Large-scale agent-based social simulation
A study on epidemic prediction and control

Zhang, Mingxin

DOI
10.4233/uuid:8d0f67a3-d8e6-43ee-acc5-1633c617e023

Publication date
2016

Document Version
Final published version

Citation (APA)
Zhang, M. (2016). *Large-scale agent-based social simulation: A study on epidemic prediction and control*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:8d0f67a3-d8e6-43ee-acc5-1633c617e023

Large-scale Agent-based Social Simulation
-A study on epidemic prediction and control

Large-scale Agent-based Social Simulation
- A study on epidemic prediction and control

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
woensdag 18 mei 2016 om 10:00 uur
door

Mingxin ZHANG
Master of Engineering
National University of Defense Technology, China
geboren te NanTong, China

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. A. Verbraeck

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. A. Verbraeck	Technische Universiteit Delft, promotor

Onafhankelijke leden:

Prof. dr. A. Bruzzone	University of Genoa, Italy
Prof. dr. ir. G.L.L.M.E. Reniers	Technische Universiteit Delft
Prof. dr. B.A. van de Walle	Technische Universiteit Delft
Prof. dr. ir. J.W.C. van Lint	Technische Universiteit Delft
dr. M.M. Dastani	Universiteit Utrecht

This work was supported by the China Scholarship Council (CSC).

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems. SIKS Dissertation Series No. 2016-28

Published and distributed by Mingxin Zhang
Email: ZH.MINGXIN@GMAIL.COM

Cover design by Ruizhi Cao
Printed by CPI Koninklijke Wöhrmann

ISBN: 978-94-6328-041-9

To my family

Acknowledgment

Since the first day I arrived at the Schiphol airport in September 2011, so many people have helped me during my stay in The Netherlands. I'd like to take this opportunity to express my gratitude to all of you.

My first and deepest thank goes to my promoter Prof. Alexander Verbraeck. Alexander, I really appreciate everything you helped me to experience and to learn during my PhD study. Starting from offering me an invitation letter, I have learned so much from talking to you in weekly meetings, assisting you in the simulation classes, discussing questions with you through emails, participating in conferences together, extreme programming, assisting you to organize conferences, joining in the trips to China and USA together, and drinking beer together. All of these trained and inspired me to be a qualified independent researcher.

I would like to thank all my current colleagues. Especially, I'd like to thank Bartel, Els, Monique and Marlies for your timely help with the extension of my working permit and residence permit, which made my defense ceremony to come true. Rens, Heide, Shalini and Xu, I am so enjoyed to work with you all the time. Lotte and Floortje, thank you so much for accepting to be my paranymphs. Sharlene, Marc, Sibel, Bert, George, Jill, Scott, Erik and many other colleagues, I am so honored to work with you in such an active team with a lot of talents.

I would like to also thank all my former colleagues. Mamadou Seck, as my daily supervisor in the first two years, we had a lot of inspiring discussions on potential research directions. Yilin, Jordan and Yakup, thank you for inviting me to Delft city center for beers in the first evening of my working day, which was the origin of my hobby on dutch beers. Martijn, Michel, Kassidy, Cagri, Amir and many others, thank you for inviting me to the "pour at 4" series, which was the fuse of my addiction on different kinds of alcohols. Diones and Everdine, special thanks to your help on arranging all the complicated administrative works for many of the Chinese visitors. Stefan, Jos, Sander, Alireza, Tanja, Farideh, Maartje, Nina and Selma, I feel so lucky to meet all of you.

The project on the study of epidemics in the artificial city of Beijing received much help from many researchers at National University of Defense Technology in China. Prof. Qiu, Bin and Rongqing, thank you for offering me the opportunity to coordinate in this project. Yuewen, Baohong, Jianbing, Yun, Shengjian, Rongqing and Qingsong, it's my pleasure to introduce you to come to TU Delft and work with you when you were visiting here.

I also enjoyed a lot on the project collaborated with Nissan research center. Mauro, Kaori and Richard, thank you for hosting me and providing me a comfortable working environment when I was in NRC.

Thanks to my friends in The Netherlands, all of you made my life whole here.

At last, I'd like to take this opportunity to express my profound appreciation to my family.

Mingxin Zhang
Delft, February 2016

Contents

1 Introduction

1.1 Research Context

> "We're presently in the midst of a third intellectual revolution. The first came with Newton: the planets obey physical laws. The second came with Darwin: biology obeys genetic laws. In today's third revolution, we're coming to realize that even minds and **societies** emerge from **interacting** laws that can be regarded as **computations**. Everything is a computation (RUCKER 2006)."

1.1.1 Motivation

Social simulation is a research field that applies computational methods to model, understand or predict a social phenomenon in human society. However, there are still many difficulties for social simulation research as it is often too complex to study social phenomena (SCHUTT 2014). For example, social phenomena can be counterintuitive and unpredictable due to ongoing dynamics of the environment (BONABEAU 2002b).

From the perspective of modeling, the difficulty comes from modeling the complex human behavior in society that depends on the interplay between people, the large number of non-linear interactions, and the dynamic evolving social structure (SQUAZZONI et al. 2013). Since complex social processes of the interactions can not be easily represented as equations (LI et al. 2008), it is hard for social simulationists to use traditional computational and mathematical models to understand how macroscopic social phenomena emerge from the microscopic level (RITZER 2001), for example, the aggregate behavior of groups of people (EASLEY and KLEINBERG 2010) when they interact over a significant period of time (SQUAZZONI et al. 2013).

With the development of artificial intelligence and computational theory, interest in agent-based social simulation (ABSS) has been growing rapidly (GILBERT and TROITZSCH 2005). ABSS was proposed as an alternative to traditional mathematical methods to model the complex behavior and study the social phenomena of human society (ADAMATTI 2014). In an ABSS society, the agents have a one-to-one correspondence with the individuals (or organizations, or other agents) that exists in the real social world, while the interactions between the agents can likewise correspond to the interactions between the real world individuals (GILBERT 2004). Through the interactions, complex behavior of the social system can emerge, which is a new way to study the macroscopic social phenomena in human society. For example, it is more natural to describe how shoppers move in a supermarket than to come up with the equations that govern the dynamics of the density of shoppers (BONABEAU 2002a). The first ABSS model was Schelling's model to study housing segregation patterns (SCHELLING 1971) while the first widely known ABSS model, Sugarscape, was developed in 1996 by EPSTEIN and AXTELL (1996). This model was introduced to simulate and explore the role of social phenomena such as seasonal migration, pollution, sexual reproduction, combat, trade and transmission of disease and culture (CASTELLANO et al. 2009).

Methodologically, AXELROD and TESFATSION (2006) considers ABSS representing an approach that could contribute to two aspects: (1) the rigorous testing, refinement, and extension of existing theories that have proven to be difficult to formulate and evaluate using standard statistical and mathematical tools; and (2) a deeper understanding of fundamental causal mechanisms in multi-agent systems whose study is currently separated by artificial disciplinary boundaries.

GILBERT and TROITZSCH (2005) presented three main objectives of scientific implementation of ABSS: (1) a way to understand basic aspects of a social phenomenon in which the resulting behavior emerges from a system that could be easily observable; (2) a prediction of real life events and phenomena; and (3) research, testing and formulation of hypotheses.

In fact, the insights on ABSS from both Axelrod and Gilbert can be convergent if we shift our focus from the goal and contribution to the category and discipline. ABSS is a cross-disciplinary research and application field. As shown in Figure 1.1, PAUL (2002) defines and differentiates the research areas that are a combination of agent-based computing, computer simulation, and social science as Social Aspects of Agent Systems (SAAS), Multi-Agent Based Simulation (MABS), and Social Simulation (SocSim). Among these, Agent-Based Social Simulation (ABSS) is the overlap between agent-based computing, computer simulation, and social science. Other than ABSS, which is introduced above, SAAS is the overlap between social science and agent-based computing that includes the study of norms, institutions, organizations, co-operation, competition, etc (SQUAZZONI et al. 2013). MABS, the overlap of computer simulation and agent-based computing, mainly includes the model realization, simulation of the agent models, the collection and analysis of outcomes, and simulation optimization. At last, SocSim is the intersection between social sciences and computer simulation. In SocSim, social phenomena in social science are modeled with the methods and tools used in computer simulation.

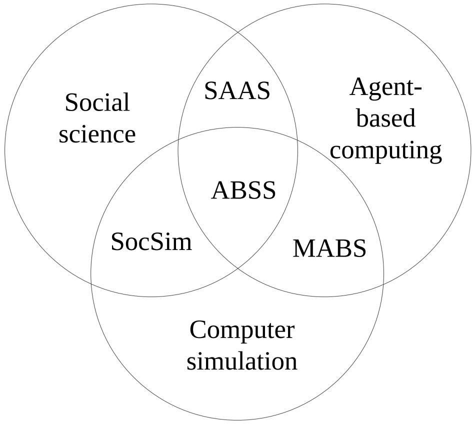


Figure 1.1: ABSS as an overlapping area of other research disciplines (PAUL 2002)

As a matter of fact, some researchers are convinced that ABSS represents an exciting, versatile methodological approach for studying human societies, which could contribute to policy making in social science (AXELROD and TESFATSION 2006), to distributed artificial intelligence and agent technology in computer science (GILBERT and TROITZSCH 2005), and to theory and modeling practice in computer simulation systems (JENNINGS 2000).

Based on this argument, we give a definition of Agent-Based Social Simulation (ABSS) from IZQUIERDO et al. (2003) as follows:

Definition *ABSS is a form of computer **modeling** of complex **social** systems, in which the **agents** within such a system are represented explicitly and individually within the model. The model agents typically represent **human individuals**, but may also represent non-human animals, or human collectives such as firms or states. Their **interactions** with each other, and often with a simulated external **environment**, take place according to **rules** that may vary between agents and change as agents **learn**.*

As we presented in Figure 1.1, ABSS is the overlap of three research areas, computer simulation, agent-based computing and social science. Thus, there are a lot of research works focusing on ABSS from these three communities.

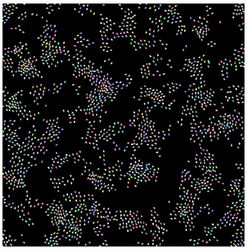
(1) The first is from the computer simulation community who mainly work in discrete event system simulation (ZEIGLER et al. 2000) and parallel and distributed simulation (FUJIMOTO 1999), where the DEVS specification (Discrete Event System Specification) (ZEIGLER and RADA 1984) is getting popular in organizing agent-based social systems (UHRMACHER and GUGLER 2000). In the computer simulation community, DEVS is an operational formalism to describe the system structure and organization of discrete event systems. Simulation implementations in this area are often programmed from scratch in Object Oriented Programming languages (OOP) (e.g. C++, Java, etc.) (LUNA and STEFANSSON 2012, WARD and HUANG 1992), as object-oriented programs show simple techniques for data collection, time management, and statistical instrumentation (TYSZER 2012).

From the standpoint of computer simulation, agents in object oriented agent-based simulation become objects and events become steps activated by loops in the program (GILBERT and TERNA 2000). Thus, many ABSS models have been programmed from scratch due to the simplification of translating the social system into a set of objects and events (RAILSBACK et al. 2006). For example, MACE3J is a Java-based MAS simulation testbed with a supporting library of components, examples and documentation for fast and flexible simulation (GASSER and KAKUGAWA 2002). More examples are given in WANG et al. (2012), LEES et al. (2005), HU et al. (2005), CICIRELLI et al. (2015). These example works may be promising from the perspective of computational efficiency. From a computational point of view, the computational efficiency of the object oriented agent-based simulation is increased with the help of techniques in parallel and distributed simulation (FUJIMOTO 1999).

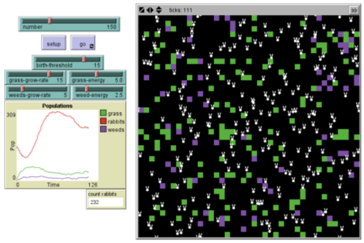
(2) The second is from the agent-based computing community who build agent-based models by an implementation using Agent-Oriented Programming (AOP) languages (SHOHAM 1997, HUNTBACH and RINGWOOD 2003). AOP can be viewed as a specialization of object-oriented programming (SHOHAM 1993), for experimenting with agents with embodied principles and concepts proposed by theorists (WOOLDRIDGE and JENNINGS 1995). For example, 3APL is an AOP language combining imperative and logic programming with a clear and formally defined semantics to model agents (HINDRIKS et al. 1999). As an extension and modification of the original version of 3APL, 2APL (DASTANI 2008) is an AOP language that facilitates the implementation of multi-agent systems with a set of individual BDI-based (Belief-Desire-Intention) agents and a set of environments in which they can perform actions. AgentSpeak can be viewed as an abstraction of one of the BDI systems and allows agent programs to be written and interpreted in a manner similar to that of horn-clause logic programs (RAO 2009). Others are AGENT-0 (SHOHAM 1993), and ConGolog (DE GIACOMO et al. 2000).

There are also models and platforms based on AOP for specific purposes. For example, the Multi-modal Agent Decision Making (MADeM) library enables researchers to model and simulate the social decisions made by each agent based on the agent interpreter Jason (BORDINI et al. 2007), implemented in the AOP language AgentSpeak (RAO 2009). The social decisions are about how to get to work every day, e.g., by train, by car, or by sharing a car (GRIMALDO et al. 2011). Wei proposed a cognitive model for robot control in which the cognitive layer is programmed in the AOP language Goal (WEI and HINDRIKS 2013). As for social simulation purpose, examples of ABSS platforms are GALATEA (DAVILA et al. 2005) and Brahms (SIERHUIS 2001).

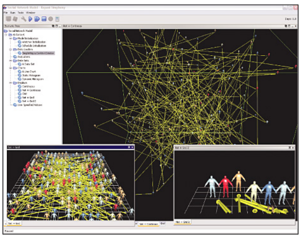
(3) The third is from the social science community who use agent-based models to analyze and study social phenomena (HELBING and BALIETTI 2013), where tools and toolkits are introduced to help social scientists who don't have much programming experience to build their own agent-based models (TAYLOR 2014). These tools and toolkits can be employed as "containers" based upon a specific shell and simplify the task of replicating simulated experiments (GILBERT and TERNA 2000).



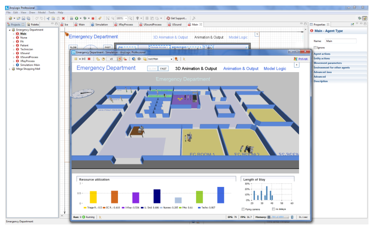
(a) Mason simulation for Craig Reynolds' Boids algorithm[1].



(b) NetLogo Interface of the Rabbits Grass Weeds model[2].



(c) Repast simulation of social influence[3].



(d) Emergency department 3d simulation with Anylogic platform[4].

Figure 1.2: Typical ABSS toolkits

Many modeling and simulation tools and toolkits exist, such as Repast (TATARA et al. 2006) which encapsulates both OOP (JAVA/C++) and AOP (ReLogo), Cormas (PAGE et al. 2000), Ascape (PARKER 2001), MASON (LUKE et al. 2004), NetLogo

[1] Retrieved from `https://cs.gmu.edu/~eclab/projects/mason/`
[2] Retrieved from `https://www.openabm.org/book/33102/42-first-steps-netlogo`
[3] Retrieved from `http://www.lionhrtpub.com/orms/orms-8-06/fragent.html`
[4] Retrieved from `http://www.anylogic.com/screentshots`

Toolkit (WILENSKI 1999), Swarm (MINAR 1996, LINGNAU 1999) and AnyLogic (BORSHCHEV et al. 2002).

There are many common features of these standard tools and toolkits, which are: (1)the basic unit in these systems is a rule-based agent (SMITH and CONREY 2007) and the modeling environment is contained in a specific shell (GILBERT and TERNA 2000), which allows users to create and test agents easily; (2)the model states are usually updated on time ticks/steps, which makes the models easier for time management (GANZHA and JAIN 2012).

With the increasing availability of powerful computing resources (e.g., GPU computing, supercomputing and cloud computing), the research interest of ABSS is shifting to large-scale social systems. Typical large-scale social systems are (1) Evolvement of societies aiming to gain a greater understanding of how human societies evolve through testing of different hypotheses and theories for urban change (CROOKS 2006); (2) Artificial society simulation usually spanning a shorter time frame compared to society evolvement systems and performs as a testbed for other high-level domain-specific research, such as epidemics (BISSET and MARATHE 2015) and rumor spread (GONG and XIAO 2007); (3) Evacuation simulation which can be used as a verification of the existing emergency plans (CAMILLEN et al. 2009) or as a prediction of the consequences of certain courses of actions to a large-scale emergency (HAWE et al. 2012); (4) Transportation and Mobility simulation, used for general traffic analysis (NAGEL and RICKERT 2001); and (5) Systems for analyzing other social phenomena such as migratory flows (FILHO et al. 2013).

Among these systems, epidemic spread on a large scale (city/country/global) has remained a research focus for an extended period. Over the last two centuries, science made enormous progress in the fight against infectious diseases, but the biggest battles may be still ahead of us (WHO 2007). With the remaining threats, much research was conducted due to the following reasons [5]:

- With the increase in plane traffic, contagious illnesses spread farther and faster than ever. During the previous decade, H1N1 Flu (STROUD and VALLE 2007) and Severe Acute Respiratory Syndrome (SARS) (HUANG 2010) infected people around the globe.
- Some diseases, such as tuberculosis (DE ESPÍNDOLA et al. 2011) are now becoming resistant to antibiotics.
- Old enemies like malaria (LINARD et al. 2009) refuse to go away.
- Others like smallpox (GRUNE-YANOFF 2010) that have been eradicated threaten a devastating comeback if released.
- New diseases are emerging at an unprecedented rate of one per year, such as MERS (Middle East Respiratory Syndrome), SARS, Ebola and new strains of Influenza.

[5]Global Public Health Threats. Retrieved from http://www.greenfacts.org/en/global-public-health-threats/

Therefore, we will focus on epidemic prediction and control as the main application area of ABSS in this dissertation.

1.1.2 Problem Statement

To get a better understanding of epidemics and to effectively support policy making during a disease outbreak, several large-scale ABSS models were developed aiming to create powerful tools that can help studying different diseases, different interventions, and different simulation scenarios (STROUD and VALLE 2007, PARKER and EPSTEIN 2011, AJELLI et al. 2010, RAKOWSKI et al. 2010, BISSET et al. 2009a, BISSET et al. 2014, GE et al. 2013, BARRETT et al. 2008). Among these models, we noticed a common characteristic that these models were all developed from scratch, and none of them were developed based on existing ABSS platforms or tools.

For existing models and platforms in the agent-based computing community, the difficulty in studying large-scale epidemic systems can be explained by the computational complexities introduced by the implemented complex agent behaviors, such as planning and reasoning, limiting the number of agents in a simulation study. To decrease the complexity, agents are often implemented as thread-based objects in languages such as Brahms (SIERHUIS 2001). However, the synchronization among large number of threads could greatly decrease system performance and limit system scale.

