the market price; further, $d_{ij} = d_{ji}$ and $d_{ii} = 0$.

chen.shuheng@gmail.com

$$M_d = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & n \end{array} \\ \begin{pmatrix} 0 & d_{12} & \cdots & d_{1n} \\ d_{21} & 0 & \cdots & d_{2n} \\ \vdots & & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & 0 \end{pmatrix} \end{array} \tag{4}$$

To visualize in a geometric space of a low dimensionality from multiple distance matrix, this method converts two-dimensional space of (5) from (4) using Multi-dimensional scaling [10]. Using this visualization technique, the position on the coordinate shows the distance between different distributions as follows.

$$\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} ms_{1x} & ms_{1y} \\ ms_{2x} & ms_{2y} \\ \vdots & \vdots \\ ms_{nx} & ms_{ny} \end{pmatrix} \end{array} \tag{5}$$

In the above equation, ms_{ix} and ms_{iy} represent the X axis and the Y axis of the distributions i of the market price on the coordinate, respectively.

3 Model

3.1 Model of Participant Nations in Emissions Trading

The agents in our model [5] was designed as the participant nations that agree with Kyoto Protocol. Our model is composed of a lot of agents that buy and sell their emission rights to achieve their emission reduction targets, which determines the market price as shown in of Fig. 1. Using this ABS, we can analyze the time series data of this market price from the viewpoint of the market price fluctuations. Concretely, the agents select their own actions from the possible alternatives (i.e., the buying, selling, and domestic emission reduction) every month (*time*), and the result derived by the selected actions is evaluated every commitment period defined by 5 years T as the same as the real international emission trading.

In our ABS, the agent consists of the two components (i.e., the *rule set* composed of a lot of state-action pairs and the *evaluation function* that the agent (participant nation) calculates the reward from his action) and two valuables (i.e., the *emission right* and the *amount of emission* that the agent (participant nation) generates in his country) as shown in Fig. 2. Since the agents aim at achieving their targets through the emission trading or reducing their own domestic emission, we employ the *reinforcement learning* [11] that enables the agents to maximize them reward defined by the successful degree of the emission trading. Concretely, Q-learning [12] is employed as the one of the major reinforcement learning mechanisms. The brief sequence of Fig. 2 is summarized as follows.

chen.shuheng@gmail.com

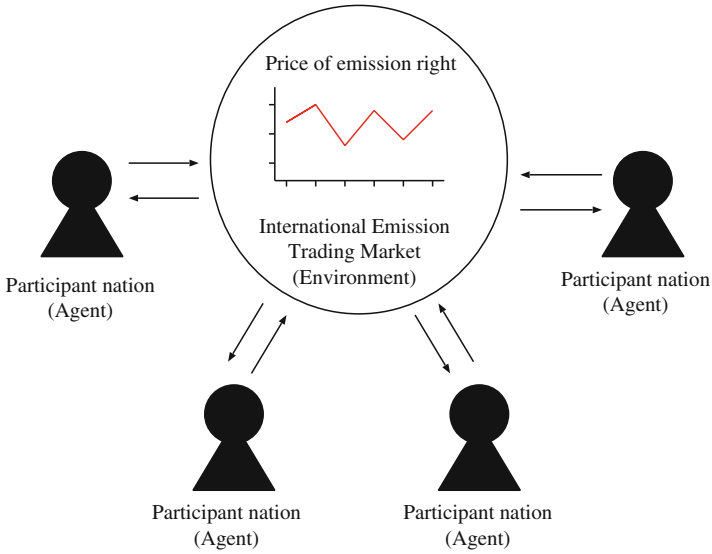


Fig. 1 Participant nation and emissions trading

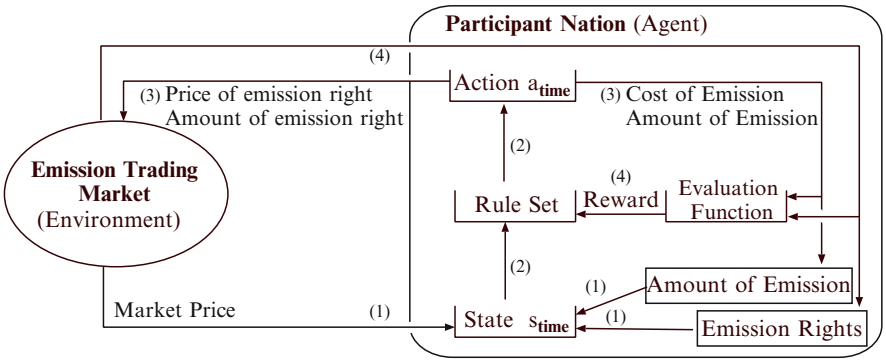


Fig. 2 Decision-making of participant nation (agent): that shows a brief decision-making sequence (1)–(4). (1) The agent recognizes a state of the market price, the amount of emission and the emission rights. (2) The agent selects an action based on the rule set of the state. (3) The action shows the bid in market trading and reduces the domestic emissions. (4) The rule set changes priority according to the reward

1. *State acquisition* The agent recognizes the market price, its amount of emission, and its emission right.
2. *Action selection by rule set* The agent determines one action from the rule set by selecting either of buying/selling the emission right or reducing the domestic emission to achieve the reduction target.

chen.shuheng@gmail.com

3. *Buying/selling in the emission trading or reducing the domestic emission* When deciding to buy or sell the emission right, both buying and selling prices of the agents are determined by the auction mechanism. When deciding to reduce the domestic emission, on the other hand, the amount of emission is reduced in the country of the agents.
4. *Changing the weight of rule set* When buying or selling the emission right, the agents evaluate their amount of the emission rights changed by the buying and selling transaction, and updates the weight of the selected rule according to the evaluation function. When reducing the domestic emission, on the other hand, the agents evaluate their amount of emission and update the weight of the selected rule according to the evaluation function. We employ the Q-learning mechanism to update the weight of them as follows.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_{a \in A(s_{t+1})} Q(s_{t+1}, a) - Q(s_t, a_t) \right] \tag{6}$$

In the above equation, the current weight and next weight indicate $Q(s_t, a_t)$ and $Q(s_{t+1}, a)$. The learning rate α $(0 \leq \alpha \leq 1)$ indicates the ratio of the learning speed, while the discount rate γ $(0 \leq \gamma \leq 1)$ determines how the reward should be considered in the t-th month by multiplying γ^{t-1} to discount it.

3.2 Compliance Mechanism

The compliance mechanism of Kyoto Protocol is employed in the real emission trading market to promote the participant nations to reduce their greenhouse gas emission. Since Kyoto Protocol establishes legally the binding commitment for reducing the greenhouse gases, the participant nations that agree with Kyoto Protocol have to set their own emission reduction targets to achieve them. To evaluate the achievement of the emission reduction targets of the participant nations, the compliance mechanism checks the participant nations by comparing its emission right with its amount of emission at the commitment period of 5 years.

There are two kinds of the compliance mechanism of Kyoto Protocol [13]. Firstly, the compliance mechanism reduces the amount of the assigned emission right of the participant nations that cannot achieve their targets in the current commitment period. Secondly, it prohibits the participant nations that cannot achieve their targets to sell the emission right to other participant nations in the next commitment period. Since the target can be achieved not only by buying the emission right from other participant nations but also by directly reducing the domestic gas emission, the participant nation has to decide either of actions (i.e., the emission right purchase or the domestic gas emission reduction). Concretely, the amount of emission and the emission right of every 5 years are calculated as follows, where T and $time$ indicate 5 years (corresponding to 60 months) and 1 month, respectively.

chen.shuheng@gmail.com

$$TofE_T = \sum_{time=1}^{T} Emission_{time} \tag{7}$$

$$TofER_T = \sum_{time=1}^{T} Right_{time} \tag{8}$$

In the above equation, $TofE_T$ and $TofER_T$ indicate the total summation of the month carbon emission $Emission_{time}$, and the total summation of the month emission right $Right_{time}$ during the commitment period T (i.e., 5 years), respectively. When the amount of emission of the participant nations does not exceed the emission right in the current commitment period, the ordinary emission trading (i.e., the buying, selling the emission right and reducing their domestic carbon emission) can be carried out in the next commitment period.

The greenhouse gas emission of the participant nation in the t-th month $Emission_{time}$ and the emission right of the participant nation in the t-th month $Right_{time}$ are respectively defined as the following equations (9), (10) and (11).

$$Emission_{time} = E_{time} - DR_{time} \tag{9}$$

$$Right_{time} = AAU_{time} \pm ET_{time} \tag{10}$$

$$AAU_T = Eyear_{1990} \times 5 \times (1 + Target)^{1+T} \tag{11}$$

In these equations, E_{time}, and DR_{time} indicate the carbon emission generated in the t-th month, and the amount of the carbon domestic reduction by the participant nation in the t-th month, respectively. AAU_{time} and ET_{time} indicate the t-th month of the assigned amount unit of the emission right (i.e., $AAU_T \div 60$ month), and the buying and selling results in the t-th month emission trading, respectively. Finally, AAU_T, $Eyear_{1990}$, and $Target$ indicate the assigned amount unit of the emission right of 5 years, the amount of emission in the base year (1990), and the carbon emission reduction target of the participant nations, respectively.

4 Experiments: Comparison of DFT and BAM

4.1 Setting

This paper investigates the times series data of the market price from the viewpoint of an influence of both the compliance mechanisms and emission reduction targets from -1 to -10%. To classify the times series data of market price with DFT and BAM, all agents in the experiments have the same emission reduction targets of which are investigated under the compliance and non-compliance mechanisms using the parameters as shown in Table 1.

chen.shuheng@gmail.com

Table 1 Parameters of participant nation, compliance mechanism, Q-Learning

Parameter	Value
Number of participant nations	39
Emissions of base year $Eyear_{1990}$	1,000
Emissions of 1 month E_{time}	65
Rate of penalty reduction $Pdeg$	1.3
Learning rate (step-size parameter) α	0.1
Discount rate γ	0.9
ε (for ε-greedy action selection)	0.2

In Table 1, the number of participant nations is set as 39 according to Kyoto Protocol, i.e., the participant nations of the post-2012 including U.S.A. $Eyear_{1990}$ indicates the amount of emission in 1990s data and E_{time} indicates the amount of emission generated by one nation. Finally, the rate of the penalty reduction $Pdeg$ is set according to Kyoto Protocol. Note that $Eyear_{1990}$ and E_{time} are tentative value, but such variable setting is enough to investigate the tendency of the emission trading market. For the parameter setting of reinforcement learning, the learning rate α, discount rate γ and ε for the ε-greedy action selection are set as shown in Table 1 which were checked in advance to converge Q-tables of the agents.

The simulation results are averaged from ten runs of simulations. In one run, the simulations are conducted until 10,000 iterations (which corresponds to 700 months in one iteration) to converge Q-tables of the agents.

4.2 Experiment Results

Figures 3 and 4 show the classification of the multidimensional scaling using DFT and BAM and the time series data. The top graph of Fig. 3 and the right graph of Fig. 4 show the distance of ratio to the standard distance between "n−1" and "n−2" within DFT and BAM. The bottom graph of Fig. 3 and the left graph of Fig. 4 show the time series data of the price of the emission right (market price). In Figs. 3 and 4, " n " denotes the non-compliance mechanism, while " c " denotes the compliance mechanism. The "number" denotes the reduction targets. For example, " c-10 " denotes simulation result of the reduction targets −10% with compliance mechanism.

From the top graph of Fig. 3, the experiment results of DFT and BAM obtain the following implications through the scatter diagram. Each graph show 20 points on the coordinate of the two-dimensional space. These graph show the similar distance and the dissimilar distance between distributions from the different wave of the multiple time series data by the 20 parameters.

We compare the time series data from the viewpoint of DFT and BAM in Fig. 3. The results of c-8 and c-9 show different distance of the price fluctuation in DFT, BAM and time series data. DFT shows the short distance, while BAM shows the long distance. In the time series data up and down fluctuation occurs at 450 times in c-9, while such fluctuation does not occur at 450 times in c-8.

chen.shuheng@gmail.com

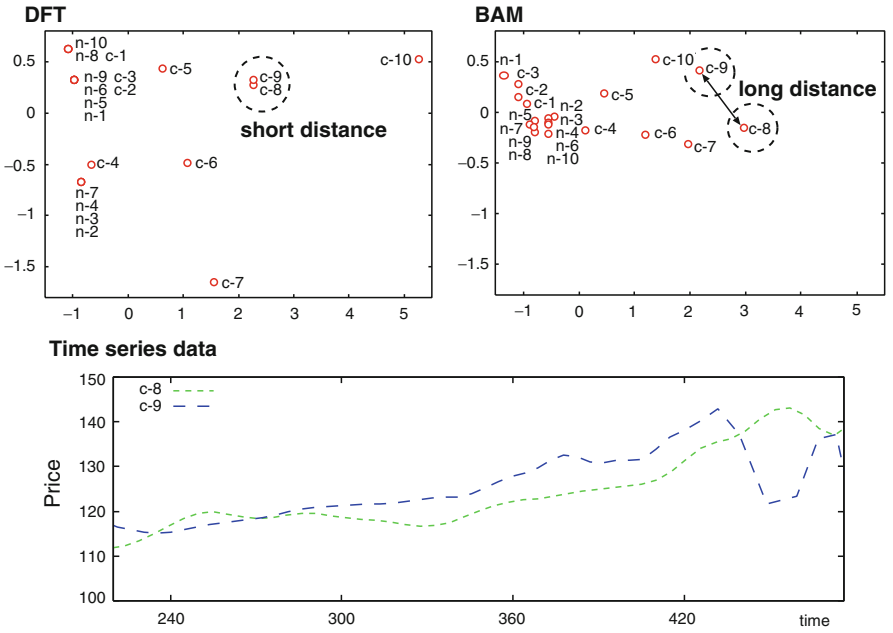


Fig. 3 Comparison of DFT, BAM and Time series data: The *top graph* show the distance of ratio to the standard distance between "n−1" and "n−2" within DFT and BAM. The *bottom graph* shows the time series data of the price of the emission right (market price)

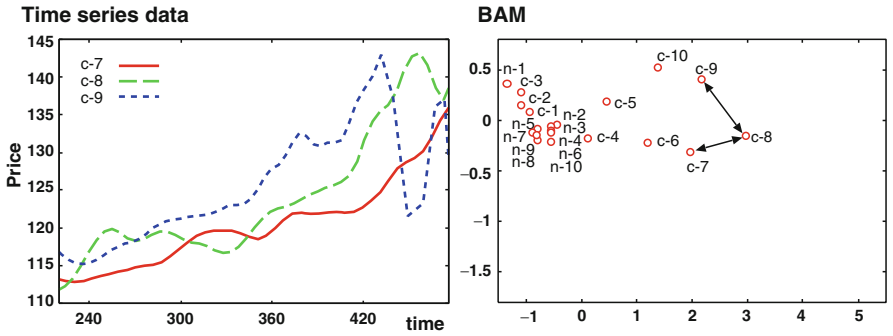


Fig. 4 Comparison of Time series data and BAM: The *right graph* show the distance of ratio to the standard distance between "n−1" and "n−2" within BAM. The *left graph* shows the time series data of the price of the emission right (market price)

In order to investigate the effect of BAM, we compare the multiple time series data from the viewpoint of BAM in Fig. 4. Figure 4 shows BAM and time series data which adds the result of c-7 in Fig. 3. Since c-7 and c-8 show similar upward

chen.shuheng@gmail.com

trend of market price, the distance between c-7 and c-8 is shorter than the distance between c-8 and c-9. While c-9 shows big fluctuation at 450 times.

5 Discussion

In Sect. 4.2, Fig. 3 shows the classification of time series data of market price using DFT and BAM. Although, we regard the price fluctuation as a complicated phenomena in the emissions trading, the analytical methods of DFT and BAM show different results. We discuss this issue in Figs. 3 and 4, respectively.

When we pay attention to 450 times in Fig. 3, the time series data shows the up and down fluctuation in c-9 while such fluctuation does not occur in c-8. DFT does not show the long distance from the upper left figure, because this fluctuation is calculated as an average of the frequency of the market price. In comparison with DFT, BAM shows the long distance, because BAM succeeds to find the sharp fluctuation of the market price by predicting the future distribution of the market price as the posterior distribution. In addition, BAM shows 20 point on the coordinate from 20 parameters, while DFT shows ten point on the coordinate from 20 parameters. This is because a point shows multiple parameters on the same coordinate by DFT. For example, one point show the same coordinate of "n−2", "n−3", "n−4", and "n−7" by DFT. However, a point shows parameters of one on the coordinate by BAM. From these results, BAM is superior to DFT in terms of classifying time series data by the different distances.

When we pay attention to the distance between parameters of simulation using BAM in Fig. 4, 1% difference between −8 and −9% of the emission reduction targets gives the larger influence to the fluctuation of the market price. In the same way, 1% difference between −7 and −8% of the emission reduction targets gives the smaller influence to the fluctuation of the market price. From these results, BAM shows 1% difference between parameters of emission reduction targets at each coordinates. Therefore, we can measure the fluctuation of 1% difference as the dissimilar distance between distributions from the different wave of the multiple time series data.

The following issues should be pursued in the near future: (1) an investigation to classify the time series data of different ABS; and (2) an investigation of guideline of analysis feature and function according to BAM. First, we can classify the distance between distributions from the different wave of the multiple time series data of same ABS. On the other hand, we donot classify them of different ABS. However, since BAM addresses the different wave of the multiple time series data, BAM is possible to measure the distance between distributions from the different wave of the multiple time series data of different ABS. Second, although DFT and BAM classify the different distance from the multiple time series data, we donot explain the analysis feature and function according to the different distance by DFT and BAM.

chen.shuheng@gmail.com

6 Conclusion

This paper proposed Bayesian analysis method (BAM) to classify the time series data which derives the complicated phenomena in the international greenhouse gas emission trading. Our investigation compared the classification of time series data at the distances using DFT and BAM. These revealed the following remarkable implications: (1) BAM is superior to DFT in terms of classifying time series data by the different distances; and (2) the different distances in BAM show the importance of 1% influence of the emission reduction targets and compliance mechanism. In addition, we need the other analytical method to classify the the different wave of the multiple time series data, since the times series data of market price is the complicated phenomena. ABS also needs analytical method to classify the the different wave of the multiple time series data, and to explain the price fluctuation as a complicated phenomena in the trading market.

The following issues should be pursued in the near future: (1) an investigation to classify the time series data of different ABS; and (2) an investigation of guideline of the analysis feature and function according to BAM.

Appendix

DFT and BAM used the following time series data in Fig. 5.

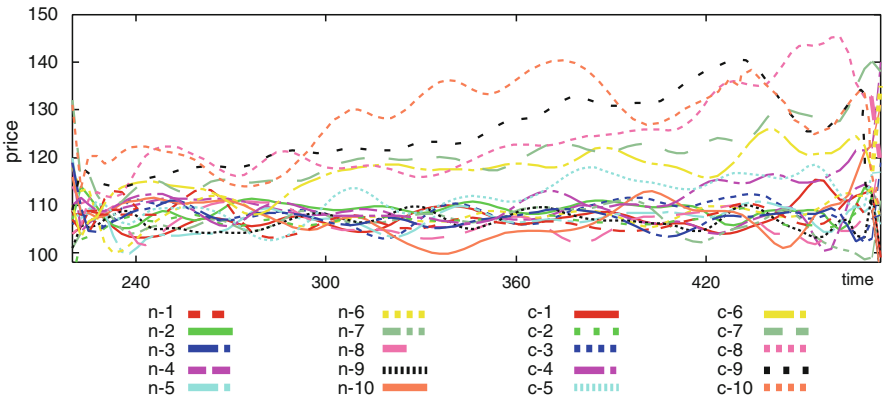


Fig. 5 Time series data of 20 parameters: It is the time series data using analysis of DFT and BAM

References

1. United Nations Framework Convention on Climate Change (1997) Report of the conference of the parties on its third session. http://unfccc.int/resource/docs/cop3/07a01.pdf
2. Arthur BW, Holland JH, LeBaron B, Palmer R, Taylor P (1996) Asset pricing under endogenous expectation in an artificial stock market. In: Santa Fe Institute Working Papers, 96-12-093

chen.shuheng@gmail.com

3. Mizuta H, Yamagata Y (2001) Agent-based simulation and greenhouse gas emissions trading. In: Winter simulation conference 2001
4. Matsumoto K (2008) Evaluation of an artificial market approach for ghg emissions trading analysis. Simulat Model Pract Theory 16(9):1312–1322
5. Nakada T, Takadama K, Watanabe S (2008) Agent-based modeling of the participant nations and the compliance mechanism in the emissions trading. In: The second world congress on social simulation
6. Arai R, Watanabe S (2009) A quantitative method for comparing multi-agent-based simulations in feature space. In: Multi-agent-based simulation 2008, LNAI, vol 5269, pp 155–166
7. Bayes T (1763) An essay towards solving a problem in the doctrine of chances. Philos Trans 53:370–418
8. McCarthy MA (2007) Bayesian methods for ecology. Cambridge University Press, Cambridge
9. Duda RO, Hari PE, Stork DG (2000) Pattern classification, 2nd edn. Wiley-Interscience, New York
10. Torgerson WS (1952) Multidimensional scaling: I. Theory and method. Psychometrika 17(4):401–419
11. Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. The MIT Press, Cambridge, MA
12. Watkins CJCH, Dayan P (1992) Q-learning. Mach Learn 8:279–292
13. United Nations Framework Convention on Climate Change (1997) Issues in the negotiating process (compliance under the Kyoto Protocol). http://unfccc.int/kyoto_protocol/compliance/items/2875.php

chen.shuheng@gmail.com

Large Scale Crowd Simulation of Terminal Station Area When Tokai Earthquake Advisory Information Is Announced Officially

Qing-Lin Cui, Manabu Ichikawa, Toshiyuki Kaneda, and Hiroshi Deguchi

Abstract On the assumption of advisory information concerning an imminent Tokai earthquake being officially announced, as a case example we developed a LSCS (Large Scale Crowd Simulator) for the Nagoya Station area, where several terminal stations are concentrated; in the model, agents played people on their way home, and such factors as the routes selected by agents and the spatial restrictions, e.g. passages, were taken into consideration. Basic on SOARS (Spot Oriented Agent Role Simulator) platform, we conducted a large-scale crowd simulation with 160,000 agents and analysis the change of space density in 1 h to compare to the estimates given by Nagoya City, we analysis the result and also refer to LSCSver1.0 for implementing much higher functions.

Keywords Spatial-spot type agent-based simulation · Tokai earthquake advisory information · Terminal station · Large area · Crowd management

1 Introduction

At present the Japanese Government has designated the possibility of a Tokai earthquake as the only example of a predictable earthquake. In 2004, the Meteorological Agency newly added "advisory information" assuming a predictable case. Advisory information will first be officially announced, followed later by a public "warning". The municipality of Nagoya expects all people, commuters and the like to return to their homes immediately. It is reasonable to make the area

Q.-L. Cui (✉) and T. Kaneda
Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 4668555, Japan
e-mail: qinglin.cui@gmail.com; cgt18509@stn.nitech.ac.jp

M. Ichikawa and H. Deguchi
Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Nagatsuta-cho, Midori-ku, Yokohama 2268503, Japan
e-mail: ichikawa@dis.titech.ac.jp; deguchi@dis.titech.ac.jp

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_12, © Springer 2011

chen.shuheng@gmail.com

around a large-scale terminal station, such as Nagoya Station, due to the massive numbers of people all returning home simultaneously; such a situation requires the application of measures to prevent crowd accidents [1].

We first do the research with a maximum scale of 6,000 people, employing cell-spatial-type agent-based simulation [2–4] and the process of such an increase and measures to prevent accidents is possible. But it is difficult to deal with 100,000 pedestrians. As a kind of network-type agent-based simulation, we conducts LSCS (Large Scale Crowd Simulator) with 160,000 agents. This time we main to analyze the congestion of complicated flows around terminal stations.

Therefore, beside the model, this research uses a spatial model for the Nagoya Station area where five railway lines converge, the research measures density at each space throughout the simulation. The research demonstrates a process in which the spatial distribution of crowd density is changing, and examines the density distribution and the tendency of changes according to different cases, e.g. flow coefficients or origin-destination data. We also examine the crowd processing performance of station and use the application of measures for crowd management.

2 A Framework of the Simulation

2.1 Characteristics of the Model Employed in the Research

To handle a large-scale crowd simulation, the research introduced two characteristics into the model. Firstly, the entire space was divided into partial spaces called spatial spots, and was not divided into cells, and a link connecting a spatial spot with an adjacent spatial spot was established; a model with this structure (spatial-spot type model) was used. This is a kind of a network-type model and has a characteristic where nodes are given spatial attributes. As a reported case study employing a model that introduced spatial attributes to nodes, the process analysis of infectious disease dissemination under typical daily activities in urban areas is well known [5].

The second characteristic is that this model includes interactions between actors; more specifically, a pedestrian selects their own walking route and this changes the density and affects other pedestrians when they select their own walking route. As a simulation that dealt with such agents, a research has been carried out that examined evacuation, which is mainly unidirectional flow [6]; however, no research has been reported that examined complicated flows such as transfers between terminal stations.

In addition, the simulation conducted in the research employed SOARS (Spot Oriented Agent Role Simulator) as a platform that has strong advantages in spot-agent representation and enables us to run large-scale agent simulation.

chen.shuheng@gmail.com

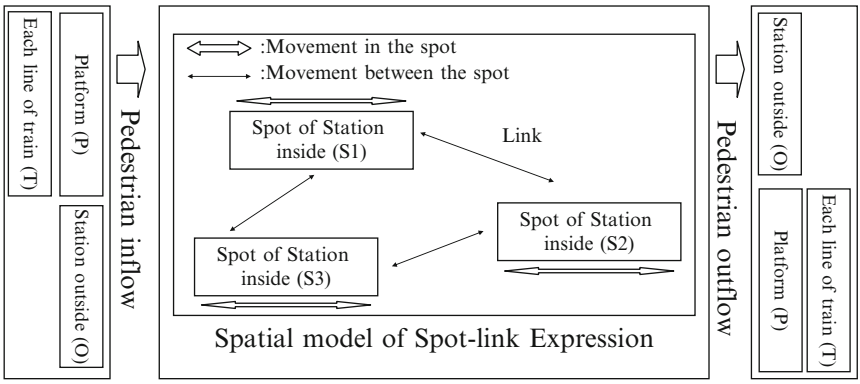


Fig. 1 Overall structure of the simulation

2.2 An Outline of the Model and Framework of the Simulation

The model used for the research consists of pedestrian agents and a spatial model. The spatial model is classified into two types of spatial spots (hereinafter, simply referred to as "spot"): internal spot S, inside the station, and external spot O, outside the station, and T, the inside of a train, and P, the platform.

Figure 1 shows the overall structure of the simulation. Firstly, before starting the simulation, pedestrian agents are given attributes to follow when they decide a walking route. In accordance with in-flow and movement conditions, and their assessment of a destination, agents walk from a departure point to their final destination by moving within and between spots.

A spatial spot records any pedestrian agent that exists within the spot and each of their steps, and based on the data it calculates the density of the spatial spot, and the distance moved per step of a pedestrian agent. Spatial spots have spatial attribute parameters and pedestrian agents have route selection rules.

3 Production Design of LSCSver.1.0

3.1 Details of the Model

This section describes details of the model from two points of view – for each of the spots and pedestrians – and then, explains the computing process by spots and the behavior rules of pedestrians.