For tools and toolkits in the social science community, large-scale ABSS platforms have been developed such as RePast HPC (ZIA et al. 2013, ZIA et al. 2012) with the help of high-performance computing middleware. However, there are several limitations for these platforms to study large-scale epidemic prediction and control. Firstly, the model concepts in these tools and toolkits are usually encapsulated and simplified to ease model development, for example, most large-scale ABSS models consider the agent environment as discrete cells/grids to help the model partitioning and reduce unnecessary inter-node communication messages. Examples include ABCCA (SUZUMURA et al. 2014) and HLA-GRID-REPAST (THEODOROPOULOS and ZHANG 2006). In simple cases, this may be a sensible approach when the movement of the agents has no particular pattern or is very restricted (WANG et al. 2012). However, in current agent-based epidemiology research, direct physical contact (e.g. touching) or vector-borne contact (e.g. a droplet) outside closed rooms is also considered to be an effective method for the spread of disease (MORSE 1995), especially in densely populated areas such as public transportation (HALL 2007). Thus, a detailed and refined representation of entities in a continuous environment is necessary, rather than cell/grid based and discrete. Secondly, agents are rule-based due to the encapsulation of agency concepts (e.g. knowledge) in these platforms (AN 2008) to ease the modeling development phase, which is insufficient to model the complex behavior of agents during an epidemic emergency. Thirdly, the tick/time step based mechanism is a popular way to advance simulation time in tools like Repast (TATARA et al. 2006) and

NetLogo (WILENSKI 1999). This design brings difficulties for these platforms to run large-scale epidemic models because of the choice of the time step (CAZELLES et al. 2007). Small steps in large-scale epidemic models with uneven event times greatly decrease system performance while large steps speed up the simulation, but may lead to wrong results (HELBING and BALIETTI 2013).

For existing platforms in the computer simulation community, object oriented methodologies have inherent advantages to deal with large-scale systems, especially because parallel and distributed simulation has been widely studied (FUJIMOTO 1999). However, the concepts in agent-based systems are greatly simplified. For example, the social networks implemented in these platforms are usually simplified and fixed, as dynamic social interactions significantly increase the number of synchronization messages if the model runs in a parallel and distributed manner.

Thus, although mathematicians, epidemiologists, computer and social scientists share a common interest in studying the spread of epidemics and rely on very similar models for the description of the diffusion of pathogens (PASTOR-SATORRAS et al. 2014), current large-scale ABSS models for epidemic predictions and control are usually developed from scratch in OOP languages. In addition, there are some limitations in current large-scale ABSS models for epidemic predictions and control.

Firstly, although system performance is guaranteed (ZEIGLER et al. 1997) by adopting distributed/parallel mechanisms, the precondition of adopting distributed/parallel mechanisms in large-scale ABSS models is usually to simplify the model components including the agent itself (UHRMACHER et al. 1997, CICIRELLI et al. 2015). Thus, the agents are usually simple reactive agents. In other words, agents for epidemic predictions are rather simple in terms of both the agent architecture and the decision-making mechanisms (HAWE et al. 2012). Agents in these models either make decisions based on simple rules or behave to initially fixed schedules. However, the results of epidemic prediction are highly related to the decisions made by individuals (FENICHEL et al. 2011). Specifically, the variance and diversity of agents' decisions can significantly impact the effect of proposed interventions during an epidemic outbreak.

Secondly, social aspects of the agent environment, such as norms and institutions that provide rules and sanctions for agents to behave during a disease outbreak (SAVARIMUTHU et al. 2008), are excluded in current large-scale ABSS models while they play an important role in disease spread as they influence agents' behavior institutionally.

At last, modeling complex social interactions on a large scale remains a challenging task as well. Current large-scale ABSS models have omitted or simplified the interaction part of social contacts in real life, for example, negotiating in planning joint social activities. Agents in existing models will only execute their scheduled activities and interact with others based on random or predefined social networks, which cannot describe the dynamic evolvement of social relations during a disease outbreak (GUO et al. 2015). However, dynamic social interactions provide a perfect fabric for fast disease propagation while they can be dramatically altered when people respond to the crisis and interventions (BISSET et al. 2009b).

To sum it up, the architectural foundation for large-scale agent-based epidemic prediction and control is still insufficient (SARJOUGHIAN et al. 2001), although there are discussions about the use of standard patterns for agent-based modeling (NORTH and MACAL 2011) based on their equivalents. We see these insufficiencies as the core problem for our research:

> *There is a gap between current large-scale ABSS models and the requirements to study specific large-scale social systems, for example, epidemic prediction and control.*

1.2 Research Objective and Questions

Research Objective *To design, implement, and test an effective conceptual model for large-scale agent-based social simulation.*

A conceptual model can be defined as a non-software specific description of the computer simulation model (that will be, is or has been developed), describing the objectives, inputs, outputs, content, assumptions and simplifications of the model (ROBINSON 2008). A conceptual model is an abstraction of a simulation model from the part of the real world it is representing (the real system) (ROBINSON 2010).

To satisfy the research objective, we will answer several research questions in different steps.

Research Question 1 *What conceptual model and model concepts can support large-scale agent-based social simulation?*

Research Question 2 *How can the components of this conceptual model be implemented in the case of epidemic prediction and control?*

Research Question 3 *How can the case of epidemic prediction and control benefit from the proposed conceptual model regarding of model outcomes and system performance?*

Research Question 4 *How can large-scale agent-based social simulation benefit from the case of epidemic prediction and control in this research?*

1.3 Research Methodology

To describe the choices that have to be made for formulating an effective research methodology, SAUNDERS et al. (2012) developed the research onion which is presented in Figure 1.3. The research onion provides an effective sequence through which a research methodology can be designed. Its usefulness lies in its adaptability for almost any type of research, and it can be used in a variety of contexts (BRYMAN 2012). This section will describe the concepts and explain the choices for the different layers.

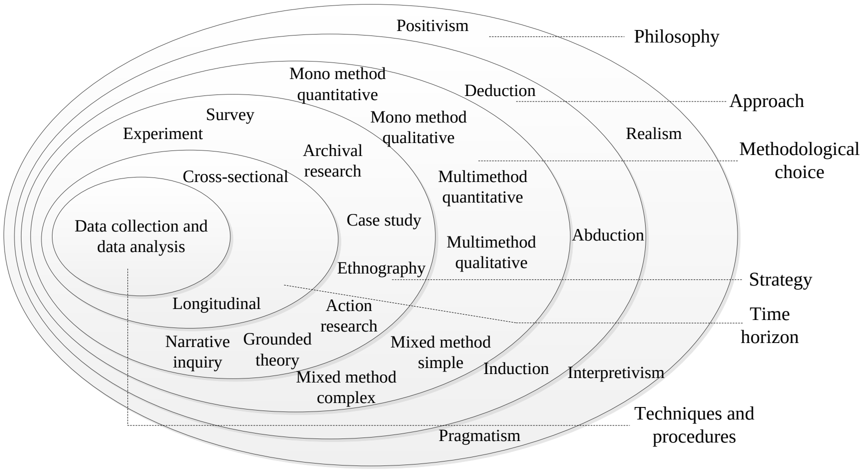


Figure 1.3: The research 'onion' (SAUNDERS et al. 2012)

1.3.1 Research Philosophy

To choose an appropriate research strategy and approach for conducting a successful research project, researchers should pay attention to research philosophy as a starting point. In general, research philosophy is built on three major components which are ontology, epistemology, and methodology. Based on these, various paradigms in different overlapping categories can be differentiated (MKANSI and ACHEAMPONG 2012). For example, ORLIKOWSKI and BAROUDI (1991) and GUBA and LINCOLN (1994) divide research philosophy into positivism, post-positivism, critical theory, and constructivism, while SAUNDERS et al. (2012) differentiates between positivism, realism, interpretivism, and pragmatism in the research onion.

Positivism, as a philosophy in both the social and natural sciences, generally confines a system to the data of experience and excludes a priori or metaphysical speculations. This view assumes that society operates according to laws like the physical world.

Realism is the viewpoint that some aspects of our reality are ontologically independent of our conceptual schemes, perceptions, linguistic practices, beliefs, etc. (DELL'AVERSANA 2013). Realists tend to believe that whatever we believe now is only an approximation of reality and that every new observation brings us closer to understanding reality (RANA 2008).

Interpretivism (also known as antipositivism or negativism) is the belief within social science that the social realm may not be subject to the same methods of investigation as the natural world [6]. Interpretive researchers assume that access to reality (given or socially constructed) can only be obtained through social constructs such as language, consciousness, shared meanings, and instruments (MYERS 2013).

Pragmatism (or Pragmaticism) is the view that considers practical consequences or real effects to be vital components of both meaning and truth (HEVNER 2007). Pragmatists believe that truth is not "ready-made", but that truth is made jointly by us and reality [7]. Some pragmatists also believe that that truth is mutable (beliefs can pass from being true to being untrue and back again), and that truth is relative to a conceptual scheme (SWANSON 2010).

Among these paradigms, positivism and interpretivism are the two dominant ontological and epistemological traditions (BECKER and NIEHAVES 2007). Although they are often seen as opposite and have many basic differences (DECROP 2006), they share the assumption that a 'real world' exists beyond the realms of human cognition (WEBER 2004), and they are frequently used in conjunction with contemporary research (CROSSAN 2003).

This research focuses on designing a conceptual model to support large-scale agent-based social simulation. A positivist view is necessary for designing the proposed conceptual model. However, the proposed conceptual model often involves subjective understanding about the requirements and supportive hypothesis, and there is no available data for validating part of the results of the simulation studies. Thus, an interpretivistic perspective is also required to evaluate the model and the model results.

1.3.2 Research Approach, Methodological Choice, Time Horizon, Strategy, and Instrument

The research approach is an overall paradigm based on the research philosophy for conceptualizing and conducting an inquiry and constructing scientific knowledge (CECEZ-KECMANOVIC 2001). There are three types of research approaches in the research onion: deduction, induction and abduction. Since there is not enough data and observations, this research will use the deductive approach to propose a conceptual model for large-scale agent-based social simulation, and use the inductive approach through testing the proposed model in the case of epidemic prediction and control.

[6] Wikipedia, s.v. "Antipositivism", last modified on 18 December 2015, http://en.wikipedia.org/w/index.php?title=Antipositivism&oldid=695740053.

[7] Retrieved from http://www.philosophybasics.com/branch_pragmatism.html, visited on 07 January 2016.

Besides the research approach, there are three recognized research methodological choices: the quantitative method, the qualitative method and the mixed method. This research will use the mixed research method, where the qualitative method is used to develop the proposed conceptual model and to implement a reference model for the case of epidemic prediction control, and the quantitative method is used to measure the model outcomes by statistical analysis of data collected through simulation studies.

Regarding the choice of time horizon in the research onion, a longitudinal method is selected in this research as both the input data and validation data are gathered over a period of time.

The research strategy will use a design science research methodology (DSRM) in VAISHNAVI et al. (2007), PEFFERS et al. (2007), which was created with objective of providing researchers with a mental model or template for a structure for research outputs. The reason for the choice is that this research is primarily driven by the needs of accounting practice and focusing on the creation of new artifacts (GEERTS 2011).

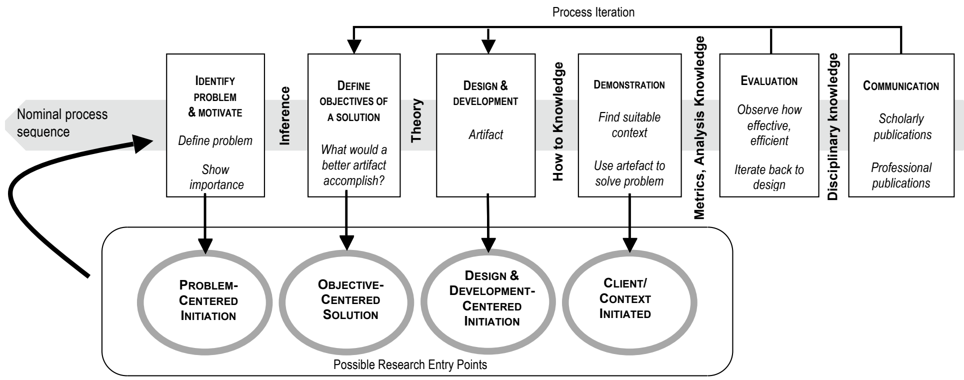


Figure 1.4: Design Science Research Methodology (DSRM) Process Model (PEFFERS et al. 2007)

There are six major phases in the DSRM process as presented in Figure 1.4. Along with the execution of these phases, different research instruments are used.

The phases start with the problem awareness phase that tries to find an interesting problem that could yield new developments in the research field (MORTON and REDMOND 2015).

For the first phase of problem awareness, this research involves a systematic *literature review* as the research instrument, where patterns of existing research on large-scale ABSS will be examined to establish the knowledge on the particular study of epidemics. The existing research on large-scale ABSS on epidemic prediction and control will also be examined.

Phase two looks at the objectives of potential artifacts that might address the problem of the first phase. This phase should be able to provide requirements to produce an artifact that can be implemented in the next phase (ibid.).

Phase three is the development of the artifact, which will be evaluated according to the functional specifications. Phase four is the application of the artifact in a proper context to solve problems proposed in phase one.

Phase five evaluates the results from the application phase, and further suggestions could be made through iteration if the results don't satisfy the requirements proposed in phase two.

For these phases, we will use *literature review, data analysis, computer simulation* and *experiments* as research instruments. A Literature review will be mainly conducted in the domains of ABSS, epidemic modeling, agent-based modeling, social networks, and other related works. Data analysis will be performed (1) on the data used as the input for a case study to identify data consistency, (2) on the data of the model output to verify model results, and (3) on the data from an actual survey about urban statistics by other independent research to validate model outputs.

Computer simulation and experiments were conducted with the implemented simulation model to evaluate the design. Computer simulation is growing in popularity as a methodological approach allowing to assume the inherent complexity of organizational systems as a given (DOOLEY and DOOLEY 2001). Experimental research is the investigation of relations between controlled variables, with tightly controlled variations, solving an artificial problem situation (epidemic prediction and control). The purpose of computer simulation and experiments is understanding the behavior of the system (SHANNON 1998) and to evaluate various strategies for controlling disease outbreak.

Phase six is the conclusion. This can indicate the end of the research cycle and completion of the DSR project by summarizing the findings in research publications.

1.4 Thesis Outline

Agent-based social simulation is an interdisciplinary area, where knowledge, theory and concepts in social science about social structures, mechanisms, and processes of interaction and communication are combined with modeling and simulation methodologies in agent-based modeling and simulation, with the objective to solve problems in social science.

This thesis will present a novel large-scale agent-based social simulation conceptual model and a reference implementation for the key model components. With the proposed reference implementations, a case study on epidemic prediction and control is conducted, and the simulation results are analyzed and discussed.

The remainder of this thesis is organized as follows. Chapter 2 presents the relevant background theories and concepts that are related to ABSS models. Chapter

3 introduces the case of large-scale epidemic prediction and control and presents the challenges and shortcomings of current solutions. Chapter 4 presents a novel conceptual model for large-scale ABSS and proposes a reference implementations for key model components. Chapter 5 describes how this conceptual model is used in the case of disease spread and policy management for the city of Beijing and what the simulation results are if relevant simulation scenarios are executed in this case study. Chapter 6 shows how this conceptual model performs when multiple simulation scenarios are executed regarding of model outcomes and system performance. Chapter 7 discusses the possibilities of using this conceptual model for other large-scale ABSS systems. Chapter 8 draws conclusions and provides suggestions for future work.

Figure 1.5 gives an overview of the structure of this thesis.

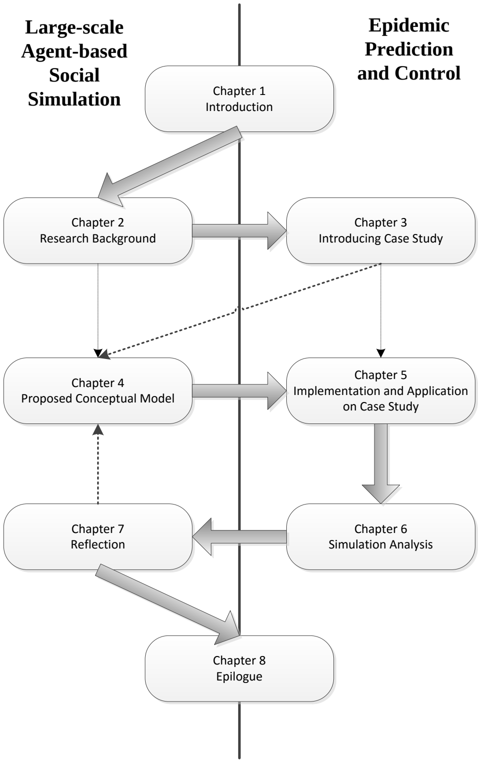


Figure 1.5: The research process

2 Background and Related Work

2.1 Application Areas in ABSS

As we presented in the first chapter, agent-based social simulation (ABSS) is a particular form of Multi-Agent Based Simulation (MABS), which is the overlay between computer simulation and agent-based computing.

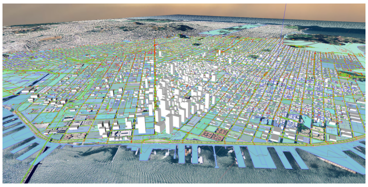
Multi Agent Based Simulation, also referred to Agent-based modeling and simulation (ABMS) (ZHENG et al. 2013), is a relatively new process of designing an agent based model composed of interacting, autonomous agents (SIEBERS et al. 2007, SIEBERS et al. 2010), and conducting experiments with this model for the purpose of understanding the behavior of the system and/or evaluating various strategies for the operation of the system (SHANNON 1998). In ABMS, individual agents and their behavior are described by an agent architecture, and agents interact with others or the environment following a set of rules. By this modeling and simulation process, the full effects of the diversity of attributes and behavior can be observed as it generates the behavior of the system model as a whole (MACAL and NORTH 2010). Self-organization can be observed in such models as well. Patterns, structures, and behavior emerge that were not explicitly programmed into the models, but arise through the agent interactions (ibid.).

Compared to agent-base social simulation, ABMS spans a broader range of areas. There are many review papers summarizing the application areas applying ABMS models (MACAL and NORTH 2010, RAILSBACK et al. 2006, CASTIGLIONE 2006, HELBING 2012, BONABEAU 2002a, GETCHELL 2008). Based on their research, we list several areas where ABMS models are typically adopted.

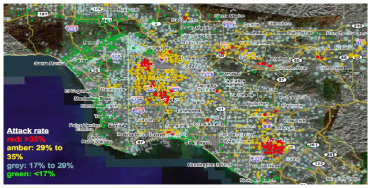
- Economics. Examples are stock markets (ZAWADOWSKI et al. 2002), self-organizing markets and trade networks (TESFATSION 2001), consumer behavior (SAID et al. 2002), and deregulated electric power markets (WEIDLICH and VEIT 2008). Among these, E-Commerce is a hot application topic (ROMAN and KATERINA 2012) where investigating market mechanisms and consumer behavior on-line are attractive for researchers, and ABMS can be adopted to model the market dynamics.
- Education and Training. Education in the 21st century can benefit from using ABMS as there is a geographical dispersion of students and teachers who spend hours a day interacting with multimedia environments (LI et al. 2008). ABMS can be adopted to decide resource allocation for education optimally (PAN et al. 2006).
- Ecology. Examples are population dynamics of fish such as salmon and trout (GRONER et al. 2013), flocking behavior in birds (RAZAVI et al. 2010), rain forest growth (DEADMAN et al. 2004), fire spread (HU et al. 2005), and insect societies (PRATT et al. 2005).
- Political Sciences. Examples are water rights in developing countries (SCHREINEMACHERS and BERGER 2011), party competition (LAVER and SERGENTI 2011), origins and patterns of civil violence (EPSTEIN 2002).
- Social Science. Examples are evolvement of societies (CROOKS 2006), urban riots (BRUZZONE et al. 2011), spread of epidemics (BISSET and MARATHE 2015), human evacuation (CAMILLEN et al. 2009, BRUZZONE et al. 2014), and traffic flow management (NAGEL and RICKERT 2001).

Among those areas listed above, the social science area is the research interest of this thesis. As mentioned in the list, social science includes a lot of sub-areas that attract research interests. We listed several example applications that can easily be elaborated as follows:

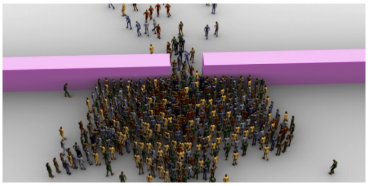
1) Evolvement of societies. Agent based models (ABM) allow for the testing of different hypotheses and theories for urban change, thus leading to a greater understanding of how cities evolve (CROOKS 2006). For example, Crooks presents an ABSS model to explore how cities change and develop by integrating his model with Geographical Information Systems (ibid.). UrbanSim is a urban simulation models studying urban dynamics and reflecting specific agents interacting with other agents, such as households, jobs, and governments (WADDELL 2012). To investigate changes in the social structure of New Zealand, the Modeling Social Change (MoSC) project simulated inter-ethnic cohabitation patterns using ABSS models populated with unit-level census data (WALKER 2009). Similar research can be found in another agent-based model for urban simulation (NAVARRO et al.



(a) ABM on Evolvement of societies[1]



(b) ABM on Artificial society with epidemic[2]



(c) ABM on Evacuation[3]

(d) ABM on Transportation and Mobility[4]

Figure 2.1: Application Areas of ABSS

2011) and simulation of residential dynamics in a city (BHADURI et al. 2014).

2) Artificial society. Different from society evolvement research, artificial society research usually spans a shorter time frame compared to society evolvement and performs as a fundamental testbed for other high-level domain-specific research, such as epidemics (BISSET and MARATHE 2015) and rumor spread (GONG and XIAO 2007). An artificial society is a multi-agent simulation where autonomous agents carry out activities in parallel, move around the environment locations and communicate with each other (SAWYER 2003). It requires individual agents representing humans that have daily behavior, together with locations (households, schools, work places, hospitals, stations, etc.) that provide space for agents' activities. Based on the artificial city models, fundamental collective behaviors "emerge" from the interaction of individual agents following a few simple rules (EPSTEIN and AXTELL 1996).

3) Evacuation. ABSS models on evacuation can be used as verification of the existing emergency plans for building evacuation (CAMILLEN et al. 2009), or to predict the consequences of certain courses of action and to respond optimally to a large-scale emergency (HAWE et al. 2012), or for analysis of realistic evacuation models at the level of large cities (ZIA et al. 2012). In the study of large-scale evacuations of cities by LÄMMEL et al. (2010), ABSS is used to understand the interdependency of infrastructure systems and their vulnerabilities for natural disasters,

[1] Retrieved from `https://www.cs.purdue.edu/cgvlab/urban/index.html`
[2] Retrieved from `http://www.lanl.gov/programs/nisac/episims.shtml`
[3] Retrieved from `http://www.geosimulation.org/disasters.html`
[4] Retrieved from `https://github.com/agents4its/mobilitytestbed/wiki`

terrorist attacks, accidents, and other incidents.

4) Transportation and Mobility. A lot of ABSS studies on large-scale transportation are conducted. For example, the highly scalable X10-based agent simulation platform XAXIS was used to implement a large-scale traffic simulation which focuses on performance issue (SUZUMURA and KANEZASHI 2012). The TRANSIMS project (NAGEL and RICKERT 2001) by the Los Alamos National Laboratory (LANL), has helped to create large-scale ABSS models to study the travel behavior of the 7.5 million inhabitants of Switzerland (RANEY et al. 2002) and for general traffic analysis in the Buffalo-Niagara metropolitan area (ZHAO and SADEK 2012). A similar research project was the large-scale multi agent-based transport simulation for Shanghai (ZHANG et al. 2013) using MATSIM. Agent-based traffic simulation was also used to capture the regional impacts of new development (ZHANG et al. 2012). Besides transportation, ABSS can be easily applied to pedestrian mobility models (ZIA et al. 2013, PELECHANO et al. 2007).

5) Other social phenomena. Besides the above hot topics, agent-based social simulation are widely used in analyzing other social phenomena. For example, educational scientists developed agent-based models to identify the casual implications of the same-race effect on the educational achievement trends (MONTES 2012). Social scientists used ABSS models to reproduce migratory phenomena to gain a deeper understanding regarding migratory flows and social networks (FILHO et al. 2013).

2.2 Conceptual Model for ABSS

A conceptual model is usually made of a composition of concepts. The general opinion about when a model should be called agent-based is that this is often decided at the conceptual and not at the implementation level (SIEBERS et al. 2010). Due to the standardization of ABM tools and toolkits, conceptual frameworks are commonly designed in platforms in the social science community which are discussed in Section 1.1.1. Among these platforms, three typical platforms are selected to analyze their conceptual models, which are Swarm (MINAR 1996, LINGNAU 1999), NetLogo (WILENSKI 1999) and Repast (TATARA et al. 2006).

Swarm was developed at the Santa Fe Institute in 1994 and was specifically designed for artificial life applications and studies of complexity. It is selected as an representative in this dissertation for analysis as it is the first re-usable software tool created for agent-based modeling and simulation (ALLAN 2009).

The two types of concepts that are particular to each Swarm simulation (STEFANSSON 1997) are Agents and Model Abstractions. Agents are actors in the artificial world or possibly auxiliary objects that control the agent behavior. Model Abstractions are objects that may collect information and respond to inquiries from the agents or perform other necessary tasks in the artificial world.

Swarm is the most mature ABMS library based framework and is stable and well organized (ALLAN 2009). However, Swarm is not suitable for large-scale social systems. One of reasons is the way it use to organize a group of agents. To manage more than one agent in a Swarm world, 'List' is used to keep track of all agents as a collection. When the list grows to millions, it will cause problems for agent interactions. There is a lack of concept in the Swarm world to represent the relationships among agents.

NetLogo is a multi-agent programming and modeling environment for simulating complex phenomena. It is selected for analysis as it is one of the most widely used multi-agent modeling tools today (RAILSBACK et al. 2006), with a community of thousands of users worldwide (BLIKSTEIN et al. 2005).

In NetLogo, the only model concept is agent. There are four types of agents in NetLogo, and each serves a different purpose in a NetLogo model (LYTINEN and RAILSBACK 2012):

- Observer. In each NetLogo simulation, there is always exactly one instance of this kind of agent. It is the only agent that can perform certain global operations in a model (e.g. clear-all).
- Patch. These stationary agents represent the agents' physical environment where each agent can only stay in one patch (grid/cell).
- Turtle. These are agents equivalent to agents in other ABM platforms.
- Link. Links, representing the relationships among agents, are agents connecting one turtle to another.

Similar to Swarm, NetLogo is difficult to apply in large-scale social systems, as well. One of the limitations is the simplification of the concept on modeling environment. NetLogo adopts grids (patches) to represent agent environment, while many large-scale systems social include continuous space.

Repast (Simphony) is an open source agent-based modeling toolkit that simplifies model creation and use. RePast is selected as it is ranked to be the first among several other ABMS tools based on the value of the weighted total score by TOBIAS and HOFMANN (2004).

In Repast, the most important concepts besides agents are contexts and projections (HOLZHAUER 2010). Contexts represent agent's environments with their own internal states which are organized hierarchically. Each agent needs to belong to at least one context. Projections specify the environment the agents are within and impose a structure on the context. This might be continuous space, grid, GIS information or a network. Figure 2.2 outlines the relations between contexts, agents and projections (ibid.).

Repast (Simphony) probably now has the greatest functionality of any AMBS package (ALLAN 2009). However, the concepts of context and projection in Repast are designed for general agent-based modeling and simulation. There is no specific concept available for modeling concepts such as social regulation in large-scale social systems.

Swarm, NetLogo and Repast are three representatives among many other platforms, and there are many review papers on all kinds of platforms (MACAL and

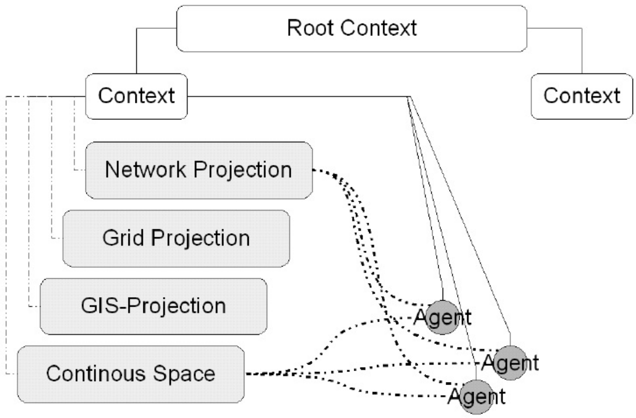


Figure 2.2: The relations between contexts, agents and projections in Repast

NORTH 2010, RAILSBACK et al. 2006, CASTIGLIONE 2006, HELBING 2012, BONABEAU 2002a, GETCHELL 2008). A review paper by RAILSBACK et al. (2006) concluded that "it would be easier to teach these platforms (and even to motivate students to bother with them) if their libraries were more clearly linked to a well-defined, standard conceptual framework". This conceptual framework can provide a common language for thinking about and describing agent-based models. MACAL and NORTH (2010) concluded that a typical agent-based conceptual model usually includes three concepts :

- Agent. Each agent has attributes and behavior and individually assesses its situation and makes decisions on the basis of a set of rules or a defined schedule.
- Interdependency. An interdependency is an underlying topology of connectedness defining how and with whom agents interact.
- Environment. The environment is the sum of entities that agents can interact with in addition to interacting with other agents.

These concepts in conceptual models used in general ABM tools and toolkits are rather popular in many application areas. Nevertheless, as we showed in Chapter 1 (pp.7), they are not suitable to simulate large-scale social systems. As far as we can see, one of the design choices preventing the conceptual models to be applied in large-scale social systems is the integration of many heterogeneous concepts into general ones. Regarding agent environment, few of these models distinguish social environment (e.g., norms) from physical environment (e.g., geographical space). These concepts and conceptual models work flexibly for diverse small-scale problems. When they are applied to large-scale social systems, difficulties arise, such as organizing the large number of agents with dynamic evoling social networks.

Therefore, we propose the following requirement for research on large-scale ABSS in this dissertation as follows:

Research Requirement - Model Architecture A conceptual model is required for large-scale ABSS.

2.3 Agent and Agent Architecture in ABSS

Over the last few years, the term 'agent' has become almost commonplace, far beyond its originating niche area of interest in artificial intelligence (LUCK and D'INVERNO 2001). Although increasingly popular, the term has been used in such diverse ways that it has become meaningless without reference to a particular notion of agenthood (SHOHAM 1993). There is a lack of agreement over what actually constitutes an agent. "Agents sound just like computer programs. How are they different (FRANKLIN and GRAESSER 1997)?"

There is no universal agreement in literature on a precise definition of an agent (MACAL and NORTH 2010), therefore many researchers provided their own definitions.

FRANKLIN and GRAESSER (1997) defines "an autonomous agent is a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future".

FERBER (1999) states "an agent can be a physical or virtual entity that can act, perceive its environment (in a partial way) and communicate with others, is autonomous and has skills to achieve its goals and tendencies".

JENNINGS and WOOLDRIDGE (2012) views an agent as "a computer system situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives".

BONABEAU (2002a) sees an agent as "an autonomous decision-making entity that assesses its situation and makes decisions on the basis of a set of rules". In addition, agents may be capable of evolving, allowing unanticipated behaviors to emerge.

Among these definitions, the most frequently mentioned keywords for defining an agent are "autonomous", "environment", "sense (perceive)", "act" and "goals (objectives)". Based on the systems we are targeting, we define an agent in the context of large-scale social systems as:

Definition *An agent is an* ***autonomous*** *entity that* ***senses*** *and* ***acts*** *upon its* ***environment*** *including other agents, physical entities, and social regulations, and directs its activity towards achieving* ***goals****.*

This definition is not concerned with how complex and intelligent agents can be. In fact, agents can be categorized based on their degree of intelligence and capability. RUSSELL and NORVIG (2003) group agents into five classes: simple reflex agents, model-based reflex agents, goal-based agents, utility-based agents and learning agents. NWANA (1996) identify seven types of agents: collaborative agents, interface agents, mobile agents, information/Internet agents, reactive agents, hybrid agents and smart agents. In both of the categorizations, agents are different in terms of application area, capability and "smartness" (CHEN 2008), which is related to the design of agent architectures.

Agent architectures can be thought of as designing agents from the perspective of software engineering. Thus, researchers in this area are primarily concerned with the problem of designing agents that will satisfy the properties specified by agent theorists (WOOLDRIDGE and JENNINGS 1995).

Over the years, a large amount of research was done to design agent architectures for various purposes. For example, mobile agents can be developed based on templates such as Aglets (LANGE and MITSURU 1998), and Web-based agents based on JAVA Agent Template(JAT). Originating from Aglets, LI et al. (2006) proposed an execution model of mobile agents called SMA which is described via SMA-DEVS. They also built a direct execution simulation environment called MADESE for performance evaluation of mobile agents.

Agents with more "smartness" can be realized through cognitive architectures such as SOAR (NEWELL et al. 1987) and ACT-R (ANDERSON et al. 2004), which provide more complete representations and reasoning frameworks than the above template architectures (BUSEMEYER and DIEDERICH 2010). For example, ACT-R supports a theory of human cognition and provides a wide range of cognitive capabilities.

Besides cognitive architectures, there are a number of multi-level/layer architectures, such as TouringMachines (FERGUSON 1992), Atlantis, and InteRRap (MÜLLER and PISCHEL 1993), providing a more flexible mechanism to solve cognitive problems. For example, SLOMAN et al. (1999) introduced a 3-layer internal architecture of individual agents required for social interaction, collaborative behavior, complex decision making, learning, and emergent phenomena within complex agents, and implemented a software toolkit, SIM_AGENT, to allow construction sets of agents in which each agent has a multi-component architecture.

Compared with the above mentioned agent architectures, Beliefs-Desires-Intentions (BDI) (BRATMAN 1999) based models provide a more formal description for reasoning agents. In BDI theory, beliefs represent the informational state of the decision-maker. More precisely, Beliefs represent what the agent believes about itself and the environment. Desires represent the motivational state of agent, that is, the goals that the agent would like to reach. Intention is constructed by a set of plans, where the agent chooses the best action to perform based on its beliefs.

BDI has been extensively explored in multi-agent reasoning research. For example, DUNIN-KEPLICZ and VERBRUGGE (2002) investigated the notion of collective intention in teams of agents involved in cooperative problem solving (CPS) in multi-agent systems. GRANT et al. (2005) presented a formal logical calculus for representing the formation of intentions by agents, which can describe the reasoning and activities of the agents. ZUCKERMAN et al. (2012) presented a formal Beliefs-Desires-Intentions (BDI) based model, called the Adversarial Activity model, for bounded rational agents operating in a zero-sum environment.

Based on the these formal BDI models, researchers contributed to architectural work and implementations. BRATMAN (1999) presented a high-level specification of the practical-reasoning component of an architecture for a resource-bounded rational agent. TAMBE (1997) implemented an agent architecture STEAM, founded on the joint intentions theory and a practical operationalization. However, the most successful architecture is PRS (Procedural Reasoning System) (INGRAND et al. 1992) which is an implementation based on the BDI theory. There are four main components in the PRS architecture: (1) beliefs; (2) desires; (3) intention; and (4) plans.

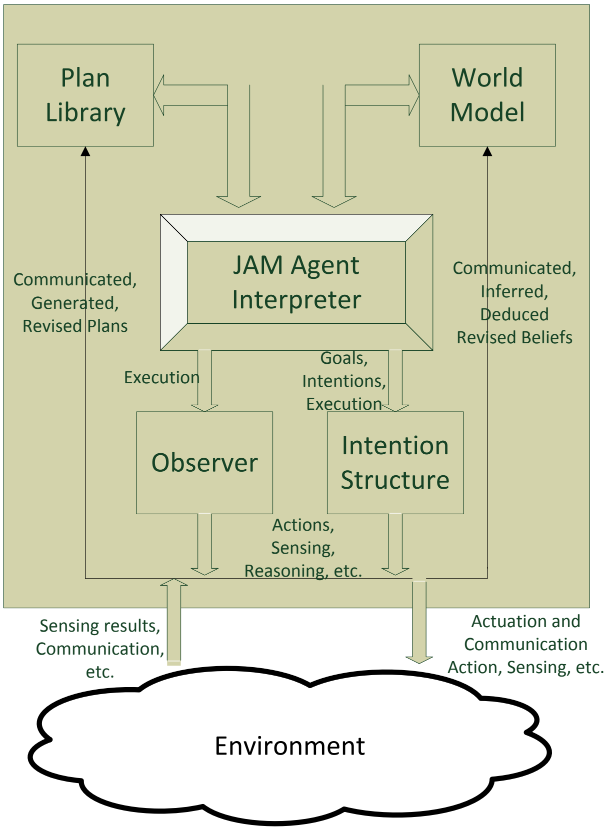


Figure 2.3: The JAM intelligent agent architecture

JAM (HUBER 1999) is a hybrid intelligent agent architecture that draws upon the theories and ideas from PRS, and some other agent architectures, presented in Figure 2.3 as an example. ARTS, a BDI architecture that was extended from PRS and JAM includes goals and plans which have deadlines and priorities, and allows the development of agents that guarantee (soft) real-time performance (VIKHOREV et al. 2009).

To deal with norm emergence, normative agent architectures were introduced and designed to enable norms to be communicated, adopted and used as meta-goals on an agent's own processes (CASTELFRANCHI et al. 2000). Typical examples are EMIL-A by ALDEWERELD (2007) and NoA by KOLLINGBAUM and NORMAN (2003).

Before choosing a proper reference architecture for agents in large-scale ABSS, the requirements need to be finalized. MACAL and NORTH (2010) consider agents to have certain essential characteristics and some optional characteristics. Since one of the problems we try to solve in large-scale social system is to improve the agent capability, we introduced all these characteristics for agents in the large-scale systems to some extent as below (ibid.):

CH1 "An agent is a self-contained, modular, and uniquely identifiable individual."

CH2 "An agent is autonomous and self-directed."

CH3 "An agent has a state that varies over time."

CH4 "An agent has dynamic interactions with other agents that influence its behavior."

CH5 "An agent may be adaptive, for example, by having rules or more abstract mechanisms that modify its behavior."

CH6 "An agent may be goal-directed, having goals to achieve (not necessarily objectives to maximize) with respect to its behavior."

CH7 "An agent is heterogeneous."

As described at the beginning of the thesis, the case system is composed of a large number of parallel autonomous agents which have a certain level of capability to reason, interact, and perform specified tasks. To model this specified type of agent, a desired agent architecture must satisfy the following requirements, which are mostly interpretations of the characteristics mentioned above. In addition to these requirement, there is also the challenge of computational complexity. Large-scale social simulations could involve a large number of agents, up to thousands or even millions. Thus, scalability is a significant issue (SUN 2009) when choosing agent architectures.