Spatial spots. The model used for the research consists of pedestrian agents and a spatial model. The spatial model is classified into two types of spatial spots: internal spot S, inside the station, and external spot O, outside the station, and spot T, the inside of a train, and spot P, platform. Spot has length L_m and width W_m; this is approximated by a rectangular space with an area of $L_m{}^*W_m$.

chen.shuheng@gmail.com

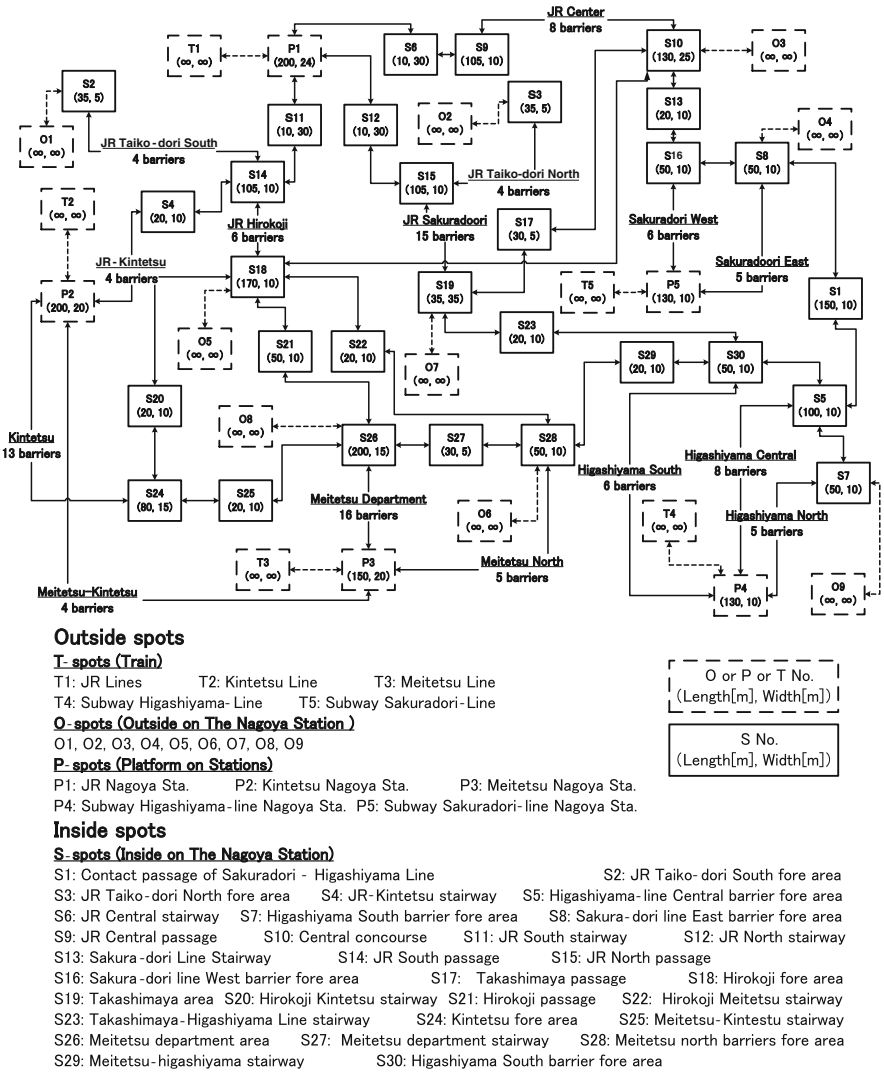


Fig. 2 Nagoya station area spatial model expressed as spatial spots

Spatial Model of Nagoya Station area. For the research, the Nagoya Station area are made by 5 T-spot, 5 P-spot, 9 O-spot and 30 S-spot which shows the spatial model representing by Fig. 2.

Pedestrian agents. Spots O and T have random variables w_1 and w_2 which belong to the agent $u \in C$. w_1 is a random variable for determining a starting point, and w_2 is a random variable for determining the destination of an agent who departs from that starting point.

A pedestrian agent selects the shortest route without the change of density from their starting point to the ending point.

chen.shuheng@gmail.com

3.2 Assessment of the Shortest Distance

Concerning the calculation of a distance, firstly it is assumed that there is a node in the center point of a crossover spot expressed as a rectangle. As a method to search for the shortest route, Dijkstra's method [7] (label-setting method) is used. The found shortest route is set as the basic walking route of an agent.

3.3 Rules Concerning Movement of Pedestrian Agents

This section describes the method to calculate assessment indicators, which affect walking and route selection, and the overall process.

Rules for movement within a spot. When a pedestrian agent enters in spot S_m, the agent reads an L_m value of spot S_m. Movement distance for each step ΔL is given by (1).One of the most influential to affect the speed is considered to be density. A straight line model measured by an actual measurement experiment for two-way flows, and an estimate equation of walking speed V is given by (2).

$$\Delta L = V \times T \tag{1}$$

$$V = a \times \rho + V_0 \tag{2}$$

$$\Delta L = (a \times \rho + V_0) \times T \tag{3}$$

Here, T: Time/step (s/step), ΔL: Movement distance/step (m/step), V: Walking speed (m/s), ρ: density (people/m^2), V_0: Standard walking speed, a: Parameter

ΔL is added to L, the distance which an agent has already moved within a spot.

When $L \geq L_m$, movement within the spot ends.

Set values for V_0 and a differ according to one-way or two-way flow. In addition, in two-way flow the set values differ according to the ratio of pedestrians moving in the opposite direction (Table 1).

In the research one-way flow and single and double pedestrian ratios for two-way flow are implemented. The ratio of number of pedestrians is rounded off, and less than one is regarded as a one-way flow, and two or more is regarded as a two-way flow.

Table 1 Walking speed used for the model [8]

Crowd flow	The number of people rate (α)	Flow of people[a]	a	V_0
One-way flow	–	All	−0.28	1.48
Two-way flow	1 time	All	−0.275	1.605
	2 times	More	−0.285	1.675
		Less	−0.39	1.958

[a] Flow of people: All, The flow of all people; More, The flow of more people; Less, The flow of less people

chen.shuheng@gmail.com

Table 2 Expression of the ticket barrier

The passage type	The effective rate of the width (γ)	Method of counting number of streets
Ticket barrier none	1	Each of the two directions
Ticket barrier	0.5	Two directions total

Rules for restricting movement between spots. A stationary state is caused by pedestrians stopping or reducing walking speed, or by a location with a width narrowing. In the research, we only consider the stationary states caused by reduction of spot width are taken into account.

When spots S_m and $S_m{}'$ are adjacent to each other, and their width W_m and $W_m{}'$ are compared, the in-flow upper limit for the side with the smaller width is applied. In this case, if $W_m \leq W_m{}'$, the in-flow upper limit Q is expressed as below:

$$Q = k \times \gamma \times W_m \times T \tag{4}$$

Here, Q: In-flow upper limit, k: Flow coefficient $(=1.5\text{people}/\text{ms})$, γ: The effective rate of the width, T: Time of 1 step $(=10\,\text{s})$

Agents more than Q stay and remain in the same spot.

The setting of the passage and the ticket barrier is explained in Table 2.

Rules for in-flow/out-flow of pedestrians. In-flow and out-flow of pedestrians is controlled by spot O outside the station and spot T on each platform and platform P.

An outside-the-station spot calculates and controls a maximum number of pedestrians that flow in or out per unit of time, based on the width of the spot for in-flow or out-flow, and a flow coefficient.

Platform spot, considered the train carry ability, each up-line and down-line controls the largest number of people per time who flow out.

A train spot on each platform allows a set number of pedestrians to flow in with 360 steps and averagely flow out.

3.4 Rules for Route Selection by Pedestrian Agents

When an agent arrives at a crossover spot, they need to determine their next route. But if the shortest route can not be chosen, people can choose the rest but the shorter line. However, the condition for being a possible route choice is any route that enables the agent to reach their destination.

For route selection, two kinds of criteria are provided: the shortest route β_1 and density assessment β_2.

When the density of the next spot located on the shortest route is less than β_1 $(\text{people}/\text{m}^2)$, the agent can now take the shortest route.

When the density of the next spot located on the shortest route is β_1 $(\text{people}/\text{m}^2)$ or more, the agent moves to an alternative route spot with a density of less than β_1 $(\text{people}/\text{m}^2)$. If it is the agent takes the shortest route. If density is β_2 $(\text{people}/\text{m}^2)$ or more, the agent remains in the present spot.

chen.shuheng@gmail.com

4 Simulation Analysis

This chapter describes the simulation analysis. It was assumed that advisory information is officially announced from 2:00 P.M. onward. The Nagoya City Local Government has given estimates of the number of people likely to converge on the station; for the simulation three cases were set: 0%, 50% and 100% of the estimated numbers. In addition, for each case four patterns of detour criteria density were provided: 0, 1, 2, or 3 $(\text{people}/\text{m}^2)$.

The simulation was conducted to confirm the agent behavior needed to analyze stationary states, and spatial bottlenecks; the time scale was set as 10 s/step with a total of 360 steps.

4.1 Rules for Route Selection by Pedestrian Agents

Parameters were set as shown in Table 3. OD data matrix and the number of agents were set as shown in Table 3 (train side), and Table 4 (outside). The number of agents were set as 158,200.

It was assumed that advisory information of a Tokay earthquake is officially announced at 2:00 P.M. on a weekday and the simulation experiment focused on the congestion state in the station space between 2:00 and 3:00 P.M. Table 5 gives classifications for each case.

This section describes the classification of cases. Assuming that all passengers traveling by train flow into the station, the number of passengers was set at 73,881, the number estimated by Nagoya City. The number of pedestrians arriving from outside the station to take a train was set at three cases – 0%, 50% and 100% of the number estimated by Nagoya City; the number of in-flow pedestrians from outside was 0 for Case 1; 45,656 for Case 2; and 91,311 for Case 3. Moreover, according to the differences of β_2, which is a density decision criteria for taking a detour, four patterns were set for each case. β_2 $(\text{people}/\text{m}^2)$ is a decision criteria for an agent to take a detour. The agent takes a detour under the following conditions: the density of the front spot – the shortest route – is β_1 $(\text{people}/\text{m}^2)$ or more, and any possible detour route has a density of β_2 $(\text{people}/\text{m}^2)$ or less. When β_2 is 0 $(\text{people}/\text{m}^2)$, no detour is necessary. The upper limit of β_1 was set as 3 $(\text{people}/\text{m}^2)$.

4.2 Results Analysis

This section illustrates the measurement data used for analysis, and the data is compared with congestion analysis using estimates given by Nagoya City. The performance of the model and analysis results are then described.

Agent behavior. Six typical agents are abstracted from simulation as an example to explain the performance of model (Table 6).

chen.shuheng@gmail.com

Table 3 Estimated OD matrix used for the simulation (train) [9]

O \ D		T-line Up train	T-line Down train	K-line Up train	C-line Up train	H-line Up train	H-line Down train	S-line Up train	S-line Down train	M-line Up train	M-line Down train	K-line Up train	Out side	Total (people)
T-line	Up train	–	0 (0%)	285 (6%)	999 (22%)	1,070 (24%)	272 (6%)	285 (6%)	0 (0%)	188 (4%)	83 (2%)	574 (13%)	772 (17%)	4,529
T-line	Down train	0 (0%)	–	630 (9%)	31 (1%)	355 (12%)	212 (7%)	71 (2%)	0 (0%)	0 (0%)	430 (15%)	681 (24%)	448 (16%)	2,858
K-line	Down train	144 (12%)	225 (19%)	–	152 (13%)	153 (13%)	0 (0%)	33 (3%)	0 (0%)	87 (7%)	227 (19%)	0 (0%)	173 (14%)	1,194
C-line	Down train	48 (%)	3,059 (39%)	1,077 (14%)	–	0 (0%)	187 (2%)	0 (0%)	0 (0%)	57 (1%)	702 (9%)	2,299 (29%)	365 (5%)	7,793
H-line	Up train	607 (9%)	920 (14%)	65 (1%)	500 (7%)	–	0 (0%)	661 (11%)	0 (0%)	1,566 (25%)	1,501 (23%)	335 (5%)	300 (7%)	6,455
H-line	Down train	1,177 (10%)	2,410 (20%)	296 (2%)	54 (0%)	0 (0%)	–	18 (0%)	13 (0%)	1,392 (12%)	3,306 (27%)	2,800 (23%)	718 (7%)	12,183
S-line	Up train	1,211 (18%)	1,343 (19%)	138 (1%)	929 (13%)	852 (12%)	114 (1%)	–	0 (0%)	348 (4%)	1,673 (24%)	444 (6%)	0 (0%)	7,052
S-line	Down train	1,038 (10%)	2,048 (20%)	378 (4%)	123 (1%)	99 (1%)	241 (2%)	0 (0%)	–	1,130 (12%)	1,919 (20%)	2,686 (25%)	490 (5%)	10,153
M-line	Up train	460 (4%)	504 (5%)	564 (5%)	687 (7%)	1,865 (18%)	560 (5%)	183 (2%)	60 (0%)	–	0 (0%)	3,766 (35%)	1,949 (18%)	8,070
M-line	Down train	0 (0%)	197 (3%)	176 (3%)	133 (3%)	687 (12%)	444 (6%)	136 (3%)	0 (0%)	0 (0%)	–	3,119 (47%)	1,562 (23%)	6,454
K-line	Up train	358 (5%)	831 (12%)	0 (0%)	911 (13%)	1,195 (17%)	73 (1%)	391 (5%)	0 (0%)	1,193 (17%)	1,685 (24%)	–	503 (7%)	7,140
Total		5,043	11,537	3,609	4,519	6,276	2,103	1,778	73	5,961	11,526	16,704	7,280	73,881

T-line: Tokaido line; K-line: Kansai line; C-line: Chuo line; H-line: Higashiyama-line; S-line: Sakuradoori-line; M-line: Meitetsu-line; K-line: Kintetsu-line

chen.shuheng@gmail.com

Table 4 Estimated OD matrix used for the simulation (outside) [9]

O \ D	T-line Up train	T-line Down train	K-line Up train	C-line Up train	H-line Up train	H-line Down train	S-line Up train	S-line Down train	M-line Up train	M-line Down train	K-line Up train	Total (people)
01	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	6,392
02	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	6,392
03	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	15,523
04	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	12,784
05	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	6,392
06	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	6,392
07	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	21,915
08	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	9,131
09	7%	11%	2%	9%	12%	2%	3%	0%	20%	23%	11%	6,392
Total	6,517	9,930	1,622	8,479	10,979	2,260	2,595	38	18,337	20,937	9,617	91,311

T-line: Tokaido line; K-line: Kansai line; C-line: Chuo line; H-line: Higashiyama-line; S-line: Sakuradoori-line; M-line: Meitetsu-line; K-line: Kintetsu-line

Table 5 Cases in the simulation

Case no.		The number of people (people) Train side	In-flow	The density standard value (people/m^2) β1	β2
Case 1	1	73,881	0	3	0
	2	73,881	0	3	1
	3	73,881	0	3	2
	4	73,881	0	3	3
Case 2	1	73,881	45,656	3	0
	2	73,881	45,656	3	1
	3	73,881	45,656	3	2
	4	73,881	45,656	3	3
Case 3	1	73,881	91,311	3	0
	2	73,881	91,311	3	1
	3	73,881	91,311	3	2
	4	73,881	91,311	3	3

4.2.1 Measurement Data

Figure 3 illustrates the results of the measured densities, and Fig. 5 illustrates changes in density distribution (Case 2–2).