Thus, agents in this research should be illustrated in Figure 2.4.

The requirements for agent architectures in the target conceptual model for large-scale ABSS are as follows:

Research Requirement - Agent Architecture The architecture for agents in the target conceptual model for large-scale ABSS should satisfy the following requirements.

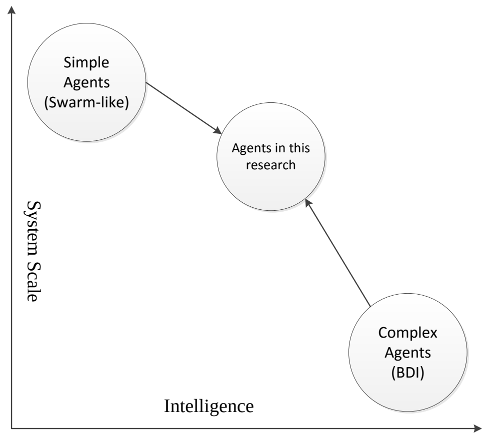


Figure 2.4: The illustrated place of agent in this research

- R1 It should be easily decomposable into components. -Interpreted from CH1
- R2 It should have its own decision-making capability. -Interpreted from CH2 & CH7
- R3 The architecture should support a mechanism to sense and reflect the surrounding environment. -Interpreted from CH3 & CH4
- R4 It should have dynamical social interaction capability. -Interpreted from CH4
- R5 It should have the ability to response to emergent events, rather than having to blindly follow a prearranged plan. -Interpreted from CH5 & CH6
- R6 The agents should be "simple" and "small". -Interpreted from scalability requirement

Based on these requirements, we conducted a comparison in Table 2.1, which compared the mentioned agent architectures above.

Table 2.1: Comparison of agent architectures

	Template architectures	**Cognitive architecture**	**BDI architecture**	**Normative architecture**	**Multi-level/layer architecture**
Example	Aglets/ JAT	SOAR/ ACT-R	PRS/ JAM	EMIL-A/ NoA	Touring Machines/ SIM_AGENT
Primary Domain	Mobile agents	General purpose AI; learning	Reasoning agents	Norm emergence in multi-agent systems	Individual agent with multi-layer architecture
Component-based(R1)	No	Yes	Yes	Yes	Yes
Decision-making Capability (R2)	Hard	Production Rule	Meta-level	Normative	Multi-level
Interaction ability to Environment (R3 & R4)	Yes	Yes	Yes	Yes	Hard
Working memory (R5)	No	Not explicitly defined	The active beliefs and intentions	Yes, Normative	Yes, exists in Specialized layers
Model Realization (R6)	Implemented through Templates	Simulated as parts of simulation	Simulated as parts of simulation	Implemented through interpreted programming languages	Simulated as parts of simulation

According to the comparison on different requirements of current agent architectures in Table 2.1, we can see that all agent architectures have their own difficulties to satisfy the requirements for our target large-scale social systems which is shown in Table 2.2.

Table 2.2: Difficulties of applying current agent architectures in large-scale social systems

Agent Architecture	Difficulties
Template agent architectures	R1,R2,R3,R5,R6
Cognitive architecture	R1,R6
BDI architecture	R6
Normative architecture	R1,R6
Multi-level/layer agent architectures	R6

2.4 Social Networks in ABSS

One of the practices of agent-based models is that agents only interact with the agent's neighbors, and local information is obtained from interactions with an agent's neighbors and its localized environment. Thus, an agent's neighbors could change rapidly as agents move through space when simulating (MACAL, HOWE et al. 2006), as they will continuously get new neighbors. However, this practice is insufficient for large-scale social systems where agents can easily contact other agents far way.

Another way to manage the relationship of agents is by a topology, such as a spatial grid or network of nodes (agents) and links (relationships). This topology describes who transfers information to whom. However, agents can interact according to multiple topologies. For example, an agent could interact only with its neighbors located close-by in physical (or geographical) space as well as neighbor agents located close-by in its social space as specified by the agent's social network (MACAL and NORTH 2010). When large-scale social systems grow to include millions of agents, assigning a topology for the whole population becomes too difficult. As a matter of fact, modeling social networks is becoming an active research area in the large-scale agent-based social simulation research (ALAM and GELLER 2012).

Basically, research on social networks can be distinguished into two parts: (1) social network structure and (2) social network dynamics (SQUAZZONI et al. 2013). Social network structure defines the interacting agents that an agent has. In general, a social network structure is a graph representation of individuals (nodes/vertices) and the relationships (lines/edges/links/arcs). For each line between two individuals, information can be specified, such as what type of connection it is, how often the agents interact with each other and what information they share.

As defined in graph theory, lines in social networks may be directed from one individual to another which means for instance that friendship may be one-sided. Generally, social network structures have two types (RONALD et al. 2012a): complete networks where the whole topology of the links among nodes are specified (PUJOL and FLACHE 2005, GABBRIELLINI and TORRONI 2014) and egocentric networks where links of nodes are described per node (CARRASCO et al. 2008, GATTI et al. 2014).

Social network structures, in the form of agent organizations in multi-agent systems, are commonly studied (DUBOZ et al. 2006). From their perspective, agent organizations can be seen as a set of agents regulated by social rules and mechanisms with which autonomous agents can achieve common goals under some kind of institutional control. In other words, the organization supports an agent to recognize its role, and the roles of others, in accomplishing those collective goals, with communication as an important means. Organizational theory is therefore increasingly used in agent-based modeling and simulation and support of social systems.

In recent years, various type of organizations were modeled by researchers, such as institutions, groups, firms, and communities. The organizational structure usually involves two fundamental concepts: agent roles and their relations with which the overall behavior of the MAS is determined (GROSSI et al. 2005). In addition to modeling agent organizations, several organizational structures specifications were introduced by KOLP et al. (2006) as the meta-class of organizational structures for MAS, which adopted concepts from organizational theories.

An example of the organizational structure specifications is the Block-like Representation of Interactive Components (BRIC) (WEISS 1999). Another specification dealing with structure specification is the Agent/Group/Role or AGR organization modeling approach (FERBER and GUTKNECHT 1998), in which an organizational structure consists of a set of groups, roles in each group and agents fulfilling roles (BOELLA and VAN TORRE 2006). GROSSI et al. (2005) argues that organizational structures should be seen along at least three dimensions, instead of just one: power, coordination, and control.

However, these research works focus on the descriptive level of organizational structures. When it comes to organization design itself, it is not always easy to model organizational structure on an operational level. The reason is that the relation between agents could change over time due to the autonomous property of agents. Thus, the primary limitation of these specifications is that they don't enable the formalization of the structural evolution of the system (i.e., the dynamics of interactions).

To formalize the system dynamics in multi-agents system, temporal modeling specification languages have been introduced (DARDENNE et al. 1993) in which dynamic properties are often specified in the form of a set of logical formulas. Hence, this method is one of the dominant approaches for specification and analysis of dynamic properties in agent-based systems. The advantage of this method is the declarative modeling of simulation models, for example, Executable Temporal Logic (AMIGUET et al. 2002) and the Strictly Declarative Modeling Language SDML (EDMONDS 2002). The shortcoming of this specification is that it usually does not provide explicitly specified organizational structure nor does it offer dedicated support for a specific type. Moreover, simulation of dynamics is the main purpose of this specification and not much formally defined support is offered for analysis of dynamics.

Instead of adopting specifications, there are many other ways to look at social networks. From the perspective of what characteristics of social networks to focus on, social networks have three different categorizations: the structure-oriented, the actor-oriented, and the actor-structure crossing (JIANG et al. 2014). From the perspective of how to add new nodes in a social network model, the social network models are classified into two main categories: network evolution models (NEMs) and nodal attribute models (NAMs) (TOIVONEN et al. 2009).

Different algorithms were developed to help generate social networks for different research purposes. BADHAM and STOCKER (2010) introduced a spatially based algorithm that generates networks with constrained but arbitrary degree of distribution, clustering coefficient and assortativity (preference for a network's nodes to attach to others that are similar in some way). This algorithm randomly creates the nodes in space and assigns a target degree to each. A different approach is based on dynamically constructing networks on the basis of the likelihood that people connect. It has been observed that the similarity between people, or homophily, increases that chance that people talk (BROWN and REINGEN 1987), and highly similar people are more likely to connect than those who are very different (KOSSINETS and WATTS 2009). Similarity is often measured in terms of attributes such as age, gender, education, or lifestyle. Greater similarity between people seems to also increase trust, understanding, and attraction between them, creating a stronger relationship (RUEF et al. 2003). This introduces the possibility of using social simulation models to look at networks based on individual preferences and characteristics rather than having to collect full network data (SQUAZZONI et al. 2013), which is difficult for large-scale systems.

Thus, we propose a requirement for the target conceptual model for large-scale ABSS as follows:

Research Requirement - Social Networks The conceptual model for large-scale ABSS should support modelers to generate dynamical social networks.

2.5 Social Interaction in ABSS

Social interaction is widely modeled in various areas. The theory of Sociology of Organized Action, also called Strategic Analysis, is a key theory, and it intends to discover the functioning of any organization beyond its formal rules, especially how social actors build the organization that in return rules their behavior (SIBERTIN-BLANC et al. 2006). For over a decade, the FIPA specification, acting as a key standard in MAS area, offers interaction mechanisms for agent to agent communications as well (POSLAD 2007). Social interaction is also modeled in agent-based argumentation (SIERRA et al. 1998), in which agents are socially embedded and exchange information by means of simulated dialogues (GABBRIELLINI and TORRONI 2014). The advantage of agent-based argumentation is that it can generate rational and explainable decisions, while it is not suitable for large-scale complex social interactions as each agent is equipped with a local argumentation theory (KAKAS et al. 2012), which will consume a lot of memory.

Both agent-based argumentation theory and other theories on social interaction mentioned above are designed for small-scale systems. Nevertheless, modeling complex human social interactions is an important part in large-scale agent-based social simulation research. For example, results of interactions (negotiations) for scheduling joint social activities could influence the future plans of the involved individuals, which has a great impact on the researches such as activity-based travel demand analysis and agent-based epidemic models.

Scheduling a joint social activity can be considered as a process of group decision, by which agents plan, negotiate and execute a joint social activity. This process is becoming more important in the field of traffic demand analysis (RONALD et al. 2012a, LIN and WANG 2014). In addition to activity-scheduling behavior, congestion levels at specific times and places emerge in traffic demand analysis system and may lead to discrepancies between scheduled and actual activity and travel times. Agents should respond to such unforeseen events by reconsidering an existing schedule (within-day re-planning) and by adapting their expectations about traffic conditions for subsequent days (learning) (ARENTZE et al. 2010).

Describing this process is a rather difficult task than it may seem, in particular when the system has a very large scale (millions of individuals). Current research efforts in large-scale ABSS ignore or simplify the negotiation/coordination part of the social interactions in order to reduce complexity, either by using fixed and predefined human daily schedules as inputs or by constraining the joint social activities (interaction purposes) into several specific types (e.g. eating out) (ZHANG et al. 2015).

Another limitation of current research practices in large-scale ABSS is that the modeled decision process is only held among family members (MILLER and ROORDA 2003). Thus, only scheduling social activities for families are studied.

Thus, we propose a requirement for the target conceptual model for large-scale ABSS as follows:

Research Requirement - Social Interaction The conceptual model for large-scale ABSS should support agents to join in different types of social activities through complex social interactions.

2.6 Agent Environment in ABSS

The environment is a fundamental concept of agent-based modeling and simulation (ABMS), in which agents exist, interact, perceive, and act (HESAN et al. 2015). In the context of a classic agent-based social simulation, the agent environment is usually represented by a geographical area that may be real or fictitious. One of the simplest methods of representing a fictitious geographical area is using grids or cells. When representing real areas, a vector Geographic Information System (GIS) based spatial environment is more popular (HAWE et al. 2012).

A grid-based or cell-based agent-based simulation is also called a situated agent-based simulation (CICIRELLI et al. 2015), by which agents can interact through information exchange in the grids or cells. An example of a situated ABMS platform is DIVAs, in which the environment (designed as a network of cells) is a distinct active entity that provides an indirect coordination mechanism for agents situated in the cell (MILI et al. 2006). Compared to situated systems, GIS-based systems provide a more realistic way to model agent environments for agent-based social simulation. Many middleware approaches are developed to link existing GIS and ABM models in order to enable interaction between geographic data (fields and objects) and agent-based process models (BROWN et al. 2005). Generally, a GIS-based agent environment is more popular in current large-scale ABSS models as it can not only deal with spatial data but also with other data such as culture, political ideology or religion (ADAMATTI 2014).

Many recent projects on social simulation involving policy makers and stakeholders have shown that environment information contains not only static elements, such as spatial entities providing GIS information, but also active and interdependent artifacts (BOERO 2006). Thus, artifacts, providing the services and functions that influence individual agents together with spatial information are both required to shape agent environment (OMICINI et al. 2008).

Besides the active and interdependent elements that influence individual agents, social norms influence agents' behavior institutionally (SAVARIMUTHU et al. 2008). In multi-agent systems (MAS) literature, an agent environment includes social aspects in addition to physical spaces (HESAN et al. 2015). However, current large-scale ABSSs usually don't include social concepts such as norms or institutions. This is because, in large-scale ABSS models, agents behave according to fixed schedules while norms and institutions can only take effect when agents have mechanisms for norm compliance (SAVARIMUTHU et al. 2008). The contradiction relies on the lack of the right kinds of beliefs (or expectations) and preferences in agents with fixed schedules, which are required for modeling people with social norm compliance (BICCHIERI 2005).

In addition to norms, there are many other concepts, frameworks and perspectives that are developed in social science areas, which should be addressed clearly in large-scale ABSS because techniques and insights borrowed from these other discipline can be beneficial (FOO and PEPPAS 2005). We will list several popular and important concepts in social science research that can be/have been imported into agent-based social simulation research.

1) Norm. Norms, as a means of regulating social systems consisting heterogeneous and autonomous agents,are prohibitions, permissions and obligations associated with agents (VASCONCELOS et al. 2012). Norms play a crucial role in social systems research because they: (i) regulate the behavior of agents, and (ii) create expectations on the behavior of other agents (VIGANÒ et al. 2006). With social norms, social order can be facilitated (SAVARIMUTHU et al. 2011).

2) Institution. Institutions are viewed as a collection of social constraints (NORTH 1990), which are used to regulate relations among agents. These constraints are either informal, for instance customs and traditions, or formally defined, like laws and regulations, which affect the formation of agent organizations (ALDEWERELD 2007). According to CRAWFORD and OSTROM (1995), the central pieces of an institution are strategies, norms and rules, which can be distinguished by their grammatical texture. Thus, an institution is a concept greater than a collections of norms.

3) Organizational concept. Organizational concepts such as groups, roles and structures, define schemes for describing agent coordination and negotiation in multi-agent systems (FERBER and GUTKNECHT 1998). Organizational concepts can be translated into norms and institutions (HÜBNER et al. 2011).

Therefore, we propose a requirement for the target conceptual model for large-scale ABSS as follows:

Research Requirement - Agent Environment The agent environment in the conceptual model for large-scale ABSS should be separated into different concepts to represent spatial information together with other independent artifacts and social concepts that can influence agent behavior.

2.7 Conclusion

This chapter presented background theories and concepts that are related to ABSS models. Firstly, the diversity of application areas of ABMS and ABSS was presented. Then, current conceptual models for ABSS were discussed together with their shortcomings. The concepts in the conceptual model, such as agent, agent architecture, social networks, social interactions and agent environment were presented and discussed. With this chapter, the requirements for a general conceptual model for large-scale ABSS were explained and presented as follows.

RR1 Research Requirement - Model Architecture A conceptual model is required for large-scale ABSS.

RR2 Research Requirement - Agent Architecture The architecture for agents in the target conceptual model for large-scale ABSS should satisfy the following requirements.

RR2.1 It should be easily decomposable into components. -Interpreted from

CH1

RR2.2 It should have its own decision-making capability. -Interpreted from CH2 & CH7

RR2.3 The architecture should support a mechanism to sense and reflect the surrounding environment. -Interpreted from CH3 & CH4

RR2.4 It should have dynamic social interaction capability. -Interpreted from CH4

RR2.5 It should have the ability to respond to emergent events, rather than having to blindly follow a prearranged plan. -Interpreted from CH5 & CH6

RR2.6 The agents should be "simple" and "small". -Interpreted from scalability requirement

RR3 Research Requirement - Social Networks The conceptual model for large-scale ABSS should support modelers to generate dynamic social networks.

RR4 Research Requirement - Social Interaction The conceptual model for large-scale ABSS should support agents to join in different types of social activities through complex social interactions.

RR5 Research Requirement - Agent Environment The agent environment in the conceptual model for large-scale ABSS should be separated into different concepts to represent spatial information together with other independent artifacts and social concepts that can influence agent behavior.

3

Large-scale Agent-based Epidemic Prediction and Control as a Case of Large-scale ABSS

3.1 Epidemics

The meaning of the term *"epidemic"* depends on the context in which it is used (GREEN et al. 2002). An epidemic can be defined as "the occurrence in a community or region of cases of an illness, specified health behavior, or other health-related events clearly in excess of normal expectancy; the community or region, and the time period in which cases occur, are specified precisely (LAST 2001)".

From the earliest times to the present, epidemics have affected human history in myriad ways: demographically, culturally, politically, financially, and biologically. Humans have never known a time in history when epidemics did not threaten them. This is as true today as it always was [1]. Epidemics have an overwhelming impact on a population both directly through damaging health and indirectly through causing panic, disrupting the social and economic structure and impeding development in the affected communities. There are quite a number of excellent introductions to the history of epidemics and its effects on history (McNeill 2010, Hays 2009, Harrison 2013).

Recent large outbreaks, just since the start of this century, have shattered a number of myths about the world's vulnerability to threats arising from new pathogens and epidemic-prone diseases like Ebola [2]. A time-line on epidemics since the start of the 21st century is present in Figure 3.1.

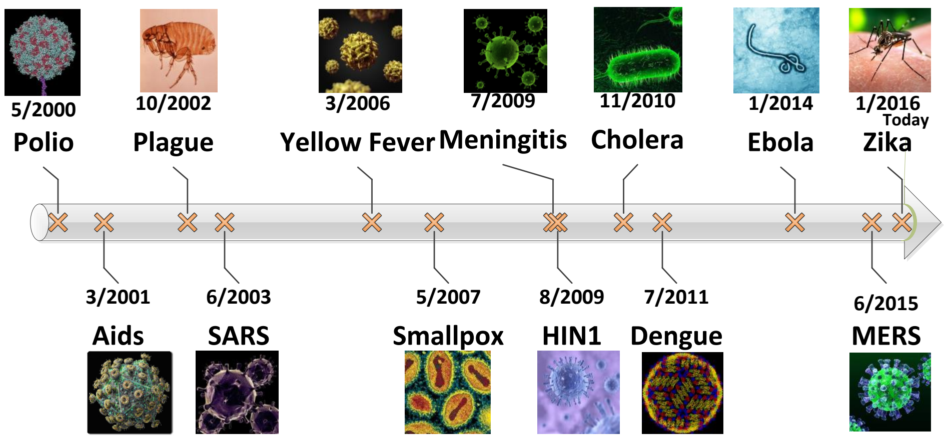


Figure 3.1: Epidemics of 21st century

A lot of research has been conducted to understand epidemics. There are more than 20 thousand results since 2015 when searching the term "epidemic" in Google Scholar. However, there are still many challenges remaining in understanding epidemics (Lloyd-Smith et al. 2015).