Examination for crowd management. In case 2–2 the number of agents flowing into Nagoya Station was set as 73,881 (in-flow passengers from trains) and 45,656 (in-flow pedestrians from outside). Each agent basically walks along the shortest route to their train or destination; however, if the /hboxdensity of the front spot is 3 people/m^2 or more, and a detour route is available with a density of less than 1 person/m^2, they can take the detour route, and a detour route is available with a density of 1 person/m^2 or more, they stay. About 5 min after the start of the simulation, S30, S23, S19, and S17 reached their upper limit and became congested; as time

chen.shuheng@gmail.com

Table 6 Typical Agent movements of simulation analysis

Agent no.	Start	End	PSN[a] (St,[b] As[c])
1	T2	O5	T2 (120, –), S13 (90, 1.1), S21 (60, 1.3), S17 (20, 1.0), S14 (120, 1.42), O5
2	O7	T2	O7 (0, –), S15 (160, 1.31), S12 (80, 1.25), S5 (210, 1.42), S14 (120, 1.42), S17 (20, 1.0), S21 (60, 1.3), S13 (90, 1.1), T2
3	T3	T4	T3 (10, –), S29 (120, 1.25), S25 (70, 1.42), S26 (20, 1.0), S27 (60, 1.3), S28 (80, 1.3), T4
4	T1	T4	T1 (2,300, –), S3 (210, 0.9), S6 (20, 10), S9 (100, 1.0), S14 (150, 1.1), S18 (90, 1.1), S23 (200, 1.0), S29 (150, 1.0), T4
5	T1	T4	T1 (2,100, –), S3 (240, 0.8), S7 (20, 1.0), S10 (90, 1.1), S15 (200, 0.9), S20 (20, 1.0), S27 (60, 1.3), S26 (30, –), S25 (90, 1.1), S29 (130, 1.2), T4
6	T2	T3	T2 (2,470, –), S13 (150, 0.6), S29 (140, 1.1), T3

[a] Passed Spot Name
[b] Spent time(s) or Waiting time(s)
[c] Average street speed (m/s). As is calculated by $As = L_m/S$

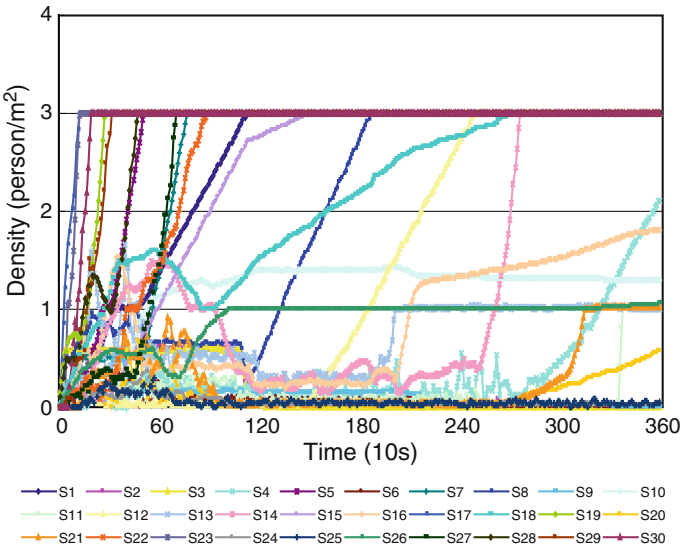


Fig. 3 Illustration of the density measurement results (Case 2–2) (color figure online)

went by, the number of congested and stationary spots gradually spread until by the end of the simulation a total of 16 spots, including the above four spots, had reached the upper limit.

When the measurement results of each case are examined, the points found by simulation analysis are summarized as follows:

1. In each case, the highest number of agents processed, and the lowest number in a stationary state were found when no detours were taken. For both β_1 and β_2, in

chen.shuheng@gmail.com

the case of 3 people/m^2, the number of agents who stayed inside increased and the number of agents processed showed the greatest decrease.

2. When the number of in-flow agents from outside increased, locations with increasing density also tended to increase.
3. When the same patterns in different cases were compared, it was found that as the number of in-flow agents from outside increased, processing capacity tended to decline.

Comparison of processing capacity. Concerning processing capacity, when there was no in-flow from outside, the number of agents processed in the simulation model was about 80% of the estimates of Nagoya City; when the time required for movement and the impact of congestion are taken into consideration, this result is reasonable. However, along with an increase of in-flow from outside, congested spaces triggered a decrease in walking speeds and consequent stationary phenomenon, and it was found that the processing capacity dropped to about 40%. Particularly, in the conditions of Cases 2–4 and 3–4 where a detour was allowed to the upper limit, it was confirmed the processing capacity decreased to 30% (Fig. 4).

The number of out-flow. The changes to the number of agents processed were examined using the graph for the number of out-flow agents (Fig. 5). Case 1 – no in-flow from outside – recorded the largest accumulated number of agents processed.

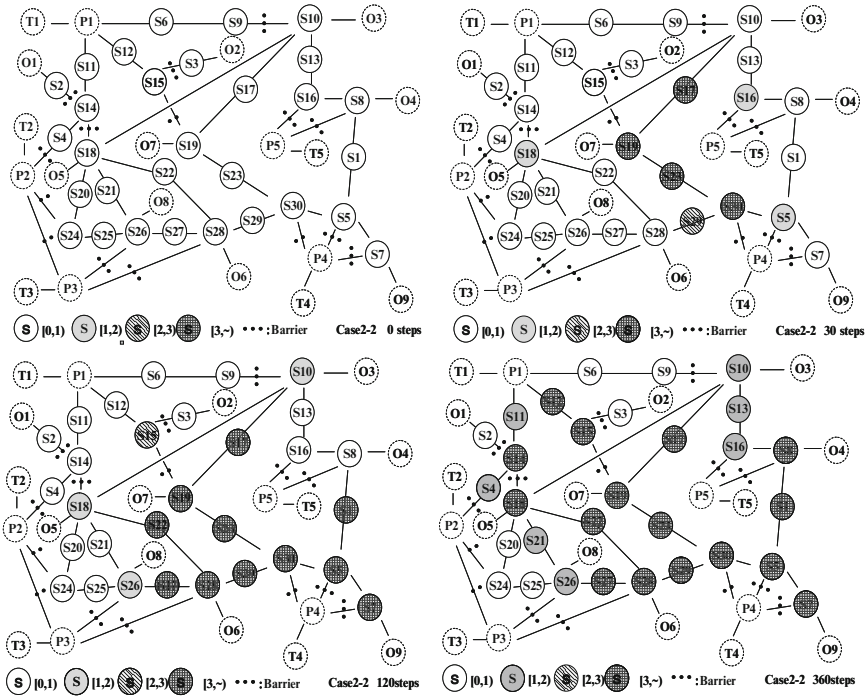


Fig. 4 Illustration of changes in density distribution (Case 2–2)

chen.shuheng@gmail.com

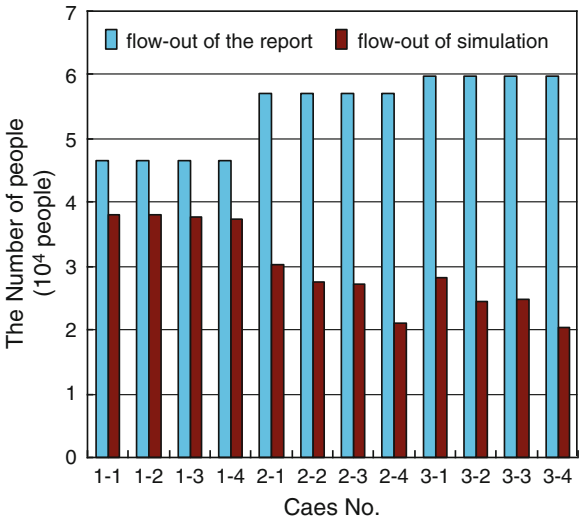


Fig. 5 Comparison with the estimates given by Nagoya City (processing capacity in the initial 1 h) (color figure online)

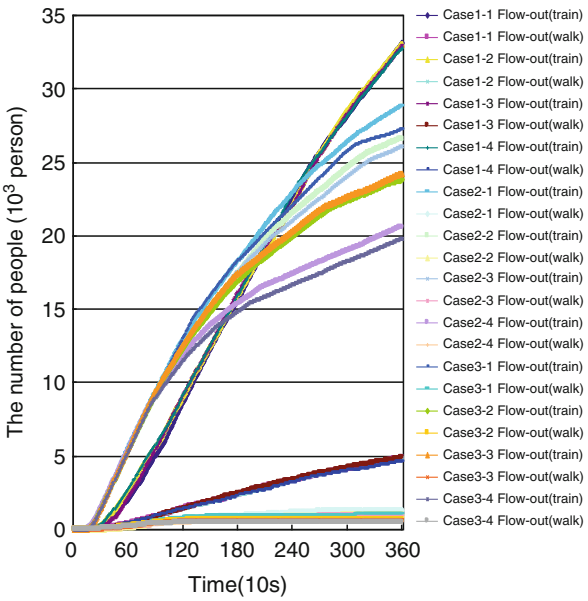


Fig. 6 Graph: number of out-flow agents (train/walking) (color figure online)

In Cases 2 and 3, up to about 120 steps, more out-flow agents than Case 1 were found, but gradually due to increasing congestion, processing capacity declined. In particular, in Cases 2–4 and 3–4, with detour behavior allowed up to the upper

chen.shuheng@gmail.com

limit, congestion in the inside area further increased, resulting in the largest decline in processing capacity. The case of pedestrian out-flow showed a similar tendency; however, compared to the number of in-flow agents their number was very low, and therefore in the second half, processing capacity remained at almost the same level (Fig. 6).

5 Conclusion

On the assumption of Tokai earthquake advisory information being officially announced, this research created LSCSver1.0 of the behavior of pedestrian agents, some of whom were eventually forced to remain in a stationary state, and developed the model to analyze an agent-based simulation at the terminal station. Through the simulation experiment the following was found: (1) due to congestion caused by agents getting off trains, increasing density spread in several areas; as in-flow agents from outside increased, the sections in front of the ticket barriers and internal passages experienced an increase of density, causing a further decline of processing capacity; (2) as a result of (1), when the simulation results for the number of agents processed during the initial 1 h are compared to the estimates released by Nagoya City, they show a figure of 30–80% of the upper limit of the estimate; and (3) even though agents who flowed in from outside were able to take a detour in response to congestion, the number of stationary agents actually increased, and taking a detour had little effect on the increasing density; therefore it is suggested to implement measures to restrict the total number of pedestrians who flow in from outside the station.

References

1. Accident Investigation Report on the Firework Festival at the 32nd Akashi City Summer Festival, pp 84–114, January 2002
2. Kaneda T (2004) Pedestrian flow agent-based simulation, measurement and control. 43(12), pp 944–949 (in Japanese)
3. Kaneda T, Suzuki T (2005) A simulation analysis for pedestrian flow management. In: Terano T, Kita H, Arai K, Degchi H (eds) Agent-based simulation from modeling methodologies to real-world applications. Springer, New York, pp 220–229
4. Cui Q-L, Taniguchi H, Kaneda T (2008) Crowd simulation of terminal station transfer passage when Tokai earthquake advisory information is announced officially. J Soc Saf Sci 10, pp 153–159 (in Japanese)
5. Deguchi H, Kanatani Y, Kaneda T, Koyama Y, Ichikawa M, Tanuma H (2006) Social simulation design for pandemic protection. In: Proceeding of the first World Congress on social simulation, vol 1, pp 21–28
6. Fujioka M, Ishibashi K, Tsukagoshi I, Kaji H (2001) A research of the multi-agent simulation model for evacuees escaping from Tsunami disaster. Paper presented at sixth inter University seminar on Asian megacities in Taipei, Taiwan, Proceedings of the sixth inter University seminar on Asian Megacities (6) pp 2–10

chen.shuheng@gmail.com

7. Japan Society of Civil Engineers (1981) Transportation demand forecasting handbook, Chapter 2, Gihodo Shuppan
8. Ministry of Construction (1982) Report on comprehensive technology development project about city disaster prevention measures technique by Ministry of Construction
9. CTI Engineering Co., Ltd (2002) Report on survey and research for measures to support pedestrians remaining in front of Nagoya Station, fiscal year, pp 80–123

chen.shuheng@gmail.com

Part VI
Miscellany

chen.shuheng@gmail.com

Boundary Organizations: An Evaluation of Their Impact Through a Multi-Agent System

Denis Boissin

Abstract Modern environmental issues imply that decision-makers consider simultaneously various dimensions, such as science and economics. To take into account opinions from experts of different fields, they can rely on boundary organizations, institutions able to cross the gap between different domains and act beyond the boundaries. By encouraging a flow of useful information, they provide a better understanding of a situation characterized by uncertainty, increasing the efficiency of the decision-making process. Though never formally proved, this hypothesis is widely accepted based on the observation of existing institutions. In this paper, we observe the impact of boundary organizations through an agent-based model of continuous opinion dynamics over two dimensions where heterogeneous experts distinguished by credibility and uncertainty interact. We conclude that boundary organizations significantly reduce the diversity of opinions expressed and increase the number of experts agreeing to emerging positions, which confirms their positive impact on the efficiency of decision-making.

Keywords Boundary organization · Opinion diffusion · Decision-making · Agent-based model

1 Introduction

Faced with modern environmental issues characterized by uncertainty and complexity, decision-makers must refer to experts from different areas of science, or even from different sciences, such as natural and social sciences, in order to consider all the aspects involved and take the most objective decision possible. Traditional decision-making provides independent advice from experts of the different dimensions involved, but these opinions may be conflicting when considered

D. Boissin (✉)
SKEMA Business School & GREDEG, UMR 6227 CNRS, c/o SKEMA Business School, BP 85, 06902 Sophia Antipolis, France
e-mail: denis.boissin@skema.edu

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_13, © Springer 2011

chen.shuheng@gmail.com

together. Boundary organizations have been designed to manage the meeting of distinct areas: by initiating and framing debates between experts, they provide decision-makers with a panel of opinions that integrate the various dimensions of an issue. The resulting eased and increased interaction between experts facilitates the decision-making process by encouraging the emergence of dominant opinions. Though this hypothesis is widely accepted, it has never been formally proved. Through an agent-based model of continuous opinion dynamics over two dimensions, it is possible to simulate a boundary organization and assess its impact on the positioning of experts: heterogeneous agents, differentiated by credibility and uncertainty, are separated in two distinct groups of experts and left free to modify their opinion through one-to-one interactions in their respective field. The boundary organization is introduced through agents open to trans-disciplinary discussion: able to cross the boundary between the two areas, they open possibilities of exchange on both dimensions between agents. This multi-agent system allows us to evaluate the impact of boundary organizations on the diversity of final opinions expressed and on the number of agents agreeing to each, testing the hypothesis supporting the existence of boundary organizations.

2 Boundary Organizations

Scientific knowledge is essential to a sound decision-making process, but science is more than a simple reservoir of knowledge, competencies and people: it includes normative concepts such as objectivity, honesty, neutrality and truth that give it an institutional and ideological privileged status in the elaboration of public policy [1]. Science includes in reality sciences, as it covers not only different areas of expertise, but also different approaches such as natural and social sciences, and decision-makers must consider all the dimensions involved in the issue considered. This implies the ability to cross the boundaries between different areas and types of sciences. The concept of boundary has been formalized by sciences in order to strengthen their differences with pseudo-sciences and scientific impostures. Boundaries protect organizations from the outside and maintain an internal order, while imposing the organization as a major actor through its relations with other organizations [2]. They allow members to affirm their authority as experts over a field challenged by others, to maintain a monopoly by excluding others, and to protect the autonomy of the members while enforcing their cohesion. Boundaries can be of three types: physical, social and mental. Physical boundaries may be real objects or structures, but also rules and regulations that frame the exchanges within the organization or between the organization and its environment. These boundaries ensure a certain predictability, synonym of stability, while creating and enforcing an image of solidity toward the outside. Social boundaries are developed so that organizations may distinguish themselves from each other: they bring a notion of identity to members while giving the opportunity to identify what constitutes others, what is not part of the organization. Mental boundaries allow for distinctions and give individuals a meaning to the world that surrounds them [2]. When two fields, under the

chen.shuheng@gmail.com

authority of different experts, are involved in an issue and brought to interact, it is a natural reaction for each to reinforce the demarcation, in order to avoid confusion and to clarify the responsibilities. While boundaries play their role to protect an organization, they also set a barrier that limits or even prevents flows of information with the outside: their reinforcement results in a lack of communication between experts of the different aspects of an issue. The maintenance of a blurry boundary, rather than the clear and intentional distinction traditionally applied, can increase the productivity of policy-making [3]. Boundary organizations have been designed to manage the meeting of distinct areas of expertise and frame interactions in order to enhance the efficiency of the decision making process: by handling debates between experts, they ensure that science brings in a pertinent and useful information while maintaining its independence.

Boundary organizations are institutions that cross the gap between different fields: they are able to act beyond the boundaries while remaining accountable to each side [4]. They encourage and manage a blurring of the boundaries, permitting an exchange to take place while maintaining the authority of experts [5, 6]. By integrating the demarcation, they allow for communication instead of division: each side can express its reactions to the other's expectations, leading to cooperation around common interests [2]. Miller defines boundary organizations as "organizations that sit in the territory between science and politics, serving as a bridge or an interface between scientific research, political decision and public action" [7], and Guston defines them as "institutions that internalize the provisional and ambiguous character of the apparent boundary between science and politics" [8]. Boundary organizations are similar to an interface established and influenced by both sides, but independent. It may look like they face a reductive double set of constraints, but groups of experts seen as independent or even opposed distinct social organizations, are more similar than it seems [7], at least in their structure and behavior. The double responsibility of boundary organizations makes them in fact stronger, as if their structure was held on both sides, giving it a unique support that guarantees impartiality [1]: this dependence of boundary organizations on each side is as important as their independence [4]. A boundary is not an established limit between two different areas of authority, but an intermediary zone of variable size: the boundary is permanently defined, criticized, challenged, defended and adjusted. Boundary organizations are not fighting against a strong solid demarcation, but in reality helping to stabilize or even create the boundary. They do not limit themselves to the zone between two areas, but extend inside each side, widening the boundary zone in order to internalize the possible areas of ambiguity. The goal is to involve both sides in the construction of a common boundary that is favorable to each perspective, while setting the limits to potential intrusion of one sphere into the other: the boundary organization must allow and encourage the interactions by increasing the permeability of the separation, while guaranteeing that they don't mix irresponsibly by limiting the porosity [9]. Unlike most organizations, their goal is not their personal benefit but that of both sides of the boundary: they do not aim at ensuring their survival, but at offering a complete and honest vision of a situation by encouraging the production and sharing of knowledge in order to guarantee a decision-making process as objective and

chen.shuheng@gmail.com

efficient as possible. Therefore they must remain impartial and neutral in the debates they frame. The boundary organization is judged on criteria of credibility, pertinence and legitimacy, similarly to the expert members and the global decision-making process. The problem of the evaluation of the efficiency of a boundary organization comes from the fact that many parameters, external to the organization and on which it has no control, may affect positively or negatively the results, such as the historical context or the characteristics of the evaluation process. Efficient boundary organizations are those remain stable despite external pressures and an internal instability of the boundary. Boundary organizations may in fact be applied to numerous cases of boundaries: between science and non-science, like historically done, between science and politics, like currently done, but also between different fields of science or different types of sciences such as natural and social sciences like modern global environmental issues may benefit from. Boundary organizations are not a new concept, but modern successful applications, such as the Health Effects Institute, the Office of Technological Assessment, the Agricultural Extension or the International Research Institute for Climate Prediction, demonstrate the diversity and utility of such institutions [1]. Boundary organizations lead to a decision-making structure able to integrate knowledge from different dimensions into a single analysis. Decision-making with respect to technological choices that enhance the well-being of society by modifying the man–environment relationship, associated with risk and uncertainty, requires to take into consideration norms and practices from natural sciences and economics. Boundary organizations appear as a solution to integrate the interactions between the different sources of information involved in environmental issues. At this date, no boundary organization has taken on this exact role, but it has been suggested as a possible evolution of existing organizations, such as the European Environment Agency [10]. As the current level of globalization increases the temporal and spatial scopes of risks, boundary organization could be an interesting solution for decision-making related to environmental issues, especially if their theoretical efficiency can be fully established.

3 Methodology

The hypothesis that supports the existence of boundary organizations is that the resulting eased and increased interaction facilitates the attainment of dominant opinions among experts of different fields. Though never formally proved, this is accepted based on the observation of existing boundary organizations [1]. Through an agent-based model, we can assess the impact of a boundary organization on the diffusion of opinions and final positioning of experts of similar and different domains. The methodology is based on simulations of opinion diffusion where experts of different fields positioned on a continuous model of opinion interact and modify their positions through series of one-to-one discussions; once the system is stabilized, we observe the number of opinions expressed, and the ratio of experts agreeing to each. A boundary organization of increasing importance is simulated to see the impact

chen.shuheng@gmail.com

on those indicators. The model relies on a multi-agent system where autonomous heterogeneous agents interact: replicated series of experimentations over ranges of parameters allow us to observe an emerging recurrent macroscopic behavior resulting from microscopic interactions that could not be deduced by simply aggregating the properties of the agents [11]. Since opinions can be more or less positive or negative, they are modeled using a continuum going from an absolute negative to an absolute positive, rather than through a binary approach. For example, positive unconditional opinions are rare as they will generally be accompanied with constraints or restrictions and be in fact more or less positive. Opinions may also change over the entire spectrum along the process. Negative positions can be definitive, or simply temporary: they can be reviewed if new information becomes available, such as through the application of the precautionary principle, and even become positive. The model is based on work done on a single dimensional model of continuous opinion dynamics [12] extended over two dimensions, representing two independent fields of expertise such as natural science and economics. Agents are positioned at random over a two-dimensional graph, where each axis represents the range of possible opinions, from −100 to 100%, in each field of expertise involved. Agents a_i have a state vector X_i (opinion attributes with respect to the axes of the graph) and a state transition function f_i at a given time unit: they are identified by their coordinates x_a and y_a reflecting their position over the different dimensions, and are left free to interact through one-to-one exchanges at each time unit, modifying this position as a result. The choice is to rely on a model without desire, intention, or motivational function for agents, but with a belief that evolves through time with respect to interactions with other experts. Agents are reactive, with a perception–action relation and no representational function of their environment: they only show a reflex behavior with respect to their encounters, as tropic agents. They are heterogeneous agents differentiated by credibility c and uncertainty u. The credibility of an agent represents how much other agents may be influenced by this agent, with respect to their own credibility. It is used as a factor, ranging from 0 to 100%, applied to the change of position, so that the sum of the credibility factor of an agent and that of its interlocutor equals 1. If we consider an agent a and its interlocutor a', the credibility effect of the agent a' over the change of position of the agent a is:

$$\frac{c_{a'} - c_a}{2} + 0.5 \tag{1}$$

The uncertainty of an agent reflects the range of accepted opinions of potential interlocutors. It acts as the maximum distance between the position of an agent a and that of its interlocutor a' so that:

$$\left.\begin{array}{l} |x_a - x_{a'}| \leq u_a \textit{ for interactions between } x - \textit{experts} \\ |y_a - y_{a'}| \leq u_a \textit{ for interactions between } y - \textit{experts} \end{array}\right| \tag{2}$$

It is also used to influence the change of opinion, based on the uncertainty of an agent over the total uncertainty of both interlocutors, so that more uncertain agents

chen.shuheng@gmail.com

will have a greater change of position than less uncertain agents. The uncertainty effect of the agent a' over the change of position of the agent a is:

$$\frac{u_a}{(u_a + u_{a'})} \tag{3}$$

Time units represent series of one-to-one interactions where each agent chooses an interlocutor to engage into discussion and modifies its position as a result. The maximum potential change of position is based on the capacity for an agent to adopt the position of its interlocutor. It is affected by a factor inversely proportional to the overlap between the distance separating an agent from its interlocutor and its uncertainty, so that the further away agents are, the smaller the change of position. The resulting basic change over the x-axis for an agent a with an interlocutor a' is:

$$(x_{a'} - x_a) \times \left(1 - \frac{x_a - x_{a'}}{u_a}\right) \tag{4}$$

The actual change is in fact the basic change, affected by the credibility and the uncertainty effects. For an interaction over the x-axis, the change of position of an agent a with an interlocutor a' would be done according to the following formula:

$$x_a^{t+1} = x_a^t + \left[x_{a'}^t - x_a^t\right] \times \left[1 - \frac{\left|x_a^t - x_{a'}^t\right|}{u_a^t}\right] \times \left[\frac{u_a^t}{u_a^t + u_{a'}^t}\right] \times \left[\frac{c_{a'} - c_a}{2} + 0.5\right] \tag{5}$$

A change over the y-axis would follow a similar formula. Due to the heterogeneity of agents, the change of position is not reciprocal, and the interaction does not imply that both agents modify their positions, since only one may lie in the zone of uncertainty of the other.

In addition, the more interactions an agent has, the smaller his uncertainty becomes, so that it tends to zero, hence the stabilization of the system. The uncertainty of an agent, after it has changed its position, is modified as follows:

$$u_a^{t+1} = u_a^t \times \left[1 - \frac{\left|x_a^{t+1} - x_a^t\right|}{2 \times u_a^t}\right] \tag{6}$$

The goal of this model is to observe the final positioning of agents representing experts interacting over an issue, and to see how it is affected by the introduction of a boundary organization. The simulation relies on NetLogo, a programmable modeling platform developed specifically for simulating both natural and social phenomena over time, making it a logical choice for a model of interaction between natural and social sciences. In addition, the Logo language relies on mobile agents called turtles who move over a grid, by modifying their direction and choosing a length of displacement, a well suited concept for simulating experts who decide to modify their position (displacement) after identifying an interlocutor (direction). Considering that NetLogo is also free of use and cross-platform, including the possibility to be embedded in web pages, it seemed the best choice, and its weaknesses with respect to the simulation capacity were not a limit for our model.

chen.shuheng@gmail.com

The simulation involves 200 agents equally spread over two fields of expertise and is left running over 1,000 time units, each time unit giving the opportunity for each agent to choose to enter into interaction with an other agent. First, only two kinds of agents (scientists and economists) are left free to interact through one-to-one exchanges in their respective field represented by each of the axes. At each time unit, each agent chooses an interlocutor of the same kind and modifies its position and uncertainty according to (5) and (6). These reference results show that the two-dimensional projection is in accordance with the single-dimensional continuous opinion model used as a basis. The boundary organization is then introduced through agents called borgs: open to trans-disciplinary discussion, they are able to cross the boundary between the two axes, creating possibilities of exchange on both dimensions, while other agents remain limited to interactions within their field of expertise with similar agents. Borgs are regular agents who gain a new property, no matter what their initial position is, mainstream or minor, extreme or average, since the boundary organization must remain impartial and neutral in the debate and allow for all opinions to be expressed to maintain a high level of legitimacy. Boundary organizations could therefore not be modeled as a spatial zone, since it would reduce the diversity of opinions that could be expressed within the organization. The ratio of borgs among the total population of agents is increased from 0 to 50%, by steps of 1% from 0 to 10% and by steps of 5% beyond 10%, with ten simulations at each value. The position of experts is recorded every 10 units of time. The results are analyzed in terms of the number of opinions expressed by a group of experts representing at least 1% of the total number of agents (i.e. at least two agents sharing the same opinion), and the ratio of experts agreeing to each of these opinions once the positions are stabilized.