Generally, transmission of an infectious disease may occur from one person to another by one or more of the following means (Straif-Bourgeois et al. 2014): direct physical contact (e.g., touching), indirect physical contact (e.g., contaminated food) or vector-borne contact (e.g., a droplet). These contacts usually occur among people in geographical spaces, either open environments or interior spaces, where people can quickly or easily get in touch with each other directly or indirectly. For example, if an infected person coughs or sneezes in a bus, the droplets containing microorganisms may enter another person's body,

[1] Christian W. McMillen, Epidemic Diseases and their Effects on History, last modified on 24 JULY 2013, retrieved from `http://www.oxfordbibliographies.com/document/obo-9780199743292/obo-9780199743292-0155.xml`.

[2] WHO, How the 4 biggest outbreaks since the start of this century shattered some long-standing myths, Retrieved from `http://www.who.int/csr/disease/ebola/ebola-6-months/myths/en`.

which causes a disease to spread. This is considered as a basic model for disease transmission. Thus, understanding of the transmission patterns of a hypothetical epidemic among a susceptible population (MOSSONG et al. 2008) is pivotal for epidemic modeling research (PEREZ and DRAGICEVIC 2009). Similar to other social system models, typical epidemic models are based on mathematical models (BOBASHEV et al. 2007) or agent-based models (AJELLI et al. 2010).

3.2 Mathematical Epidemic Models

The first mathematical approach to the spread of a disease was proposed by BERNOULLI (1760) in 1760, while KERMACK and MCKENDRICK (1927) defined the modern mathematical modeling of infectious diseases in 1927. Mathematical epidemic models generally assume that the population can be divided into different compartments depending on the phase of the disease (ANDERSON et al. 1992) such as susceptibles (denoted by S, those who can be infected), infectious (I, those who are infected and contagious), and recovered (R, those who recovered from the disease). Another assumption is that the total population (denoted by N) in the system is fixed, which means other demographic processes such as migrations, births, and deaths are ignored.

Based on these assumptions, many mathematical epidemics models have been proposed with different transition phases.

The Susceptible-Infected-Susceptible (SIS) model (IANNELLI et al. 1992, ALLEN 1994) separates the population into two compartments, susceptibles and infectious. Thus, there are only two possible transitions among the population, either from S to I when a susceptible individual interacts with an infectious individual and becomes infected or from I to S when the infectious individual recovers from the infection. This model assumes that people who are infected will never become immune.

Another mathematical model for epidemics is the Susceptible-Infected-Recovered (SIR) model (SHULGIN et al. 1998) which separates the population into three compartments, susceptibles, infectious and recovered. In an SIR model, an infectious individual can only transit to recovered. It assumes that people can either acquire a permanent immunity or be removed from the total population (e.g. because of death). Thus, this model is more realistic for certain diseases (e.g., chickenpox) than for others (e.g., the flu).

The SIR model is often used to study the epidemics by tracking the number of susceptible, infectious and recovered population through three differential equations for $S(t)$, $I(t)$, and $R(t)$:

$$\frac{dS}{dt} = -\beta S(t)I(t) \tag{3.1}$$
$$\frac{dI}{dt} = \beta S(t)I(t) - kI(t) \tag{3.2}$$
$$\frac{dR}{dt} = kI(t). \tag{3.3}$$

In this equation set, k is the recovery rate and β is the probability of transmitting the disease when an infected interact with a susceptible person.

A more complex mathematical model is the Susceptible-Exposed-Infectious-Recovered (SEIR) model (D'ONOFRIO 2002). Compared with the SIR model, the phase of exposed is included in the transition process to represent those individuals who have been infected and are undergoing an incubation period. Individuals in the phase exposed are not infectious yet and no symptoms are shown.

With this added phase, the SEIR model provides a more accurate abstraction of the various stages of some diseases, such as Ebola. The dynamics of population in different phases of diseases in the SEIR model can be modeled by the following set of differential equations:

$$\frac{dS}{dt} = -\frac{\beta SI}{N}, \tag{3.4}$$
$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E, \tag{3.5}$$
$$\frac{dI}{dt} = \sigma E - \gamma I, \tag{3.6}$$
$$\frac{dR}{dt} = (1-f)\gamma I. \tag{3.7}$$

In this equation set, β is the transmission probability of a disease, σ is 1 over the incubation period, γ is the recovery rate, and f is the fatality rate.

In addition to the SIS, SIR, SEIR model, there are many other mathematical models designed for various disease transmission processes, such as SEIS (FAN et al. 2001), SEI (LI and ZHEN 2005), SIRS (XU and MA 2009) and SEIRS (COOKE and VAN DEN DRIESSCHE 1996).

With assumptions and simplifications of the complex spreading process of epidemics, mathematical models have good performance in theoretical analysis of macroscopic regularities of epidemic diffusion, such as the epidemic threshold and final epidemic size (DUAN et al. 2015). However, mathematical models estimate the speed of a disease outbreak based on the basic reproduction number which depends on the number of adequate contacts (DEL VALLE et al. 2007), while the contact details often rely on priori contact assumptions with little or no empirical basis (MOSSONG et al. 2008) in the form of a set of parameters, for example, household contact rates, school contact rates and workplace contact rates

(GREFENSTETTE et al. 2013). Thus, current mathematical models rarely consider the heterogeneous process of disease transmission (GUO et al. 2015), such as heterogeneous contagiousness of infectious individuals, and do not reveal realistic contact patterns due to the difficulties in modeling demographic stochasticity and spatial heterogeneity (BEN-ZION et al. 2010). Furthermore, a small set of parameters in mathematical models are inadequate to capture the variety of factors associated with the epidemic spread process, especially the determining factors, human behavior and heterogeneous links between individuals (DUAN et al. 2015).

In recent years, a second type of epidemic modeling has become popular. The real-world accuracy of the models used in epidemiology has been considerably improved by the integration of large-scale datasets and the explicit simulation of entire populations down to the scale of single individuals (PASTOR-SATORRAS et al. 2014), which is agent-based epidemic modeling.

3.3 Agent-based Epidemic Models

Due to the increasing threat from epidemics and the shortcomings of the mathematical models, agent-based epidemic models are getting more popular. To name a few, the disease that has been studied using agent-based methods include tuberculosis (TB) (DE ESPÍNDOLA et al. 2011), smallpox (GRUNE-YANOFF 2010), SARS (HUANG 2010), malaria (LINARD et al. 2009), CA-MRSA (MACAL et al. 2012), HIV (MEI et al. 2010), measles (PEREZ and DRAGICEVIC 2009), AHC (CHEN et al. 2014) and H1N1 (STROUD and VALLE 2007).

The popularity of adopting agent-based methods comes from the fact that ABM can characterize each agent with a variety of variables that are considered relevant to model disease spreading such as mobility patterns, social network characteristics, socio-economic status, health status, etc. (FRIAS-MARTINEZ 2011).

Recently, due to the growth of computational power, large-scale agent-based modeling and simulation have become possible for epidemic models (STROUD and VALLE 2007, PARKER and EPSTEIN 2011, AJELLI et al. 2010, RAKOWSKI et al. 2010, BISSET et al. 2009a, BISSET et al. 2014, BARRETT et al. 2008).

As we stated in Chapter 1, current large-scale ABSS models for epidemic predictions and control are usually developed from scratch in the OOP languages which causes large amount of repetitive works in studying large-scale epidemics. To tackle this, the objective of this research is to design and test a conceptual model architecture that can help simulation modelers to construct models of large-scale agent-based epidemic and prediction. The conceptual model architecture should not only meet the general requirements for large-scale ABSS presented in Chapter 2, but also the requirements for this specific case, large-scale epidemic prediction and control, in the following aspects.

The first is related to disease modeling. New diseases are emerging at an unprecedented rate of one per year, such as MERS (Middle East Respiratory Syndrome), SARS, Ebola and the variance of the novel Influenza. Although mathematical researchers have developed many disease spread models, such as SIR and SEIR, these models are usually disease-specific, which means a lot of more efforts are needed to develop a new disease spread model when a new disease with a different phase transition process emerges. Thus, the requirement for modeling diseases in large-scale agent-based epidemic prediction is:

RR6 Research Requirement - Disease in the Case Study Large-scale epidemic prediction and control should have a flexible mechanism to model new diseases with a different phase transition process.

The second is about policy modeling. Although current agent-based epidemic models contribute to help health-care policy makers investigate the effects of different interventions, they have inherent limitations on modeling new policies during an epidemic outbreak (GREFENSTETTE et al. 2013), such as the difficulty to test a combination of multi policies. Thus, the requirement for modeling policies for large-scale agent-based epidemic prediction is:

RR7 Research Requirement - Policy in the Case Study Large-scale epidemic prediction and control should have a flexible mechanism to model new policies with different settings.

More requirements for the case of large-scale agent-based epidemic prediction and control are reflected in the model components.

3.3.1 Agent Behavior

Agent behavior in agent-based simulation is usually specified by agent architectures (ATKIN et al. 2001). To model the behavior of agents for epidemic prediction and control, the agents should be able to make complex decisions in daily life, interact for social activities and respond to emergency during the period of an epidemic outbreak. However, agents in current large-scale epidemic prediction models are rather simple in terms of both the agent architecture and the decision-making mechanism (HAWE et al. 2012). Agents in these models either make decisions based on simple rules, or behave according to initially set fixed schedules.

The reasons are two-fold. Firstly it seems there is no necessity for modeling complex reasoning agents for epidemic predictions. The interest in the current large-scale agent-based epidemic models stems from the requirement to understand the correlation between human behavior, the epidemic dynamics and the potential interventions (MACAL et al. 2012, PEREZ and DRAGICEVIC 2009, ZHOU et al. 2012). Since individual responses can be categorized into several common types and incorporating these response types into either simple rules or fixed schedules seem to be adequate to predict the common human behavior, the mechanisms and processes for agents to reason and to make complex decisions are no longer necessary.

Another consideration is from the perspective of simulation performance. Complex reasoning and decision-making process involves a large amount of coordination among the components of the agent architecture, which will cause countless communication messages and greatly decrease the simulation performance (SWARUP et al. 2014). In addition, agents' local argumentation knowledge will decrease simulation performance as well, as they will consume a lot of memory when the system scale increases. In large-scale epidemic predictions, even simple rule-based agents become a luxury due to the performance limitation. Thus, fixed schedules for agents start to serve as a common law for agent-based large-scale epidemic predictions.

However, modeling agents' complex decisions to some extent is essential for epidemic prediction. Due to the uncertainties of new emerging diseases and the fast information acquisition and progression, human responses to unforeseen events are becoming difficult to predict. Thus, the resulting emergent human behavior through the interactions of simple agents is becoming less trustworthy.

Thus, the requirement for agents in large-scale agent-based epidemic prediction is:

RR8 Research Requirement - Agents in the Case Study Agents for large-scale epidemic prediction should have capabilities to make complex decisions to some extent during a disease outbreak.

3.3.2 Agent Environment

The environment is a fundamental concept of agent-based modeling and simulation (ABMS), in which agents exist and interact, can perceive and act. Thus, a common agent environment in ABMS is considered as a physical space (typically a 2D grid) or as a virtual space that supports agent-to-agent interaction (HESAN et al. 2015).

In current agent-based epidemiology research, an agent environment represented by physical rooms is a key component as the majority of transmissions of contagious diseases is thought to occur as a result of sustained indoor contacts (ANDREWS et al. 2013), such as students taking classes in classrooms and workers working in offices. However, direct physical contact (e.g. touching) or vector-borne contact (e.g. a droplet) outside rooms can be an effective method for disease spread as well, especially in densely populated areas such as public transportation.

Some historical observations examined the role of public transportation for epidemic outbreaks and studied the effect of some interventions related to public transportation. For example, CONDON and SINHA (2010) observed the facemask usage on public transportation in Mexico City in April/May 2009 during a 2-week period and the results showed mask usage rates matched the course of the H1N1 epidemic. WANG (2014) showed that the dynamics of the Taipei underground usage during the 2003 SARS epidemic in Taiwan were closely linked to the daily reported probable SARS cases.

Some mathematical epidemic models also investigated the possible effect that using public transportation may have on the spread of contagious diseases. The model and results from ANDREWS et al. (2013) indicate that public transportation may play a critical role in the transmission of tuberculosis. An investigation by XU et al. (2013) gave the conclusion that increasing transportation efficiency and improving sanitation and ventilation of the public transportation system decrease the chance of an outbreak occurring. ZHOU et al. (2012) used a mathematical model called SIS (susceptible-infectious-susceptible) and studied the impact of the preference of an individual for public transport on the spread of infectious diseases. More generally, several mathematical epidemic models were proposed to investigate disease transmission among regions with infection during travel, such as SIS (CUI et al. 2006, TAKEUCHI et al. 2007), SEIS (WAN and CUI 2007), SEIRS (DENPHEDTNONG et al. 2013) and SIQS (LIU and TAKEUCHI 2006).

The popularity of examining the role of public transportation (or travel in general) by mathematical models comes from the fact that they can easily estimate the likelihood of a disease outbreak based on the basic reproduction number which depends on the number of adequate contacts (DEL VALLE et al. 2007), while the contact details often rely on a priori contact assumption with little or no empirical basis (MOSSONG et al. 2008) in the form of a set of parameters, for example, contact rates (GREFENSTETTE et al. 2013). Thus, mathematical models can not reveal realistic contact patterns due to the difficulties in modeling demographic stochasticity and spatial heterogeneity (BEN-ZION et al. 2010).

In agent-based epidemic research, GREFENSTETTE et al. (2013) and PARKER and EPSTEIN (2011) use gravity models with simplified assumptions to model the general travel patterns in order to recreate random contacts during travel. RAKOWSKI et al. (2010) considered only the intermediate breakpoints (transfer cities) between endpoints (the origin and the target cities) to determine the number of co-travelers for each traveling agent during his travel. PEREZ and DRAGICEVIC (2009) modeled a transportation network to represent the movement path as a trajectory in space for disease propagation, while disease doesn't propagate during the transportation. All these researches claim that these methods are sufficient to model disease spread during general travel (public transportation is not even mentioned) to some extent, however, none of them offer the ability to test some specific interventions on controlling individual travel behavior (e.g., shut down one metro line).

According to our findings in the literature, COOLEY et al. (2011) developed an agent-based model of New York city that incorporates subway ridership, which simulates the interactions of subway riders and examines the impact that a severe influenza epidemic would have on NYC and the potential effects of different hypothetical subway-related disease control measures. In our opinion, the scarcity of including public transportation or travel in general does not mean it is under-appreciated in agent-based epidemic models, but the reason can be explained as the consideration of simulation performance. Due to the high resolution of simulation time in a microscopic traffic model, the inclusion of the traffic component can greatly decrease the simulation performance.

Thus, including complete physical spaces into agent environment for epidemic research still remains a challenging task for current large-scale models as they typically omit or simplify many of the temporary contacts during traveling.

Thus, the requirement for an agent environment for large-scale agent-based epidemic prediction is:

RR9 Research Requirement - Agent Environment in the Case Study The agent environment for large-scale epidemic prediction should model different kinds of physical spaces including movable spaces for disease transmission.

3.3.3 Social Interaction

Besides the agent environment, modeling complex human social interactions is also an important part in agent-based epidemic and pandemic predictions (STROUD and VALLE 2007, MOSSONG et al. 2008, RAKOWSKI et al. 2010, GE et al. 2013) due to the fact that social interactions provide a perfect fabric for fast disease propagation while they can be dramatically altered when people respond to the crisis and interventions (BISSET et al. 2009b). For example, results of interactions (negotiations) for scheduling joint social activities could influence the future plans of the involved individuals, which has a great impact on the contacts of people. To describe these interactions is a more difficult task than it may seem, in particular when the system has a very large scale (millions of individuals). Current research efforts ignore or simplify the negotiation/coordination part of the social interactions in order to reduce complexity, either by using fixed and predefined human daily schedules as input or by constraining the joint social activities (interaction purposes) into several specific types (e.g. eating out). Thus, to model complex social interactions, social structures in current large-scale agent-based epidemic models are not implemented independently of individual agents, but as properties of agents (RONALD et al. 2012b).

In the model EpiSimS (STROUD and VALLE 2007), there are no predefined or dynamically generated social networks in the model. To eliminate the need to simulate every single agent's day-to-day activities, explicitly stored social networks and random contacts were considered in a global-scale model (PARKER and EPSTEIN 2011) to replace social interaction. In the model of (AJELLI et al. 2010, RAKOWSKI et al. 2010), no social networks are discussed. In EpiFast (BISSET et al. 2009a), INDEMICS (BISSET et al. 2014) and CHEN et al. (2014)'s work, social contact networks representing proximity relationships between individuals of the population were considered as input data.

Furthermore, since using a single computing core may be inadequate (HAWE et al. 2012) to deal with the scalability issue, some of the implementations are based on distributed architectures. However, there could be a huge overhead for enabling coordination between agents on distributed architectures as it increases the number of communication messages and leads to a higher communication complexity. As a matter of fact, to balance between performance and accuracy for large-scale agent-based models, reducing complex social interactions is an often used compromise (STROUD and VALLE 2007, PARKER and EPSTEIN 2011, AJELLI et al. 2010, RAKOWSKI et al. 2010, BISSET et al. 2009a, GE et al. 2013).

Thus, the requirement for social interaction for large-scale agent-based epidemic prediction is:

RR10 Research Requirement - Social Interaction in the Case Study Large-scale agent-based epidemic prediction and control should have capabilities to model complex human social interactions.

3.4 Conclusion

This chapter presented background theories and concepts that are related to the case study of epidemic prediction and control. Firstly, the basic context for epidemics was given. Then, current mathematical models for epidemics were discussed together with the shortcomings. At last, we presented a review of agent-based epidemic models in terms of agent behavior, agent environment and social interaction. With this chapter, the requirements for a general conceptual model architecture for large-scale epidemic prediction and control were clearly explained and presented as follows:

RR6 Research Requirement - Disease in the Case Study Large-scale epidemic prediction and control should have a flexible mechanism to model new diseases with different phase transition process.

RR7 Research Requirement - Policy in the Case Study Large-scale epidemic prediction and control should have a flexible mechanism to model new policies with different settings.

RR8 Research Requirement - Agents in the Case Study Agents for large-scale epidemic prediction should have capabilities to make complex decisions to some extent during a disease outbreak.

RR9 Research Requirement - Agent Environment in the Case Study The agent environment for large-scale epidemic prediction should model different kinds of physical spaces including movable spaces for disease transmission.

RR10 Research Requirement - Social Interaction in the Case Study Large-scale agent-based epidemic prediction and control should have capabilities to model complex human social interactions.

4

A General Conceptual Model for Large-scale ABSS

4.1 Model Concepts and A Conceptual Model for Large-scale ABSS

In this section, we will provide a proposed conceptual model for large-scale agent-based social simulation, which borrows many concepts from other popular ABM conceptual frameworks.

The proposed conceptual model for large-scale social systems is presented in Figure 4.1. This conceptual model separates general concepts into concrete concepts which are designed for large-scale model development.