4 Results

Considering the number of distinct final opinions expressed, the impact of a boundary organization of increasing importance is significant. The total number of opinions representing each at least 1% of all experts is reduced by 11% when 5% of agents are borgs; 10% of borgs represent a reduction by 22 and 30% lead to a decrease of 32%, with no variation beyond this level. Figure 1 shows this impact through a logarithmic regression.

This global reduction of final opinions expressed is really due to a decrease of the minor opinions, i.e. distinct opinions that represent at least 1% but no more than 2% of all experts. Their number is reduced by 11% with 5% of bores, by 31% with 10% of borgs, to reach a maximum decrease of 58%, which is represented by a logarithmic regression of greater slope on Fig. 2.

If we consider the distinct final opinions representing each at least 5% of all agents, the impact of the boundary organization is completely inverse: their number is slightly positively affected, with a linear relationship to the size of the boundary organization as shown by Fig. 3.

chen.shuheng@gmail.com

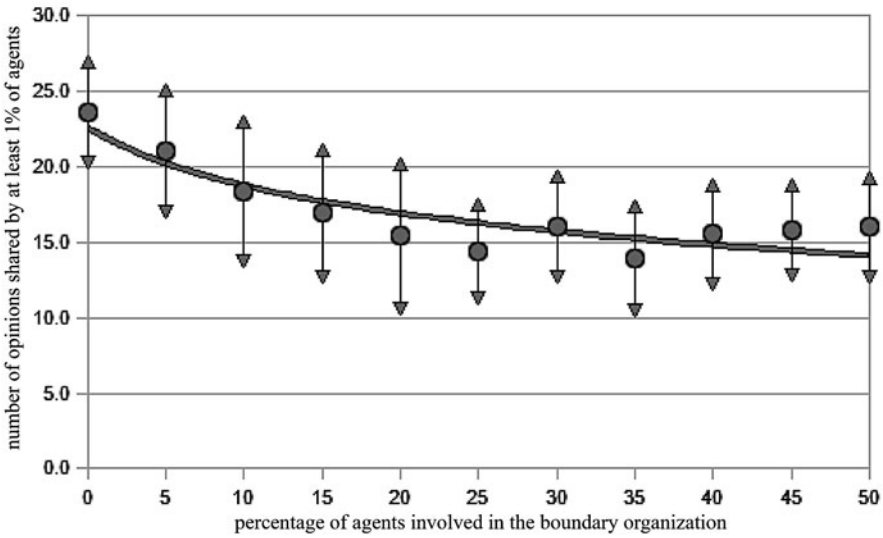


Fig. 1 Average number, standard deviation and logarithmic regression of final opinions expressed shared by at least 1% of all agents with respect to the size of the boundary organization

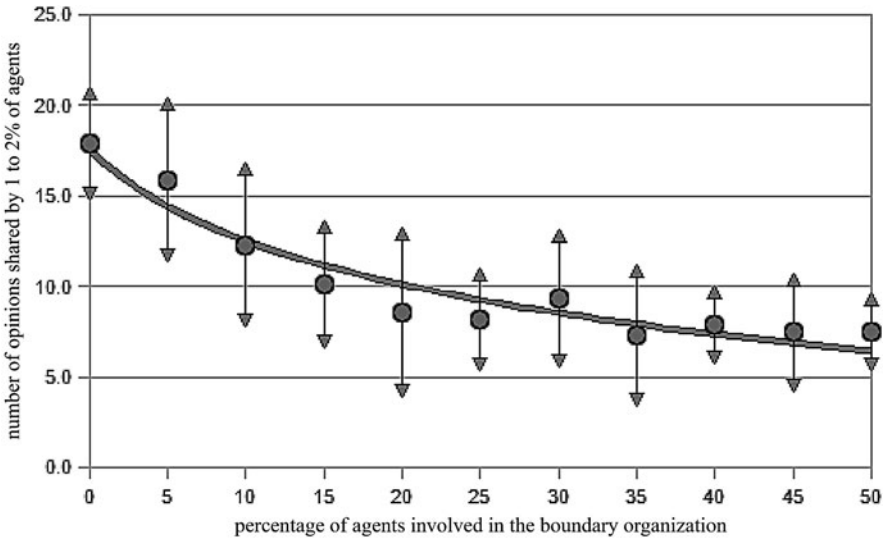


Fig. 2 Average number, standard deviation and logarithmic regression of final opinions expressed shared by 1–2% of all agents with respect to the size of the boundary organization

chen.shuheng@gmail.com

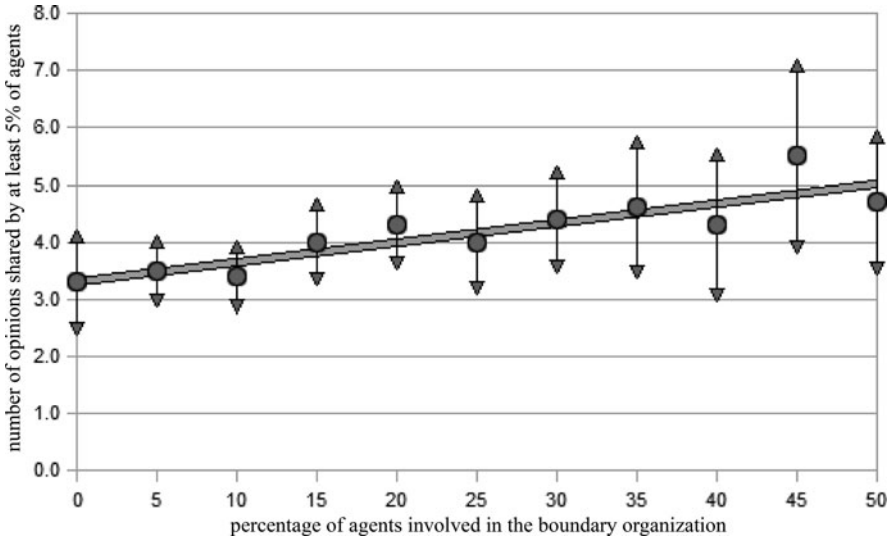


Fig. 3 Average number, standard deviation and linear regression of final opinions expressed shared by at least 5% of all agents with respect to the size of the boundary organization

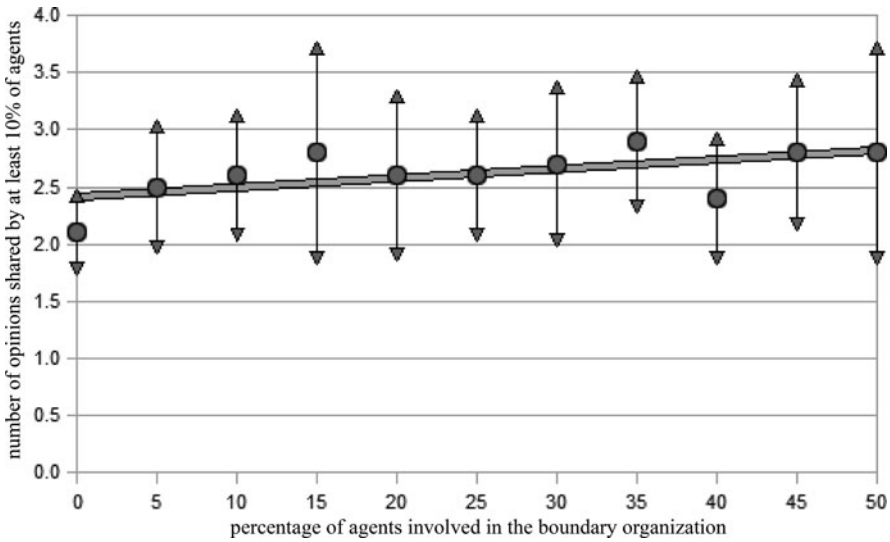


Fig. 4 Average number, standard deviation and linear regression of final opinions expressed shared by at least 10% of all agents with respect to the size of the boundary organization

Figure 4 reveals a similar situation for opinions supported by at least 10% of agents: they become slightly more numerous with a boundary organization of increasing importance.

chen.shuheng@gmail.com

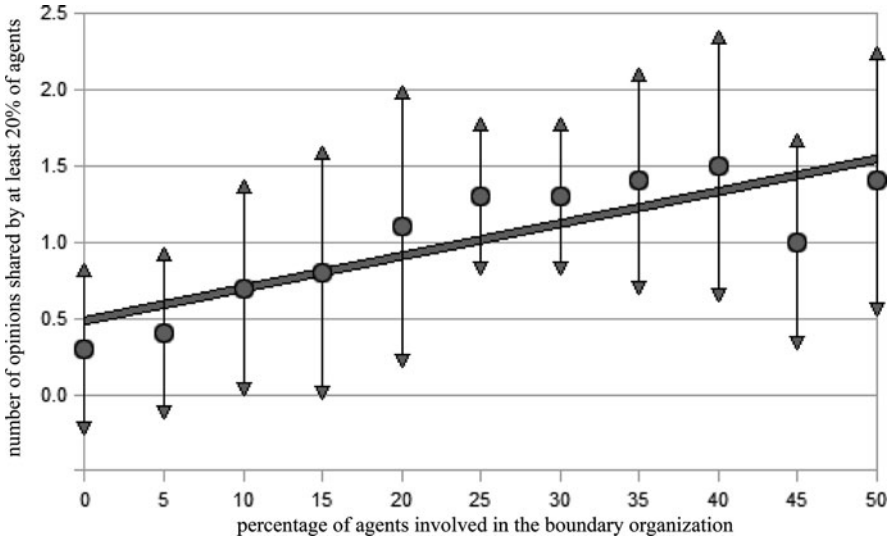


Fig. 5 Average number, standard deviation and linear regression of final opinions expressed shared by at least 20% of all agents with respect to the size of the boundary organization

The greatest positive impact is on opinions gathering at least 20% of the agents, since they can be multiplied by up to 3 as seen on Fig. 5.

The impact of a boundary organization of greater importance on the reduction of the number of final distinct opinions expressed is significant and not linear, with the stabilization of the impact at a certain level. At low realistic levels of 10–20% of agents involved in the boundary organization, the impact is immediate on the reduction of the number of different minor opinions expressed and on the increase of the number of opinions gathering the largest shares of experts, while average opinions are not greatly affected: the global reduction of the diversity of opinions is an apparent transfer from minor to dominant opinions. When we consider the ratio of agents agreeing to each of the final opinions expressed, the impact is significant. If we consider the dominant final opinion, 5% of borgs are sufficient to increase the number of agents agreeing by 19% to reach a maximum increase of 25%. If we consider the sum of agents agreeing to the two main opinions, the increase is reduced to 17% at 5% of borgs but reaches 23% at 10%, for a maximum of 43%. A similar situation is observed for the five main opinions with an increase of 20% at 5% of borgs and a maximum of 41%. The sum of the ten main opinions increases by 15% with 5% of borgs, by 20% with 10%, and by 35% with 20% to reach a maximum of 45%. Figure 6 shows the total share of agents agreeing to the 1, 2, 5 and 10 dominant opinions with respect to the size of the boundary organization.

chen.shuheng@gmail.com

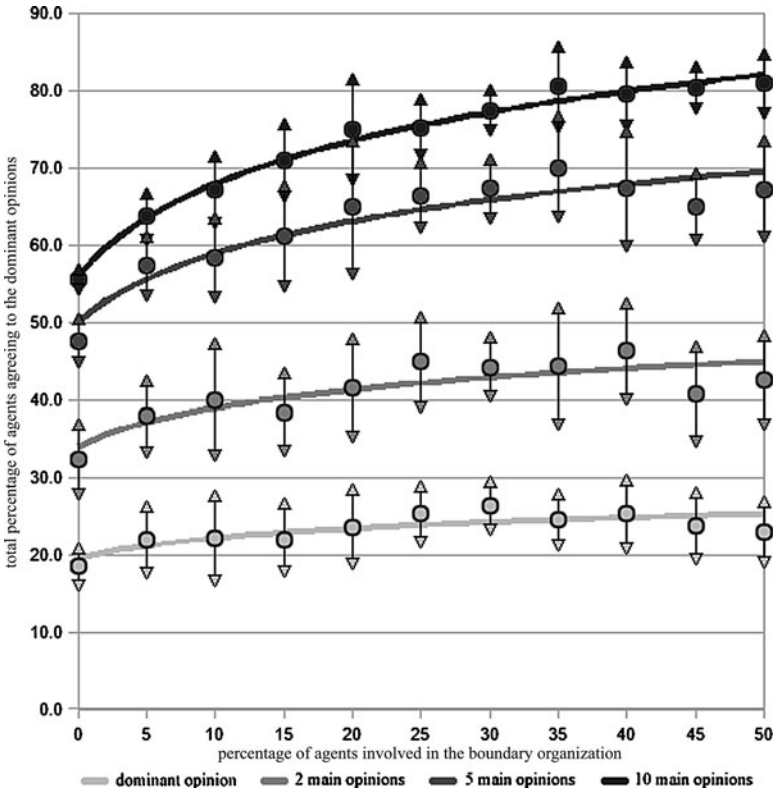


Fig. 6 Average value, standard deviation and logarithmic regression of the ratio of agents agreeing to the 1, 2, 5 and 10 main final opinions expressed with respect to the size of the boundary organization

5 Conclusion

We can confirm the impact of the boundary organization on the share of agents supporting dominant opinions by looking at the number of different opinions necessary to represent at least 50% of all agents. Without a boundary organization, we need to aggregate the seven main opinions to represent 50% of the experts, when it requires six with 1% of borgs, five with 3%, four with 5% and only three with 15%, with no variation beyond this level. The concentration of agents around dominant opinions is significantly increased by the existence of a boundary organization.