There are five concepts in this conceptual model, which are:

- Agents. Agents represent human beings in the real world.
- Physical container. The concept of a physical container is used to represent the physical environment in which an agent stays. Physical containers are hierarchically organized and every agent has to stay in at least one phys-

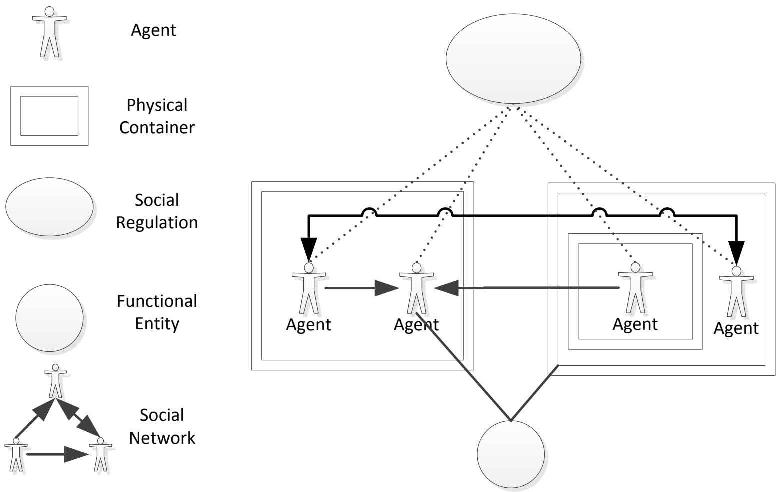


Figure 4.1: Conceptual Model for Large-scale Agent-based Social Simulation

ical container at any time. Typical physical containers are bedroom, office, classroom, bus, etc.

- Social regulation. The concept of social regulation is designed to model social concepts such as norms and institutions that can guide and influence human behavior globally. For example, the policy 'closing public transportation during an epidemic outbreak' can be modeled as a social regulation in the system of epidemic prediction and control.
- Functional entity. Functional entities are those extra objects in the system that can influence or directly change attributes of either agents, physical containers or social regulations. For example, a disease is modeled as a functional entity to change agents' healthy status. A storm is modeled as a functional entity to change the temperature of a room (physical container).
- Social network. The concept of social network defines agents' social relations in different categories, such as family member, classmates and friends.

Compared to the model concepts in a general ABM conceptual model discussed in Section 2.2, the major difference is that the concept of agent environment is separated into physical container, social regulation and functional entity. This separation overcomes the limitations on environmental completeness in other ABM models and provides flexibilities in simulating different system scenarios. On the other hand, the concepts themselves were also refined. For example, this thesis will show that physical containers can be movable to represent transportation vehicles which are difficult to implement in general ABM platforms. Moreover, this thesis will also prove that theories and concepts on social regulation from artificial intelligence can be easily implemented and integrated into an agent-based model while showing reasonable performance, which will be presented in Section 7.1.1.

Besides the general ABM concepts, the proposed conceptual model borrows concepts from classic object-oriented simulation models as well, which can be seen as a hybrid conceptual model (ZHANG et al. 2014). This design choice is shown in Figure 4.2.

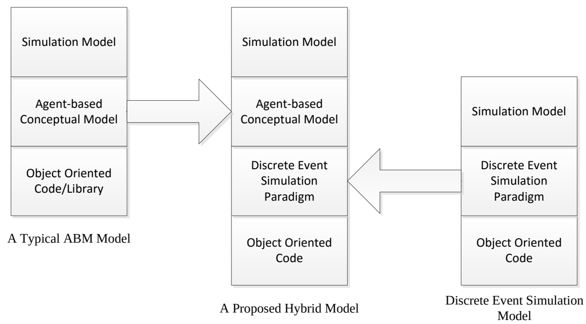


Figure 4.2: Proposed hybrid model for large-scale agent-based social simulation

Historically, the well-known UML-compliant diagrams are widely used in object-oriented discrete event simulation to represent the required conceptual information in three models: the Object Model, the Dynamic Model and the Functional Model (PASTOR et al. 1998), where the Object Model declares system classes including their attributes and services, the Dynamic Model specifies inter-object interactions, and the Functional Model captures the semantics associated with the changes of state of the objects motivated by the service occurrences.

Inspired by this concept of Functional Model in the object-oriented conceptual model, the concept of a functional entity is borrowed from classical ABM conceptual model and introduced in our proposed conceptual model for large-scale ABSS. With this clear separation of concepts which is much easier for implementation using the object-oriented paradigm, experimental results in the following chapters show that models adopting this conceptual model are reasonably more efficient in terms of system performance and low agent-to-agent communication cost than a pure ABS model (ZHANG et al. 2014) while keeping high model fidelity and the same agent capabilities.

Figure 4.1 provided the key concepts of the proposed conceptual model for large-scale ABSS. How each concept can be designed and implemented will be illustrated in the following sections.

4.2 Agent

Similar to the agent concept in the general ABM conceptual model (KROGSTIE et al. 2007), the agent concept in a large-scale ABSS model should own individual behavior and internal architecture before it can be implemented.

4.2.1 Agent Architecture

From the comparison of current agent architectures in Table 2.2 in Chapter 2, it can be seen that multi-level architecture and BDI architecture have less difficulties when applying to large-scale social systems. However, neither of them are "simple" and "small" enough for implementation in large-scale agent-based social simulation. Thus, a novel agent architecture was designed in this research for large-scale agent-based social simulation, which simplifies the traditional multi-level architecture but still keeps certain level of decision-making capability.

Large-scale social systems usually contain millions of agents and their schedules. Typical implementations of behavior schedules of agents in large-scale social systems research are activity-based, where all activities for the whole simulation are predefined in the input data source (GE et al. 2014, BARRETT et al. 2008) or generated before the simulation run (STROUD and VALLE 2007) which consumes a lot of memory. Assume there are around 20 million agents and each agent has 10 activities per day, then the total number of activities for a 4 weeks simulation period is 5.6 billion. To reduce memory consumption, activity pattern-based approach was introduced into multi-level architecture in this thesis, which is inspired by MOSSONG et al. (2008)'s research that human behavior patterns are remarkably similar among people in different countries and the patterns are highly correlated with age.

Activity pattern-based approach emerged in the 1970s, and is becoming increasingly popular in the travel planning area (ISTRATE et al. 2006, ARENTZE and TIMMERMANS 2000). The sequence of activities and travel that a person undertakes is defined as the individual's activity-travel pattern for the day (BHAT et al. 2004). The experiments show that models developed by this approach considerably affect activity planning and rescheduling behavior of individuals (PAS 1984), although large number of individuals will share same patterns which makes the individual travel model much simpler.

Inspired by this approach, we constructed an agent model architecture for large-scale agent-based social simulation (presented in Figure 4.3), which is a integration of an activity pattern-based approach and a multi-level architecture.

This agent architecture consists of three main parts: (1) agent object, (2) activity pattern, and (3) multi-level decision-making module.

An agent object, as part of the agent architecture, is the body which is responsible for updating the agent status as the carrier, receiving, processing and forwarding input messages to corresponding decision-making modules, and enabling agents to behave according to activity patterns.

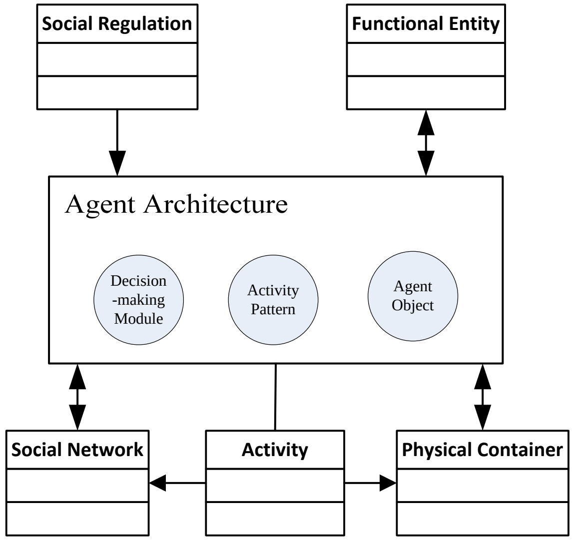


Figure 4.3: Agent Model Architecture

For a given agent, an activity pattern specifies which behavior schedule will be conducted. Based on the activity pattern, agents will mainly perform their activities according to the initial predefined sequences of activities. However, this schedule doesn't specify how long, when, where and with whom these activities take place, which are decided by the decision-making module.

The decision-making module serves as the brain of the agent architecture which is specially designated for some proposed decision-making problems. It's considered as a supplement to agent's behavior pattern.

The detailed definition and formalization of these three parts are given in the following sections.

4.2.2 Agent Object Definition

An agent object is designed as the fundamental component of the proposed agent architecture. It is designed as the behavior body and status indicator while the concrete tasks such as decision-making are undertaken by other components. As a matter of fact, each agent object is associated with a set of attribute variables and status variables. The attribute variables mainly include the agent's individual information such as gender, age, activity pattern, home location and work/school location. The status variables mainly include agent's current behavior and status, such as its current physical container, current activity, and current active social network. Take the epidemic prediction problem as an example, an agent object in such systems can be defined as follows if it's implemented in an object oriented language (e.g., JAVA).

```
public abstract class Agent{
/** <Variable> gender */
private final Gender gender;

/** <Variable> age */
private byte age;

/** <Variable> home location as a 'PhysicalContainer'*/
private PhysicalContainer homeLocation;

/** <Variable> work location as a 'PhysicalContainer'*/
private PhysicalContainer workLocation;

/** <Variable> current physical container where the agent is staying */
private PhysicalContainer currentLocation;

/** <Variable> activity pattern */
private ActivityPattern activityPattern;

/** <Variable> current activity in the pattern */
private Activity activity;

/** <Variable> healthy status */
private DiseasePhase diseasePhase;

/** <Variable> active Social Network*/
private SocialNetwork activeSNS;}
```

In this definition example of an agent object, the variables such as 'age', 'gender', 'homeLocation' and 'workLocation' are the information of the agent that will remain fixed for a long period during the simulation run, in which 'age' and 'gender' may influence the infection possibilities of some diseases. The variable 'activityPattern' specifies an initial behavior schedule for an agent. During the simulation, agent will perform his/her activities according to a defined schedule referenced by this variable. However, this variable can be replaced when the decision-making module makes a decision under certain regulation or influenced by functional entities.

The variable 'activity' indicates the agent's current activity in the pattern, while 'currentLocation' points to the physical container where the agent performs the current activity. The variable 'diseasePhase' demonstrates the agent's current health status during a disease outbreak. It's important in the agent-based simulation not only for specifying agents' status but also for the results when agents make decisions. The 'activeSNS' is a variable designed for indicating the agent's current active social network.

The above definition for an agent objects is presented as an example in simulation for epidemic predictions. The detailed working mechanism will be explained in the following subsections.

4.2.3 Formalization of Activity Pattern and Activity

As discussed above, a typical method for modeling agents' behavior on a large scale is predefining all activities for the whole simulation period, which consumes a lot of memory. To reduce memory consumption, agents in this research will behave based on activity patterns. This design is based on MOSSONG et al. (2008)'s research results that the mixing patterns and contact characteristics are remarkably similar among people and the patterns are highly assortative with age. APOLLONI et al. (2013) confirmed the coupling relations between the age-dependent mixing profiles and the conditions for the spatial invasion of an emerging influenza pandemic.

Thus, each agent in this research will be assigned a social role according to the attribute 'age' of the agent before assigned a week pattern, as the age of a person is highly related to the social role KITE 1996. An example mapping between age, social role and family role can be found in Table 4.1 as an example. There are 4 social roles, which are infant, student, worker and elder(retired), and 3 family roles, which are child, parent, grandparent. The mapping from social roles to family roles are quite simple. Any student or worker will remain as a child until he/she has a baby and becomes a parent. Any parent will shift to grandparent only after his/her child becomes a parent.

Table 4.1: Mappings between age, social role and family role

Age	**Social role**	**Family role**
0-2	Infant	Child
3-5	Kindergarten student	Child
6-11	Elementary school student	Child
12-17	High school student	Child
18-20	University student	Child
20-25	University student/Worker	Child/Parent
25-54	Worker	Parent
55-59	Worker/Elder	Parent/Grandparent
60+	Elder	Grandparent

Then, agent behavior is represented as an activity pattern according to its social role while each social role can be mapped to several activity patterns. For instance, a worker who has a car will drive to work every working day while one of his/her colleagues may prefer taking public transportation to the office, then these two workers will be assigned two different worker activity patterns. Likewise, a student will be assigned one of the student patterns according to his/her school schedule and his/her preference.

Even when agents have the same activity pattern, they show diverse behavior in terms of activity location and duration. When an activity in a pattern is executed, activity location and duration will be dynamically calculated based on the agent's attributes. Moreover, agents can change their activity patterns if necessary.

The activity pattern can work on a specified time duration basis (e.g., hourly, daily and weekly). The choice depends on the problem domain and time range of the simulation study. For instance, a university student can be assigned a university student week pattern, and a worker can be assigned a worker week pattern if they are modeled for a long-term epidemic problem. A pattern can be shared by multiple agents. To increase the heterogeneity and richness of these schedules, more than one week pattern was designed for each social role. A pattern can also be organized hierarchically. For example, a week pattern is made up of seven day patterns, which can be formalized as follows:

$$\text{weekPattern} = \{\text{ID}, \text{Name}, \text{dayPattern}[7]\} \tag{4.1}$$

For a worker who generally works in the office, the first five days patterns can remain the same, and the last two days can be two repeated weekend day patterns. In the week pattern for retired people, the seven day patterns can be the same. A day pattern can be formalized as follows if it doesn't include any lower time range patterns:

$$\text{dayPattern} = \{\text{Name}, \text{List} < \text{Activity} >\} \tag{4.2}$$

In this case, a day pattern is a sequence of linked activities. However, a day pattern can also be formed by multiple short time range patterns (e.g., 24 hourly patterns) if needed. Typical activities for agents in epidemic models are sleeping, staying at home, working, shopping, eating in a restaurant, going to school, visiting a doctor, etc. Table 4.2 gives one example day pattern for students who prefer taking public transportation (metro and bus) to their school.

Every activity in the above table belongs to an activity type, and the activity types can be defined into three activity root categories in Figure 4.4, which are regular activity, travel activity and social activity. Typical activities, such as sleeping, staying at home, working, shopping and attending school belong to regular activity category.

All activities are extended from an abstract activity. This abstract activity is formalized as:

$$\begin{aligned}\text{AbstractActivity} = \{&\text{Name}, \text{currentLocator}, \text{f(getActivityLocation)}, \\ &\text{f(getDuration)}\}\end{aligned} \tag{4.3}$$

Table 4.2: A day pattern example for students

No.	Activity Name	Activity Type	Duration
1	sleep	StochasticDurationActivity	Triangular(6.0, 7.0, 7.5)
2	transport to school	PublicTransportActivity	based on simulation
3	study	UntilFixedTimeActivity	until 12:00 am
4	lunch and rest	FixedDurationActivity	1 hour
5	study	FixedDurationActivity	4 hours
6	transport to home	PublicTransportActivity	based on simulation
7	family dinner	FamilySynchronizedActivity	Fixed(19:00-21:00)
8	homework	StochasticDurationActivity	Uniform(1.0,2.0)
9	sleep till midnight	UntilFixedTimeActivity	until 24:00

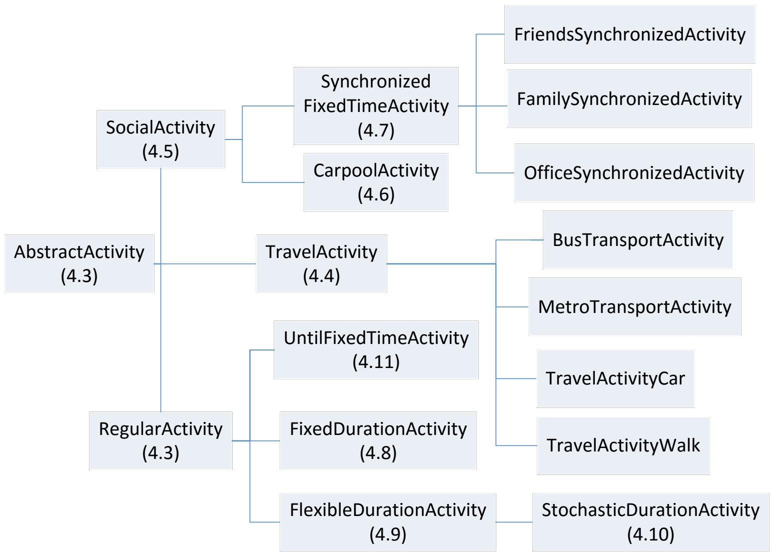


Figure 4.4: Category of activities

An abstract activity defines the capability of using a locator interface to calculate the activity location and duration during the execution phase of the activity, rather than in the definition phase. Extended from this abstract activity, a travel activity is formalized as:

$$\begin{aligned}\text{TravelActivity+} = \{&\text{startLocator}, \text{endLocator}, \text{f(getStartLocation)},\\ &\text{f(getEndLocation)}\}\end{aligned} \quad (4.4)$$

The symbol '+' here means the activity is extended from the abstract activity, same with the other activities below. In addition to the abstract activity, a 'startLocator' and an 'endLocator' interface are added to help the model calculate the start and end location for the execution of the travel activity.

A social activity can be formalized as:

$$\text{SocialActivity+} = \{\text{f(getSocialContacts)}\} \quad (4.5)$$

A carpool activity, as a special form of a social activity, can be formalized as:

$$\begin{aligned}\text{CarpoolActivity+} = \{&\text{startLocator}, \text{endLocator}, \text{carpoolLocator},\\ &\text{f(getStartLocation)}, \text{f(getEndLocation)}, \text{f(getCarpoolLocation)}\}\end{aligned} \quad (4.6)$$

Likewise the travel activity, a carpool activity also owns a 'startLocator' and an 'endLocator' interface. The difference is that a carpool activity has an additional 'carpoolLocator' interface which helps calculate the carpool location.

A 'SynchronizedFixedTimeActivity' is another type of social activity modeled in this research, which specifies the start time and end time of the activity explicitly. However, these time points can be negotiated during the execution phase as agents have corresponded decision-making module to deal with it when receiving unscheduled social activity invitation. It also offers an function to get the activity members when it's finalized properly. A typical 'SynchronizedFixedTimeActivity' is 'a staff meeting between 2pm and 3pm'. The formalization of the activity is as follows:

$$\begin{aligned}\text{SynchronizedFixedTimeActivity+} = \{&\text{startTime}, \text{endTime}, \text{f(getStartTime)},\\ &\text{f(getEndTime)}, \text{f(getActivityMembers)}\}\end{aligned} \quad (4.7)$$

A 'FixedDurationActivity' belongs to the regular activity type, which directly defines the duration of the activity. A regular activity can be formalized as same as an abstract activity. A typical 'FixedDurationActivity' is 'play football for an hour'. It is formalized as:

$$\text{FixedDurationActivity+} = \{\text{duration}\} \quad (4.8)$$

A 'FlexibleDurationActivity' is another type of regular activity, which provides the estimated activity duration while the actual duration can be calculated during the simulation execution. A typical 'FlexibleDurationActivity' is 'go shopping for around two hours'.

$$\text{FlexibleDurationActivity+} \ = \ \{\text{estimatedDuration, f(getEstimatedDuration)}\} \tag{4.9}$$

A 'StochasticDurationActivity' is a special form of a 'FlexibleDurationActivity' where the actuation duration can be calculated by a duration distribution. It is formalized as:

$$\text{StochasticDurationActivity+} = \{\text{durationDistribution}\} \tag{4.10}$$

A 'UntilFixedTimeActivity' is formalized as:

$$\text{UntilFixedTimeActivity+} = \{\text{untilHour}\} \tag{4.11}$$

This formalization simply specifies the fixed end time of the activity, which can be realized for activities such as 'sleep till 7am'.

With these definitions of activities, one major novel design about agents can be concluded as the stateless activities in agents' behavior pattern. We show this design concept in Figure 4.5, where three agents share a same activity pattern. However, each agent will get different location and duration when executing the activities.

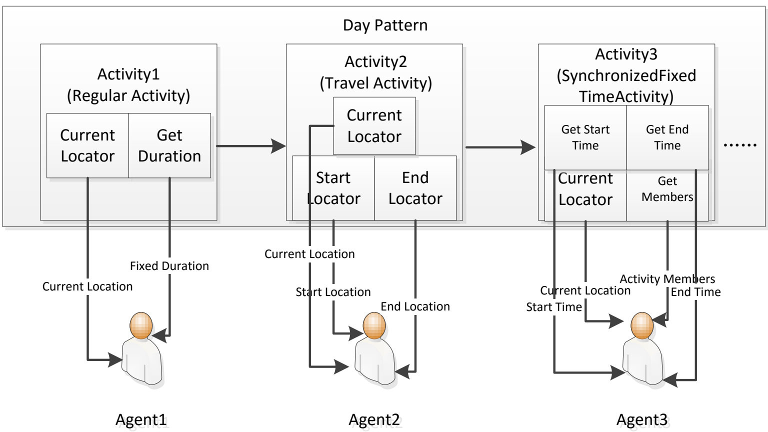


Figure 4.5: Runtime calculation of location and duration among agents with a same activity pattern

4.2.4 State Update Scheme

Agents change state, either spontaneously or through interactions. The update scheme deals with how frequent, and in what order, each agent updates its state (GUO and TAY 2008). In this research, we designed an implicit scheme for updating the agent's state by which agents can change state both spontaneously and through interactions.

Thus, the current explicit state (current activity) of an agent is set to be unknown for memory-consuming consideration. Different from keeping the current activity, a current index of the activity in an activity pattern for agents is recorded. When executing an activity, the activity itself or the activity executor (if the activity is a travel activity or social activity) will specify a duration for this agent to schedule its next activity. During this period, the agent remains in an implicit state which is shown in Figure 4.6.

Figure 4.6 gives one example day pattern for students who prefer taking public transportation (metro and bus) to school. There are 9 activities in this pattern. With the execution of these activities, agents will change their status in a cyclic way.

When executing regular activities, such as 'sleeping', agent will stay in a 'waiting' state till the end of the activity. This 'end' is informed by an 'end' event trigged when the 'sleeping' activity starts to be executed. After the 'end' event is received by the agent, the agent will start to execute the next activity in the pattern and change the state accordingly. When executing a travel activity or a social activity, the agent will stay in the state of 'transit' or 'suspended'. Both of the states are passive, which means the agent has authorized other objects to control its state temporarily (e.g. an activity group for a social activity and a travel executor for a travel activity). Take the social activity for example, when the activity group has detected that the social activity is finished, it will ask the agent to execute its next activity and change its state.

Much alike the agent life cycle in a FIPA agent (POSLAD 2007), an agent realized in this model has a implicit life cycle describing the agent states with the execution of activities (see Figure 4.7).

The difference between the life cycle of FIPA agents and agents in this model is how states are transited. Each FIPA agent keeps the exact current state in its life cycle and needs a specific transition instruction for updating to the next state. To achieve this, every agent should keep a list of future instructions which consumes a lot of memory. In our model, the current state of the agents is not clear as there are no explicitly defined states in the agents. Instead we keep a current activity index within the current activity pattern of an agent. When executing an activity, the activity itself or activity executor (if this activity is a travel activity or social activity) will specify a duration for this agent to schedule its next activity. During this period, the agent remains in an implicit state (e.g., suspended), which is shown in Figure 4.7. Based on this design, the activity pattern is reusable for agents who have the same social role, which considerably reduces memory usage compared

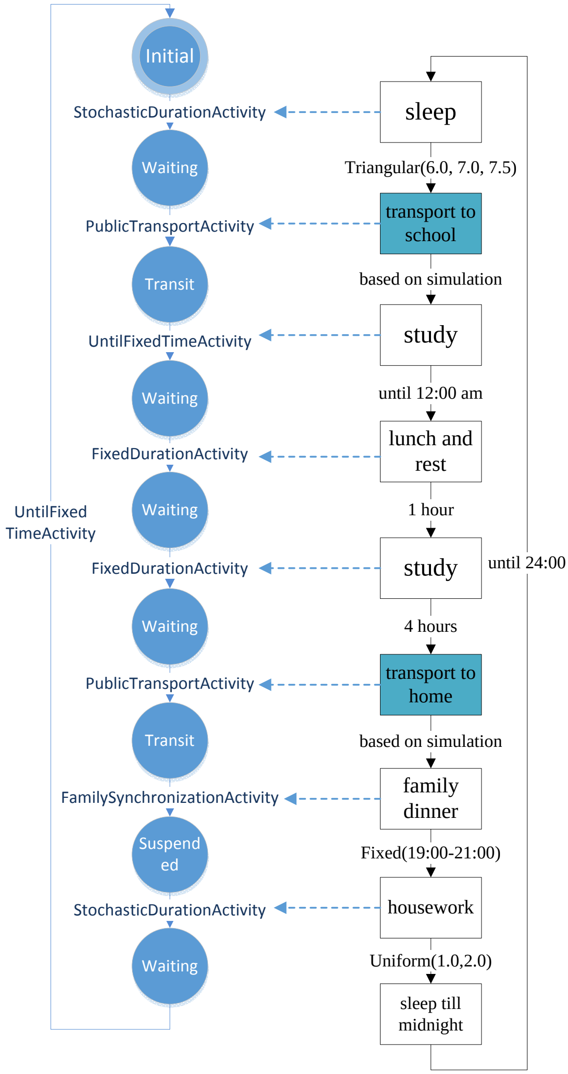


Figure 4.6: An example day pattern for students

to the FIPA solution. Take the same assumption mentioned above, assume there are around 20 million agents and each agent has 10 activities per day, then we can design 100 day patterns instead of the initial 5.6 billion activities for a 4 week simulation period, which are only around 1000 activities in total. Moreover, the week pattern of an agent in our model can be changed as a result of the state of the system (e.g. a policy intervention) as the week pattern is treated as an index attribute

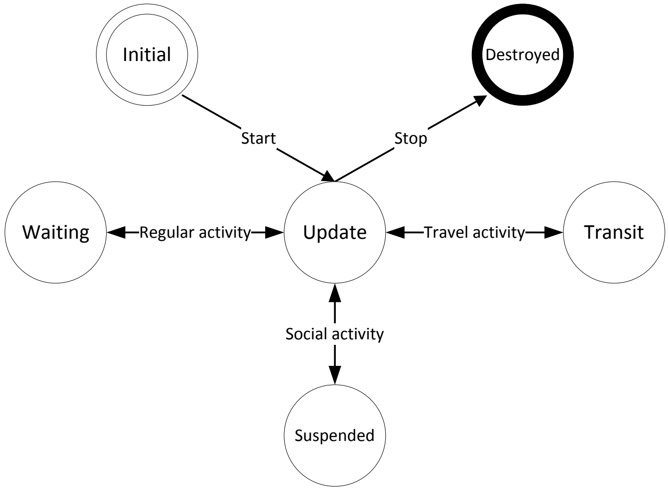


Figure 4.7: Agent life cycle

for an agent, which increases the flexibility of the model.

4.2.5 Multi-level Decision-making Capabilities

Agents realized by the proposed conceptual model are supposed to have three-level decision-making capabilities, which is mainly gained by the concept separation of agent environment into three parts.

Firstly, although the sequence of the activity list is predefined in the activity pattern for agents, agents can still adjust their preference for activity location and duration during the execution of the regular activities. This process is defined as operational level decision-making in this research, which is shown in Figure 4.8. Operational decision-making is mainly supported by physical containers.

Besides activity location, determination of activity duration is another aspect of operational level decision-making for agents. In the formalization of 'FlexibleDurationActivity' (see Formalization 4.9), the duration is defined to be flexible. However, this amount should be weighted carefully in some special activity patterns. Take the Figure 4.9 as example, the get-up time for a parent who needs to send the child to school in the morning should be calculated to meet the constraints that both himself/herself and the child can't be late for work/school.

From Figure 4.9, we can find that the parent has to get up around 7am although the duration of his/her 'sleep' activity is set to be flexible. The operational level decision-making module in this research is realized to solve this time constraint through a scanning-predicting-deciding procedure, which is described as follows:

1. Before an agent starts to execute a 'FlexibleDurationActivity', firstly it will

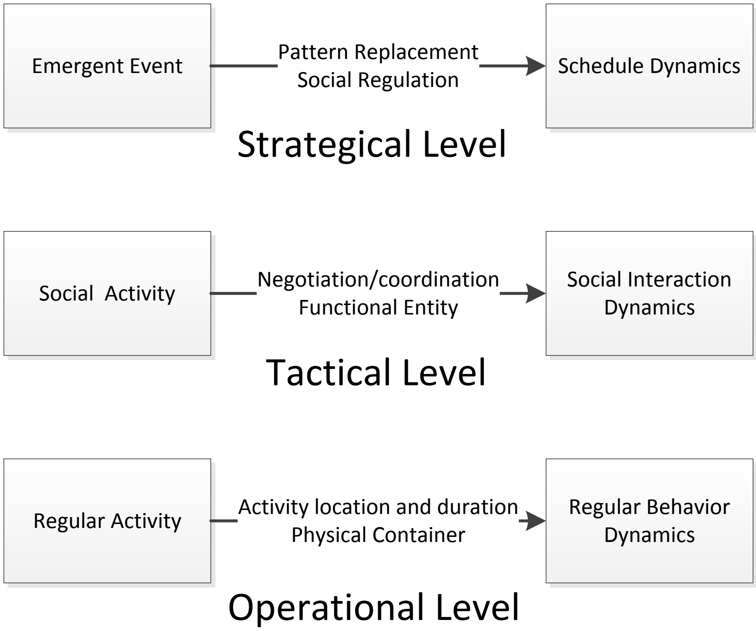


Figure 4.8: Agent's decision-making capability

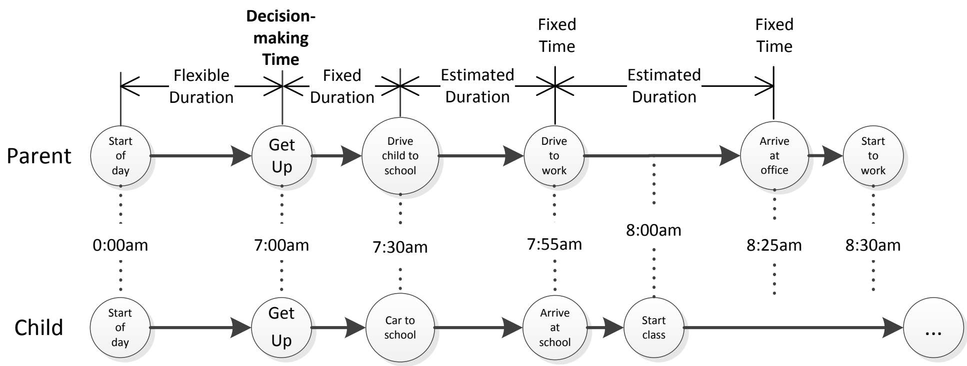


Figure 4.9: Decision-making for parent's getting up time(parent who sends child to school)

scan the future activities in the current day pattern and pick out the first activity with fixed start time or end time.

2. Secondly, the total duration required for executing all activities between current 'FlexibleDurationActivity' and the next activity with fixed start/end time can be predicted.

3. At last, the duration for the 'FlexibleDurationActivity' can be decided by a

mathematical calculation.

In addition, agents can decide to join in social activities or decline them when they receive social activity invitations which are not scheduled in their activity patterns. This is a capability of within-day re-planning to solve social interaction dynamics. When agents follow their day patterns, they can still receive social invitations from friends, which don't exist in the schedule. How agents respond to these invitations is defined as tactical level decision-making in this research, which will be explained in detail in the following sections. To realize this capability, agents need to coordinate with functional entities heavily.

Despite these, in order to increase the heterogeneity of agents, it is necessary to take into account also strategical level decision-making for agents. According to Figure 4.8, the strategical level focuses on the replacement of week pattern that an agent will follow when they break rules or encounter some emergency events, such as a mandatory policy intervention during an epidemic outbreak. This capability is achieved by the regulation of social regulations.

Using three levels of decision-making is not a new idea for agent-based modeling. However, it is very difficult to develop a comprehensive agent model including all the three levels of decision-making, yet we define these three levels of decision-making capability for agents to solve different levels of behavior dynamics problems when processing different types of events, as we believe the higher level, such as strategical level, can be progressively fulfilled by more situations and solutions in further research. How each of these dynamics problems are solved by the coordination between agent's decision-making modules and agent environment concepts is present in the following subsections.

4.3 Agent Environment

4.3.1 Separation of Concepts

In this research, the concept of agent environment is separated into physical container, social regulation and functional entity. The physical container is used to represent the physical environment where an agent stays. Physical containers are hierarchically organized and every agent has to stay in at least one physical container at any time. Social regulation is designed to model norms and institutions that can guide and influence human behavior globally. Functional entities are those extra artifacts in the system that can influence or directly change attributes of either agents, physical containers or social regulations.

This separation overcomes the limitations on environmental completeness in other ABM models and provides flexibilities in simulating different system scenarios. In other words, the concepts are also refined.

4.3.2 Physical Container

A physical container represents the physical environment where agents stay. Typical physical containers are such as school, classroom, office, bedroom, and train, etc. A physical container can be formalized as follows:

$$\text{PhysicalContainer} = \{\text{ID}, \text{Type}, \text{GISInfor}, <\text{List}>\text{PhysicalSubContainers}, \\ \text{area}, \text{status}, <\text{List}>\text{agents}\} \quad (4.12)$$

Each physical container is characterized by its *ID, type, GIS information* (longitude and latitude), the total *(*area) in square meters, its current *status* (e.g., full, closed), a list of *sub-physical containers* and *hosting agents*.

The *type* attribute is used to categorize physical containers in the simulated system. In an epidemic model, the types can be e.g., home, working places and hospitals. In a traffic model, the types can be e.g., private cars, buses, trains and road lanes. GIS information provides the accurate location of the physical container in the map.

The total *area* attribute to a physical container is unique in this definition and differs from other similar researches, which is used to generate physical sub-containers (e.g., classrooms in a school).

Physical containers are organized hierarchically. Each physical container can be partitioned into physical sub-containers by giving each physical container an attribute *list of physical sub-containers*. Examples are classrooms in a school, stores in a shopping mall, or offices in a working place. Agents can have different forms of contacts when they are in the different level of physical contain hierarchy.

Each physical container is assigned a variable *status*, which is used to specify its current availability for agents to enter or leave. All agents staying in the present physical container can also be retrieved by the variable *hosting agents*.

Besides the definition, physical containers show "behaviors" just like agents. An important "behavior" 'Calculate Distance' calculates the distance between two physical containers based on the GIS coordinate information (latitude and longitude). Since the process of calculating the distance between two physical containers is an indispensable step for many social systems, this "behavior" is most frequently shown during a simulation run.

Besides the default "behaviour", other frequently showed "behaviour" are such as 'Find Nearest Physical Container' and 'Find Physical Containers within certain Distance'. The first "behaviour" gives out the nearest physical container, and the second returns a list of physical containers within a max distance to any specified physical container. These two "behaviours" will be frequently asked by agents due to the fact that people are more willing to visit nearby places for certain activities, such as shopping, eating and leisure when they have no particular preference.

We will use an example algorithm to explain how 'Find Nearest Physical Container' works when an agent (worker) tries to find a nearest restaurant for dinner. This algorithm is shown in Algorithm 4.1.

Algorithm 4.1 Determine a nearest restaurant R for a worker

Require:
The start Physical Container L_S
Ensure:
The nearest restaurant R to the start Physical Container L_S
Calculate the key k_n of the L_S for the nearest restaurants cache map RCM_n;
if RCM_n contains key k_n **then**
  get the restaurant R from the RCM_n;
  return R;
else
  Calculate the key k_g of the L_S for the restaurants grid cache map RCM_g;
  get all the restaurants R_g in the same grid with L_S from the RCM_g;
  if R_g not empty **then**
    min Distance D_m = Double.MAX_VALUE;
    for all $L \in R_g$ **do**
      Calculate the distance D_L between L and L_S;
      if $D_L < D_m$ **then**
        $R = L$;
        $D_m = D_L$;
      end if
    end for
    add R in the nearest restaurants cache map RCM_n;
    return R;
  else
    get all the restaurants R_d within a certain distance(e.g. 1km) to L_S in the restaurants distance cache map RCM_d;
    if R_d is empty **then**
      get all the restaurants R_d in the map;
    end if
    min Distance D_m = Double.MAX_VALUE;
    for all $L \in R_d$ **do**
      Calculate the distance D_L between L and L_S;
      if $D_L < D_m$ **then**
        $R = L$;
        $D_m = D_L$;
      end if
    end for
    add R in the nearest restaurants cache map RCM_n;
    return R;
  end if
end if

In the above example algorithm, there is a supporting caching mechanism which is designed for large-scale systems. This three-level cache mechanism is creatively designed to achieve balance between cpu utilization and memory usage.

- The first cache is the nearest cache, which stores the nearest physical container of the current physical container type to a certain physical container. New items will be added into this cache only after they have been calculated for a first time.
- The second cache is the grid cache. The whole map can be divided into grids and keep indexes of physical containers in the grids (similar to Quadtree).
- The third cache is the distance cache, which is used when no results can be found in the nearest cache or the grid cache. To any specific physical container, this cache can keep nearby physical containers ordered by distance.

Based on this design, the other "behaviour" 'Find Physical Containers within certain Distance' is listed in algorithm 4.2.

Algorithm 4.2 Get Location Array within Max Distance

Require:
Start location SL; Max Distance D;
Ensure:
Location array L_d within D to SL
Calculate key k_d of SL for distance cache map M_d;
if M_d contains key k_d **then**
    get location array L_d from M_d;
    return L_d;
else
    Calculate all grids G_s within distance D to SL;
    for all $G \in G_s$ **do**
        Calculate key k_g for each G;
        get all locations L_g in G from grid cache map M_g;
        for all $L \in L_g$ **do**
            Calculate distance D_L between L and SL;
            if $D_L <= D$ **then**
                add L into location array L_d;
            end if
        end for
    end for
end if
return L_d;

Besides an effective mechanism to organize physical containers in large-scale systems, the concept of physical container separating from the general agent environment concept makes it much easier to include a transportation component in a social simulation model. This is achieved by considering vehicles as movable physical containers in the model.