The rising interest for boundary organizations, supported by observed successful cases, is confirmed by our agent-based model: boundary organizations do not require the involvement of a large share of experts to show a significant impact on the reduction of the diversity of final different opinions expressed and on the increase of the concentration of experts around dominant opinions, resulting in an apparent transfer from minor to major opinions. It is then easier for decision-makers

chen.shuheng@gmail.com

to analyze the different aspects of an issue altogether: they face less opinions to consider, and are presented with more affirmed dominant opinions among the experts consulted. Boundary organizations seem to be able to increase the scale of confrontation between groups of opinion: their agents do not emerge as opinion leaders, but encourage the exchanges between experts by easing and increasing the transfer of information from one sphere to the other, which results in more affirmed positions of experts over the different dimensions of an issue.

References

1. Guston D, Clark W, Keating T, Cash D, Moser S, Miller C, Powers C (2000) Report of the workshop on boundary organizations in environmental policy and science, 9–10 December 1999. Global Environmental Assessment Project, John F. Kennedy School of Government, Harvard University, Cambridge
2. Davenport S, Leitch S (2005) The role of boundary organisations in maintaining separation in the triple helix. In: Triple Helix 5 Conference, Turin
3. Jasanoff S (1990) The fifth branch: science advisers as policy makers. Harvard University Press, Cambridge
4. Guston D (2001) Boundary organizations in environmental policy and science: an introduction. Sci Technol Human Values 26(4):399–408. SAGE Publications
5. Cash D, Clark W, Alcock F, Dickson N, Eckley N, Guston D, Jäger J, Mitchell R (2003) Knowledge systems for sustainable development. Science and technology for sustainable development special feature ecology. Proc Natl Acad Sci USA 100(14):8086–8091
6. Clark W, Mitchell R, Cash D, Alcock F (2002) Information as influence: how institutions mediate the impact of scientific assessments on global environmental affairs. Faculty Research Working Papers Series, RWP02-044, John F. Kennedy School of Government, Harvard University, Cambridge
7. Miller C (2000) Boundary organizations: strategies for linking knowledge to action. Draft based on the Dec. 9 workshop on boundary organizations in environmental policy and science, Global Environmental Assessment Project, John F. Kennedy School of Government, Harvard University, Cambridge
8. Guston D (2000) Between politics and science: assuring the integrity and productivity of research. Cambridge University Press, New York
9. Socci A (2001) Boundary organizations and ethics in science. In: Global Climate Change and Society proceedings, Boulder
10. Scott A (2000) Scoping report for the European Environment Agency, Experts' corner report. SPRU (Science and Tecnhology Policy Research), Environmental issues series no. 15
11. Axelrod R, Tesfatsion L (2006) A guide for newcomers to agent-based modeling in the social sciences. In: Tesfatsion L, Judd K (eds) Handbook of computational economics, vol. 2: agent-based computational economics. Amsterdam
12. Deffuant G, Neau D, Amblard F, Weisbuch G (2001) Mixing beliefs among interacting agents. Adv Complex Syst 3(1):87–98. World Scientific Publishing, Singapore

chen.shuheng@gmail.com

A Bibliometric Study of Agent-Based Modeling Literature on the SSCI Database

Shu-Heng Chen, Yu-Hsiang Yang, and Wen-Jen Yu

Abstract The purpose of this study is to investigate the characteristics of the international literature using Agent-Based Modeling (ABM) in SSCI during 1997–2009. The results of this study reveal the fact that the growth of international literature using ABM is still well perceived. Most of the literature is from various institutions in the USA. According to Bradford's Law, eight core journals in ABM are identified and analyzed. Moreover, the frequency distributions of the author productivity match the generalized Lotka's Law. Applications of ABM are mainly found in the fields of social science/interdisciplinary studies, economics, and environmental studies.

Keywords Bibliometrics · Agent-Based Modeling · Social Sciences Citation Index (SSCI) · Bradford's Law · Lotka's Law

1 Introduction

Since several pioneering papers on Agent-Based Modeling (ABM) were published in early 1970s [7, 8], the exploration of the ABM literature has seen vigorous development, especially in the last decade owing to availability of various ABM software. This paper employs a bibliometric methodology that is geared towards a review of literature productivity, and an observation of the trend in ABM. In order to have a better understanding of the quantitative aspects of recorded information and discover literature features and forecast the research trends in the near future, this paper uses Lotka's Law to perform author productivity analysis and applies Bradford's Law to the core journals in this field from 1997 to 2009.

S.-H. Chen
AI-ECON Research Center, Department of Economics, National Chengchi University, Taipei, Taiwan
e-mail: chen.shuheng@gmail.com

Y.-H. Yang (✉) and W.-J. Yu
Department of Management Information System, National Chengchi University, Taipei, Taiwan
e-mail: yuxiang1001@gmail.com; yuwenjen@gmail.com

S.-H. Chen et al. (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VI: Post-Proceedings of The AESCS International Workshop 2009*, Agent-Based Social Systems 8, DOI 10.1007/978-4-431-53907-0_14, © Springer 2011

chen.shuheng@gmail.com

2 Agent-Based Modeling Description

Agent-Based Modeling (ABM) refers to the computer simulation of agents (representing individual roles) in a dynamic social system. Here, agents refer to different "representatives" who interact with each other or the environment based on pre-set rules. Through these representatives, we may be able to observe the emergence of certain macro behaviors. Derived from the Schellings Segregation Model (SSM) [8], ABM has been applied in sociology, economics, political science and ecology to explore the phenomenon of Complex Adaptive Systems (CAS), and has gradually been more widely used in almost every field of social sciences for a deeper understanding of social phenomena as emergent properties. Chen [3] points out that the economic system in agent-based economics is composed of heterogeneous agents and that those aggregate variables are the results of these heterogeneous agents' interactions. ABM serves as an ideal tool for us to advance our thinking from the micro to the macro perspective, and to observe the links and relationships between these two levels. Unlike the "top-down" mode of thinking in traditional macroeconomics, ABM has introduced a "bottom-up" style of thinking to macroeconomics, which presents both challenge and opportunities to most economists.

There is also a growing trend toward the application of ABM in political studies, for it does not focus on the causal relations between variables, as statistics and econometrics do. Instead, it is mainly concerned with addressing "how" or "what-if" questions – observing how the complicated social/political phenomena in question have been formulated through the interaction between the simulated agents. In addition, the patterns being discovered through such observations may be used either to test existing theories or to explore new ones [2]. Robert Axelrod argues that ABM can be used to describe certain fundamental questions in many fields, thereby promoting interdisciplinary cooperation. Moreover, when existing mathematical methods fall short, ABM presents itself as a useful tool to reveal the underlying unity behind various academic fields.

3 Overall Analysis of ABM literature

This paper utilizes the Social Sciences Citation Index (SSCI) of the Web of Science created by the Institute for Scientific Information. An empirical method of retrieval was used by Topic = ("agent based(*)") OR Topic=("multi-agent simulation") to retrieve data related to ABM. A total of 1,051 papers in ABM published during 1997–2009 were found. Figure 1 indicates the number and growth of published papers in ABM. According to the numerical data, a large quantity of research papers published during 2005–2009 has been catalogued in the SSCI database. These together with their respective distribution rate set against the total number of papers are 93 (8.85%), 112 (10.66%), 159 (15.19%), 200 (17.88%) and 196 (18.65%) for those 5 years. Meanwhile, Fig. 2 shows the annual citations of the published papers in ABM. The results appear to suggest that the number of papers in ABM

chen.shuheng@gmail.com

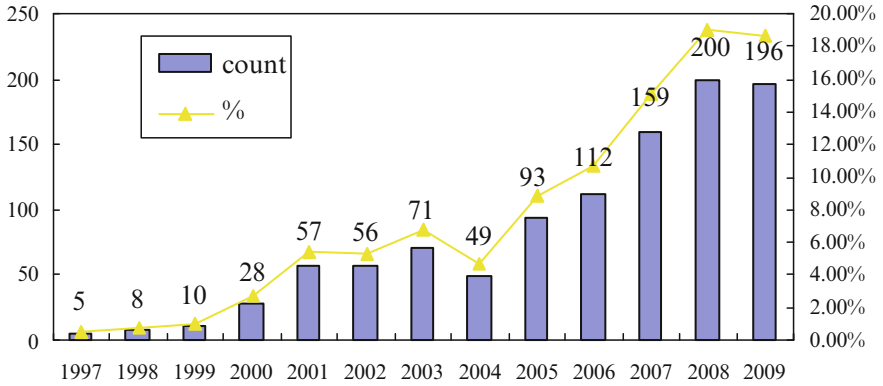


Fig. 1 Number of published papers in ABM

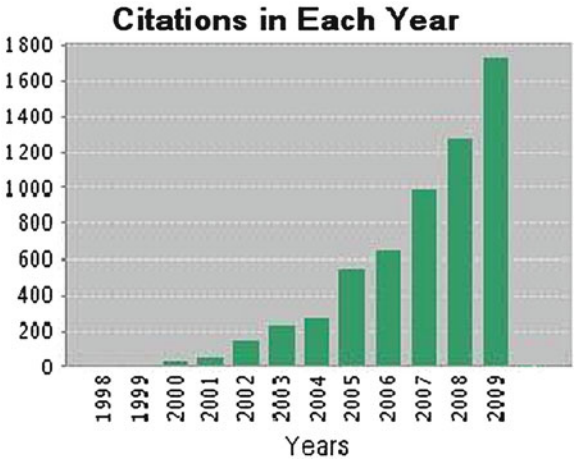


Fig. 2 Citation in each year (Source: SSCI database)

has distinctively increased since 2001, and that respective citations have also increased in each year. It appears that ABM has received much attention from researchers, which has led to a rapid growth of related papers and their citations.

With regard to the countries over which the 1,051 papers distributed, the top ten countries ranked in the most publicized catalogues in the SSCI database during the 1997–2009 period are illustrated in Fig. 3. The figure shows that the US is the dominant country (401 papers; 38.15%) in terms of the number of published papers in ABM, followed by England (110 papers; 10.47%), Germany (88 papers; 8.37%) and so on.

Table 1 offers a closer look at the distribution of academic institutions by which the indexed papers were submitted. It is observed that, among the 11 institutions whose numbers of indexed papers are greater than or equal to 11, the USA is the most productive country with eight institutions listed in the study of ABM.

chen.shuheng@gmail.com

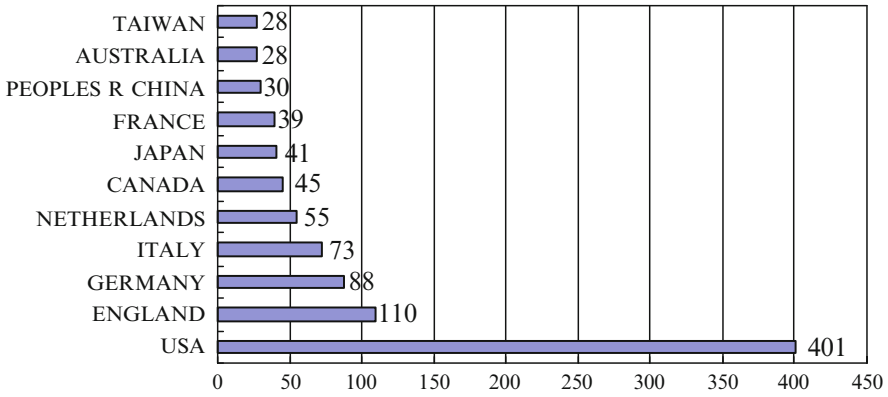


Fig. 3 The top ten numbers of published papers in ABM based on countries during 1997–2009

Table 1 Leading institutions with the most published papers in ABM during 1997–2009

Institution name	Count	%	Comprising % of the country	Country
Univ Michigan	25	2.38	6.23	USA
Univ Penn	20	1.90	4.99	USA
Univ Illinois	16	1.52	3.99	USA
Harvard Univ	14	1.33	3.49	USA
Santa Fe Inst	14	1.33	3.49	USA
UCL	14	1.33	12.73	England
Univ Groningen	14	1.33	25.45	Netherlands
Indiana Univ	13	1.24	3.24	USA
George Mason Univ	12	1.14	2.99	USA
Natl Chengchi Univ	12	1.14	42.86	Taiwan

Table 2 offers an investigation into the authors who have written more than six papers in ABM during 1997–2009. The most productive authors are Jager (10, Netherlands), Jassen (10, USA), Chen (9, Taiwan), Gallegati (9, Italy), Brown (8, USA), Bunn (8, England), and Izquierdo (8, Spain). The data show that the corresponding ratios for Netherlands, Taiwan, Italy, and Spain are much greater than that for the USA, indicating that these authors in their country dominate the academic research in the ABM field. Table 2 also shows that the research in economics using ABM is in the mainstream. This may result from the fact that ABM offers a macroscopic view of evolution through a microscopic view of agents' interactions.

With respect to the future directions of ABM research, the emphasis of the discussion here is on the ABM applications. Based on the results we retrieved from the SSCI database, Table 3 provides the top ten subject areas in which ABM is most widely utilized. The first ranked subject area, Social Sciences/interdisciplinary, is comprised of 213 papers (20.27%) against the total of 1,051 papers retrieved. The second ranked area is Economics, with 154 papers (14.65%) followed by Management with 91 papers (8.66%) related to ABM. If we group the three subject areas of computer science as one subject area, it is obvious that the published

chen.shuheng@gmail.com

Table 2 The top-ranking published papers in ABM based on authors during 1997–2009

Author	Count	%	% at his/her country	Country	Institution	Subject area
Jager, W	10	0.95	25.64	Netherlands	Univ Groningen	Economics
Janssen, MA	10	0.95	3.16	USA	Arizona State Univ	Ecology
Chen, SH	9	0.86	36	Taiwan	Natl Chengchi Univ	Economics
Gallegati, M	9	0.86	15.79	Italy	Univ Politecn Marche	Economics
Brown, DG	8	0.76	2.53	USA	Univ Michigan	Natural Resources and Environment
Bunn, DW	8	0.76	8.7	England	London Business Sch	Management
Izquierdo, LR	8	0.76	44.44	Spain	Univ Burgos	Social Sciences
Bhavnani, R	7	0.67	2.22	USA	Michigan State Univ	Political Sci
Dawid, H	7	0.67	10.94	Germany	Bielefeld University	Economics
Lebaron, B	7	0.67	2.22	USA	Brandeis Univ	Economics
Miodownik, D	7	0.67	70	Israel	Hebrew Univ	Polit Sci
Polhill, JG	7	0.67	7.61	England	Macaulay Institute	Computer Science

Table 3 The top ranking of published papers in ABM based on subject areas during 1997–2009

Subject area	Count	%
Social Sciences/Interdisciplinary	213	20.27
Economics	154	14.65
Environmental Studies	91	8.66
Mathematics, Interdisciplinary Applications	88	8.37
Management	81	7.71
Computer Science/Interdisciplinary Applications	74	7.04
Operations Research & Management Science	74	7.04
Computer Science/Artificial Intelligence	68	6.47
Computer Science/Information Systems	61	5.8
Business	55	5.23
Social Sciences/Mathematical Methods	55	5.23

papers in computer science will be more in number than those of other subject areas because ABM relies heavily on programming. By referring to Fig. 4, the top five subject areas are gradually growing year by year.

To sum up, the ABM literature is still being developed based on the retrieval performance observed across the SSCI database. The top three countries with the most published papers in the field of ABM are the USA, England, and Germany. In recent years, the published papers using ABM methodology have still been growing because of the maturity and ease of use of some ABM simulation platforms such as SWARM of the Santa Fe Institute (SFI), Starlogo of MIT, Netlogo of Northeastern University, and REPAST of the University of Chicago. ABM tools enable subject areas such as social science/interdisciplinary study, economics, management, mathematics/interdisciplinary applications and environmental studies to conduct research in natural and social science in order to provide human beings with knowledge.

chen.shuheng@gmail.com

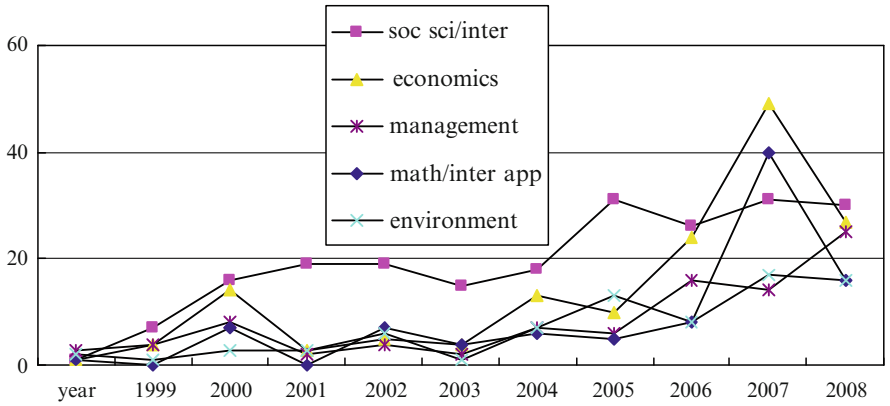


Fig. 4 Yearly distribution of top five subject areas during 1997–2009

4 Bradford's Law and Journal Literature

Samuel C. Bradford (1878–1948) introduced Bradford's Law in 1934 which is based on a pattern that estimates the exponentially diminishing returns of extending a search for references in science journals. The principle imposes a formulation that if journals within a field are sorted based on the number of articles into three groups, with each group which is concerned with comprising approximately one-third of all articles, then the number of journals in each group will be proportional to $1{:}n{:}n^2$ [9]. Bradford's Law (about scattering of subjects in information sources) is often mentioned together with Zipf's Law (about word frequencies in natural language texts) and Lotka's Law (with regard to the distribution of authors' productivity) as one of the three most important bibliometric laws. These three laws are often considered to be the best models or examples of research resources that are available within the Library and Information Sciences.

The 1,051 published papers referred to in this study are distributed among 335 journals. Table 4 provides the number of published papers in each journal and other information ranked by the number of published papers according to the zoning of Bradford's Law. Besides, the number of published papers in the top eight journals is about one third of the 1,051 published papers (334, 32%). The other 717 published papers (68%) are distributed among 327 journals, including one published paper in each of 198 journals. The results show that the distribution of published papers in ABM is decentralized.

Table 5 also provides the ratio comparisons of three zones, that is the ratio of published papers in each zone for zones A, B, C is 8:48:279. It is almost equal to $8{:}48{:}288$ or $1{:}6{:}6^2$. That is, $\text{A:B:C} = 1{:}n{:}n^2$. The result matches the explanations of Bradford's Law. Table 6 specifies eight leading journals which have published the most research papers in ABM. According to the data, ABM papers published in these journals take up to nearly one-third of the total amount. *Journal of Artificial*

chen.shuheng@gmail.com

Table 4 The distribution of ABM journals during 1997–2009

	(A)	(B)	(C)	(D)	(E)	(F)
(I) Core	174	1	1	174	174	2.241
	29	1	2	29	203	2.307
	28	1	3	28	231	2.364
	25	2	5	50	281	2.449
	21	1	6	21	302	2.480
	17	1	7	17	319	2.504
	15	1	8	15	334	2.524
(II) Relevant	12	2	10	24	358	2.554
	11	3	13	33	391	2.592
	10	5	18	50	441	2.644
	9	2	20	18	459	2.662
	8	7	27	56	515	2.712
	7	7	34	49	564	2.751
	6	3	37	18	582	2.765
	5	8	45	40	622	2.794
	4	11	56	44	666	2.823
(III) Marginal	3	25	81	75	741	2.870
	2	56	137	112	853	2.931
	1	198	335	198	1051	3.022

Column (A) number of articles. (B) number of journals. (C) accumulation of journals. (D) (A)×(B), subtotal of articles. (E) accumulation of articles. (F) $\log^E$

Table 5 A brief distribution of the literature on ABM based on the journals

	No. of journals	No. of articles	Range of no. of articles	Average articles
A	8	334	15–174	42
B	48	332	4–12	7
C	279	385	1–3	1

Table 6 The eight core journal titles and their statistics in ABM

Journal title	Count	%	Acc. %
JASSS – The Journal of Artificial Societies and Social Simulation	174	16.56	16.56
Environment and Planning B-Planning & Design	29	2.76	19.32
Physica A – Statistical Mechanics and Its Applications	28	2.66	21.98
Journal of Economic Behavior & Organization	25	2.38	24.36
Journal of Economic Dynamics & Control	25	2.38	26.74
Advances in Complex Systems	21	2	28.73
Social Science Computer Review	17	1.62	30.35
Computational Economics	15	1.43	31.78

Societies and Social Simulation (JASSS), which tops the list, has 174 published papers (16.56%) in ABM. It largely outnumbers that of the second journal on top, *Physica A* (29 papers, 2.76%). It is also observed that the main subject areas of the eight listed journals are Economics and Social Science/Interdisciplinary.

chen.shuheng@gmail.com

5 Lotka's Law and Author Productivity

5.1 Distribution of Scientific Productivity of Authors with Equality of Chances of Participation

With 1,051 papers retrieved, based on the Equality method, there are altogether 1,887 authors who jointly contributed to a published paper, there being an average of 0.56 papers per author. Within these retrieved papers, one author per paper (with no jointly accumulated authors) applies in the majority of cases, which comprise a total of 1,887 persons, or 80.23%. Meanwhile, there are 12 papers with seven or more accumulated authors per paper. By assuming that each accumulated author generated one paper, there would be 2,536 accumulated papers, as shown in Table 7. However, in actual fact there are only 1,051 papers. As a result, the overall number of estimated authors per paper will be 2.41, means the research teams will usually consist of two or three persons.

5.2 Lotka's Law

Lotka's Law regarding the scientific productivity of authors is a good example with respect to such empirical laws. Lotka [6] deduced an inverse square law relating the authors of published papers to the number of papers written by each author. Using the data specifically represented in the decennial index of Chemical Abstracts and Auerbach's Geschichtstafeln der Physik as the name index, Lotka plotted the number of authors against the number of contributions made by each

Table 7 Distribution of author productivity among ABM papers

% of Author (A)	% of Acc. Author (B)	(C)	(D)	(E)	(F)	(G)	(H)
10	2	20	20	0.79	2	0.11	0.11
9	2	18	38	1.5	4	0.11	0.21
8	3	24	62	2.44	7	0.16	0.37
7	5	35	97	3.82	12	0.26	0.64
6	14	84	181	7.14	26	0.74	1.38
5	10	50	231	9.11	36	0.53	1.91
4	20	80	311	12.26	56	1.06	2.97
3	77	231	542	21.37	133	4.08	7.05
2	240	480	1022	40.3	373	12.72	19.77
1	1514	1514	2536	100	1887	80.23	100

Column (A) number of papers published. (B) number of authors. (C) (A) × (B) subtotal number of papers published. (D) accumulated number of papers published. (E) (D)/2536, % of total number of papers. (F) accumulated number of authors (G) (B)/1887, % of total number of authors (H) (F)/1887, % of total number of authors

chen.shuheng@gmail.com

author on a logarithmic scale. Lotka proposed that these points are closely scattered around a straight line having a constant slope of approximately *negative two*. It has been shown to hold for the productivity patterns of chemists, physicists, mathematicians, and econometricians [5]. Lotka's inverse square law of scientific productivity has been shown to fit data drawn from several widely varying time periods and disciplines [1].

Many studies have confirmed the validity of the law, and have often found that the constant c is not always 2 but is rather a variable value. The empirical observations as concluded by Lotka provided the following equation [4]:

$$a_n = a_1/n^c, n = 1, 2, 3, \ldots. \tag{1}$$

where a_n = the number of authors publishing n papers, a_1 = the number of authors publishing one paper, and c = a constant (in Lotka's case, $c = 2$). Taking the log of both sides of (1), we obtain

$$\log(a_n) = \log(a_1) - c\log(n). \tag{2}$$

In the computation of the best empirical value, the constant c for the data is related to ABM by fitting a line to the empirical frequency distribution. The regression results show that $c = 2.97$. If the estimated a_1 is 0.8282, then (1) will be stated as follows: $a_n = 0.8282/n^{2.97}$ It is possible to check whether the ABM literature matches the Lotka's Law by performing the Kolmogorov–Smirnov (K-S) statistical test. According to the K-S test, as demonstrated in Table 8, if $Dmax = 0.0259$ and the sampling number is bigger than 35, then the threshold value will be $1.63/1887^{1/2} = 0.0375$, while the number of accumulated authors will be 1,887. Although *Dmax* is less than the threshold value, the result matches the generalized Lotka's Law, which indicates that Lotka's Law is based on the author productivity distribution data in the ABM literature.

Table 8 Author distribution of Lotka's Law

Count	Observation by author(s)	Acc. value ($Sn(X)$)	Expected value by author	Acc. value ($Fo(X)$)	Absolute value $\mid Fo(X) - Sn(X) \mid$
1	0.8023	0.8023	0.8282	0.8282	0.0259 (*Dmax*)
2	0.1272	0.9295	0.1055	0.9337	0.0042
3	0.0408	0.9703	0.0316	0.9652	0.005
4	0.0106	0.9809	0.0134	0.9787	0.0022
5	0.0053	0.9862	0.0069	0.9856	0.0006
6	0.0074	0.9936	0.004	0.9896	0.004
7	0.0026	0.9963	0.0025	0.9922	0.0041
8	0.0016	0.9978	0.0017	0.9939	0.004
9	0.0011	0.9989	0.0012	0.9951	0.0038
10	0.0011	1	0.0009	0.9959	0.004

chen.shuheng@gmail.com

6 Conclusion

The purpose of this study is to investigate the characteristics of the international literature that applies agent-based modeling (ABM) to social sciences for 1997–2009. The results of this study reveal that the growth of the international literature using ABM is still well perceived. Most of the core literature belongs to various USA institutions. According to Bradford's Law, the three zone ratio comparisons are almost equal to the ratio $1{:}6{:}6^2$, means the data does match Bradford's Law. The eight core journals in ABM are identified and analyzed. Moreover, the frequency distributions of the author productivity match the generalized Lotka's Law. The three applications of ABM are mainly in the fields of social science/interdisciplinary studies, economics, and environmental studies.

Acknowledgements The first author is grateful for NSC research grant No. NSC 98-2410-H-004-045-MY3, 2009–2012, and No. NSC 98-2911-I-004 -007, 2009–2012.

References

1. Allison PD, Stewart JA (1974) Productivity differences among scientists: evidence for accumulative advantage. Am Sociol Rev 39:596–606
2. Axelrod R (2006) Agent-based modeling as a bridge between disciplines. In: Judd KL, Tesfatsion L (eds) Handbook of computational economics, vol 2: agent-based computational economics, handbooks in economics series, pp 1566–1584. Elsevier, New York
3. Chen S-H (2008) Computational intelligence in agent-based computational economics. In: Fulcher J, Jain L (eds) Computational intelligence: a compendium, Chapter 13, pp 517–594. Springer
4. Chung KH, Cox RAK (1990) Patterns of productivity in the finance literature: a study of the bibliometric distributions. J Finance 45(1):301–309
5. Krisciunas K (1977) Letter to the editor. J ASZS 28:65–66
6. Lotka AJ (1926) The frequency distribution of scientific productivity. J Wash Acad Sci 16:317–323
7. Sakoda J (1971) The checkerboard model of social interaction. J Math Sociol 1:119–132
8. Schelling TC (1971) Dynamic models of segregation. J Math Sociol 1:143–186
9. Tsay MY (2003) The characteristic of informetrics and bibliometrics. Hwa-Tai bookstore, Taipei

chen.shuheng@gmail.com

Author Index

chen.shuheng@gmail.com

Keyword Index

chen.shuheng@gmail.com

chen.shuheng@gmail.com