4.3.2.1 Movable Physical Container

Human mobility and, in particular, commuting patterns have a fundamental role in understanding social systems (GARGIULO et al. 2012), such as epidemics. In current agent-based epidemiology research, agent environment represented by physical rooms is a key component as the majority of transmission of contagious diseases is thought to occur among sustained indoor contacts (ANDREWS et al. 2013), such as students taking classes in classrooms and workers working in offices. However, direct physical contact (e.g. touching) or vector-borne contact (e.g. a droplet) outside rooms can be an effective method for disease spread as well, especially in densely populated areas such as during transportation. Thus, it is also important and necessary to include a transportation component in a social simulation model.

KOPMAN et al. (2012) present a basis for a mobile epidemic simulation framework by creating a simulation model of individuals walking in a defined space. MEI et al. (2015) described an approach to study the spread of airborne diseases in cities by combining traffic information with geo-spatial data, infection dynamics and spreading characteristics. However, none of these tryouts was able to cover the fundamental structures of the route network and to run basic microscopic traffic simulations successfully in large-scale social systems which are well studied in large-scale traffic engineering areas (FELDKAMP and STRASSBURGER 2014).

There are many papers on large-scale agent-based transportation simulation (see e.g., (RANEY and NAGEL 2003, NAGEL and RICKERT 2001, ZHANG et al. 2013)). These papers mainly focus on the prediction of traffic peaks and congestion. In order to make this research applicable in other social systems, we introduced the movable physical container concept in the large-scale agent-based conceptual social simulation model. To make this concept practical, an implementation of a microscopic public transport system (subways and buses) is present.

The public transportation system is closely associated with the execution of travel activities, which are considered as a connection between two activities of agents in two different physical containers. An agent that has to commute by public transport between two physical containers to conduct its next activity, is required to authorize the transportation system to execute a travel activity, which helps this commuting agent to determine a route and calculate the travel duration. For example, the agent in Figure 4.6 will ask the public transportation system to take him from household to school when executing the activity 'transport to school'.

The public transportation component is microscopic as we can model all lines and stops for the metro and bus system. In each simulation day, modeled buses and metro trains will execute their schedules on these routes based on timetables.

The geographic information and routing data of the transportation infrastructure network can be acquired from maps such as OpenStreetMap. The map offers stops as nodes and routes as links which can be connected in a graph. This graph shows the topology of the whole public transportation network.

The metro stops and bus stops of the public transportation system are modeled as movable physical containers which are extended from the general physical containers. In addition to the behavior of a physical container, a bus/metro stop can 'move' the waiting agent from the current stop to the arriving transporter (bus/-metro train) when this transporter has enough space and is on the right route for the waiting agent in the stop. Moreover, in order to keep the agents 'simple' enough for large-scale simulation while 'heterogeneous' enough for public transportation, only the stops know and record transfer information of the waiting agents, and will pass the information to the transporter when the agents are on board. Then the transporter will 'move' the agent from the bus to a stop when it arrived at the right transfer or destination stop.

Agent transfer in and between stops causes realistic delays, while the transporter also takes an amount of delay to 'move' agents out and accept new passengers when arriving at a stop. In order to be realistic, the bus or metro train can be also enabled to operate through a timetable. This data driven method enables this public transportation model to be able to simulate people's real travel behavior.

The modeled traffic infrastructure components can't offer routing information. Thus, a graph for routing should be constructed to connect all bus/metro stops. All successive stops of the same bus/metro line will be linked and the edge of the link will be assigned a travel duration. Stops that are not on the same route but within walkable distance should also be linked, and assigned a estimated walking duration on this edge of the link. This graph can offer a shortest (in travel duration) path to a potential public transport user. Since this graph will be called many times per simulation run, a cache can be added in each node (stop) to store the next transfer stop information with the destination node as the key in the cache. The structure of this public transportation system is shown in Figure 4.10.

However, there is a big challenge for an agent to use this graph to get a travel route, which is to find the first stop to use as there could be more than one public transport stop close to the agent. An explicit solution is comparing all the nearby stops for every travel request. This could decrease the simulation performance drastically. This challenge can be solved by creating '*GridZones*' as nodes and adding them to the existing graph. The map is divided into grid cells, and the resolution of the grid can be set flexibly. The center of each cell can be called '*GridZone*'. Each '*GridZone*' is a node and is linked to the graph by linking the '*GridZone*' with all stops in this grid cell. The weight of each edge will be assigned an estimated walking duration. When an agent plans to use public transport, the public transportation model will use the agent's current '*GridZone*' as the start node to calculate the shortest path. The destination is treated in a similar manner. The details are shown in Figure 4.11.

Besides public transportation, an agent can also choose to commute by his

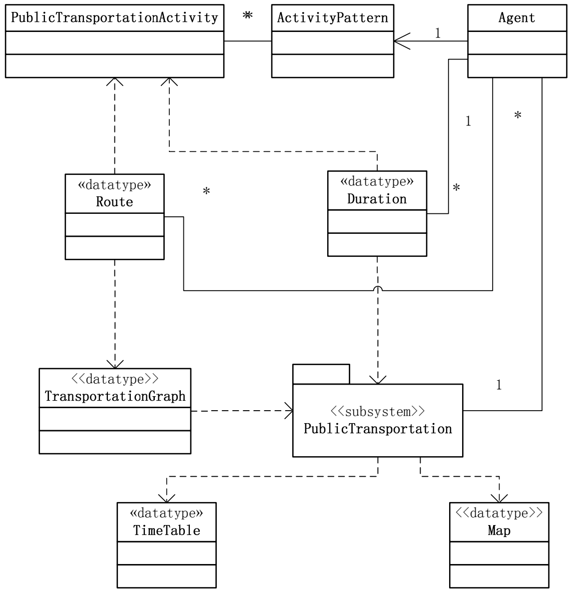


Figure 4.10: The structure of the public transportation component

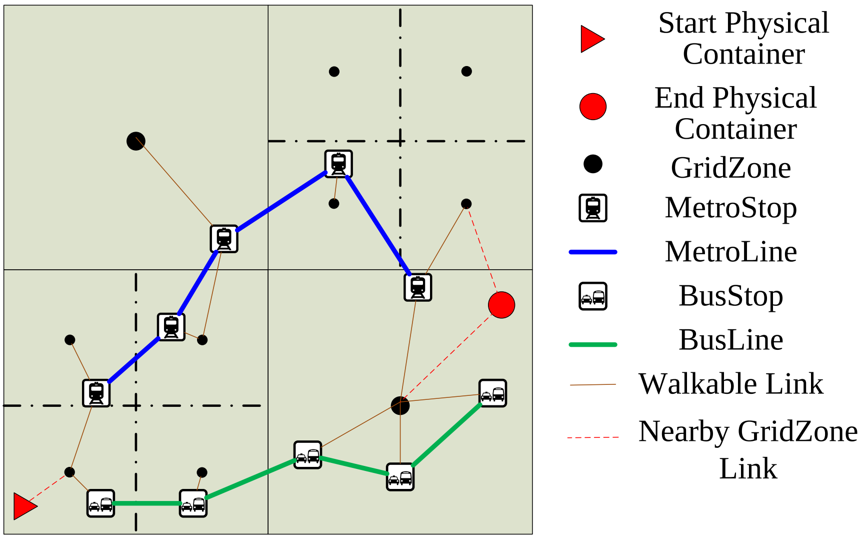


Figure 4.11: Part of the graph for public transportation using '*GridZones*'

or her own private car or taxis. When the physical container of an agent's next activity is within walkable distance, a travel activity 'walk' is conducted. This enables people to meet others when walking. Thus a 'walk' physical container can be modeled with a large area into which walking agents and cars will be put.

4.3.3 Social Regulation

In a real human society, social regulations protect public interests through restricting behaviors that directly threaten public health, safety welfare or well being (SALAMON 2002). Particularly in health-care areas, understanding the complex interplay between human disease and social environment is attracting a lot of interests (BARABÁSI et al. 2011).

In ABMS literature, there are already some researchers trying to model part of the function to simulate social regulations. An example is DIVAs, where users can create events that influence the environment at run-time, and the environment is able to enforce rules and constraints on the agents using the influence combination function (STEEL et al. 2010). OKUYAMA et al. (2009) presented an approach to integrate the modeling of environments and organizations, using a normative infrastructure that is composed of normative objects and normative places to distribute normative information over an environment. However, both of the approaches have no clear way to relate organizational and normative structures to the model of the environment where they are to be situated and operate.

On the other hand, the typical approach in multi-agent systems is that social regulation can be viewed and used at the level of individuals through the definition of entities such as norms and institutions (HESAN et al. 2015). As a matter of fact, the social regulation concept is separated from the general agent environment concept in this research, which makes large-scale normative agent-based social simulation possible.

A social regulation can be formalized as follows:

$$\text{SocialRegulation} = \{\text{ID}, \text{Type}, \text{Monitor}, \text{Standards}, \text{Operations}, <\text{List}>\text{Agent}\} \tag{4.13}$$

In this formalization, the field *Type* categorize the social regulations. A coarse-grained categorization on *Type* can be "sanction" and "enforcement". *Monitor* is used to observe agents' behavior and status, analyze the results and compare with *Standards*. Based on the comparison, social regulations can trigger various *Operations* to the agent society in order to regulate agents' behavior.

One of the referenced implementation on *(*Operations) in a large-scale agent-based social simulation is to switch agents' activity patterns when the objective agents don't comply with any of the standards. With this process, agents can respond to different situations during a simulation run. For example, regulating agents' behavior during a disease outbreak is an indispensable part to model large-scale agent-based epidemic predictions. How agents would respond to a disease outbreak is a lightweight strategical level decision-making process as it would have a big impact on the agent's behavior.

Figure 4.12 presents the coordination process between the strategical level decision-making module of agents and a social regulation.

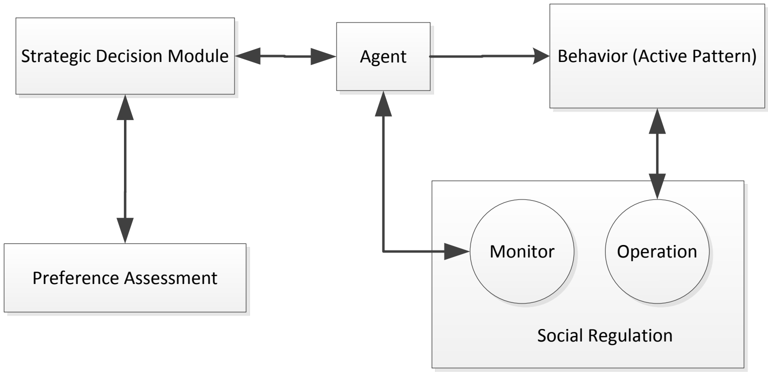


Figure 4.12: Elements involved in strategic decision-making procedure

How the process in Figure 4.12 works is explained in the procedure below, which presents the details.

1. A monitor is used by a social regulation to observe agents' states or system state, collecting results and compare the results with predefined standards.

2. When an agent's or the whole system's status is not compliant with any of the standards, the agent will be notified and the process is triggered using two steps.

3. The first step takes place in agent's decision-making module before social regulation triggers an operation, which makes a preference assessment to calculate a preference possibility indicating whether the agent would change behavior by shifting to another activity pattern.

Take an example of an agent's response during disease outbreak, where the preference assessment may work based on the probability $P(A)$ on Equation 4.14, where A_p represents the importance degree of the current activity pattern to the agent A, A_s represents the severity of the agent's new healthy status, A_a is the age of the agent, ϵ is the average infected age among people with the disease and λ is the weight coefficient. The closer to 1 the calculation result is, the higher possibility that the agent would shift to a temporary activity pattern.

$$P(A) = 1 - e^{-\lambda \cdot A_p \cdot A_s \cdot |A_a - \epsilon|} \tag{4.14}$$

4. The second step takes place after social regulation triggers an operation, which means the agent has to decide if it will comply with the regulation. The decision-making module will ask the preference assessment module to give out a compliance preference possibility.

Take the same example as above, this assessment may be based on the probability $P(A)$ on Equation 4.15. The added parameter compared to Equation 4.14 is P_c, which represents the mandatory degree of the regulation.

$$P(A) = 1 - e^{-\lambda \cdot A_p \cdot A_s \cdot P_c \cdot |A_a - \epsilon|} \tag{4.15}$$

5. Two results are possible after an operation of a social regulation is fired, the agent changes its behavior or the agent follows the original one.

4.3.4 Functional Entity

Besides agents representing human beings, many other entities are relevant in agent-based social simulation models as well. In epidemic models, researchers showed that climatic conditions greatly influence vaccine effectiveness (ESTRADA-PEÑA et al. 2014), temperature influences the dynamic of epidemic significantly (GUO et al. 2015), climate affects malaria transmission (BOMBLIES and ELTAHIR 2009), and influenza virus transmissibility differs at various temperature and humidity conditions (ZUK et al. 2009). The common way is to model these as a built-in components or properties in agent environment. This will limit the scalability and performance of large-scale models.

Thus, we separate these entities from the agent environment concept and define that functional entities are those extra objects in the system that can influence or directly change attributes of either agents, physical containers or social regulations. For example, a disease is modeled as a functional entity to change agents' healthy status. A storm is modeled as a functional entity to change the temperature (functional entity) of a room (physical container).

Although the functionalities of the functional entities are diverse, it can be formalized as follows:

$$\text{FunctionalEntity} = \{\text{ID}, \text{Type}, \text{Status}, \text{Targets}, \text{Operations}\} \tag{4.16}$$

In this definition, *Type* can be "Agent-related", "Physical Container-related" or "Social Regulation-related" where the classification criteria focus on the targets. *Targets* are the entities being affected by the *Operations* conducted by functional entities with changing *Status* in the simulation models.

4.4 Social Networks

A social network is a social structure made up of a set of social actors (such as individuals or organizations) and a set of the dyadic ties between these actors. Social networks and the analysis of them is an inherently interdisciplinary academic field which emerged from social psychology, sociology, statistics, and graph theory. For social scientists, the theory of social networks has been very rewarding,

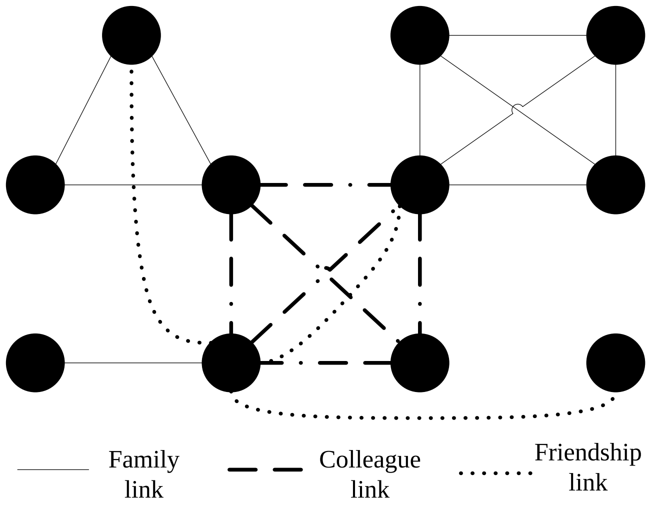


Figure 4.13: Types of social networks

yielding explanations for social phenomena in a broad range of disciplines from psychology to economics (BORGATTI et al. 2009).

4.4.1 Generating Social Networks

Social networks usually contain three types of relations: family, colleagues/ classmates, and friendships, which are shown in Figure 4.13.

Family, colleagues and classmates relations can easily arise from defining a complete topology that clearly specifies all relation connections. Friendships, as the most complex social relation, are relatively difficult to define. The topology of friend connections changes over time due to the dynamics of friendship relations (PUJOL and FLACHE 2005). This is even more complicated on a large scale (GATTI et al. 2014).

WANG and COLLINS (2014) argues that, in social networks, people's intention to connect is not only affected by popularity, but also strongly affected by the extent of similarity. The authors propose that in forming social networks, agents are constantly balancing between instrumental and intrinsic preferences. SUTCLIFFE et al. (2012) described a computational model for the development of social relationships based on agents' strategies for social interaction that favor more less-intense, or fewer more-intense partners. This model does not account for spatial effects, which have been modeled by BELTRAN et al. (2006) who demonstrated that groups can emerge from interactions within a lattice topology as a consequence of agents' preferences to maintain personal or social space. The MOCA platform presents a theoretical model that allows agents to dynamically choose to create, join or quit social networks based on their individual and collective recurrent patterns of behavior, called respectively roles and organizations (AMIGUET et al. 2002).

Based on this research, this thesis uses the concept of social reach in the social circle model and the aggregation utility between two agents (NAVARRO et al. 2011, HAMILL and GILBERT 2010), and proposes the concept of 'social similarity' to dynamically generate a special type of social networks-friendship. Then, egocentric friend networks are generated to represent friendship connections dynamically. In this thesis, friendships will be generated before planning and negotiating social activities based on an algorithm that we will present below. The candidates for the friends come from three kinds of sources: neighbors, classmates/colleagues and a random selection. When agent A is planning a social activity, the algorithm for generating friends can be described as follow:

First, the number of friends N_s is assigned to A which follows a power-law distribution (HAMILL and GILBERT 2010).

Second, the percentage of A's friends from different sources is calculated according to a combination of uniform distributions (see Table 4.3) as the source composition of A's friends may differ from another agent. For example, agent A may like to make friends with neighbors while agent B may prefer making new friends randomly in places like shops or restaurants.

Table 4.3: Distribution of agent's friends

Item	Number
Total number	N_s
Number of friends from neighbors N_n	$Uniform[0, N_s]$
Number of friends from classmates/colleagues N_c	$Uniform[0, N_s - N_n]$
Number of friends from random selection N_r	$N_s - N_n - N_c$

Third, select one candidate randomly from the source and calculate the possibility that the candidate and agent A are friends. If the calculation result exceeds a predefined threshold (e.g., 0.25 as an initial setting), put the candidate in agent A's friends list. Otherwise, select a new candidate and repeat the calculation process till all A's friends are generated. If the new friends list is still not full, increase the threshold and repeat the calculation process again. The calculation process is based on a concept called 'social similarity', which is proposed in this research. It calculates the similarity between two agents. The considered variables include age, social role (week pattern), family role and the number of friends. In this research, the similarity $S(A, B)$ between two agents A and B is evaluated by a weighted Euclidean distance which is shown is in Equation 4.17, where a represents age, s represents social role (converted to an index), f represents family role, n represents the agent's friends size and μ represents the weights for different variables. The result of $S(A, B)$ is bounded on [0,1]

$$S(A,B) = 1 - \sqrt{\sum_{i=a,s,f,n} \mu_i (A_i - B_i)^2} \tag{4.17}$$

4.4.2 Modeling Process of Social Interactions

The mechanism for implementing social interactions can be separated into two parts, social learning and social influence (MONTGOMERY and CASTERLINE 1996). Social learning is considered as the process to make decisions by taking external information, while in social influence process individuals always try to avoid conflicts. MARSELLA et al. (2004) claimed that a key factor in human social interaction is the beliefs about others. Whether we believe a message depends not only on its content but also on our model of the communicator.

Based on these arguments, the interaction mechanism in this research contains two stages, as well. The first stage is equation based, by which agents calculate the possibility to join in proposed social activities according to their own preference (belief of others and input messages). The second stage is decision-tree based, by which agents make decisions according to rules (social norms, sanctions) and try to avoid conflicts with their own schedules. The decision can be a full agreement or a decline.

The challenges for implementing these two stages are extensive. The first is that no friendship social network can be predefined in the initial data. All friendship social networks should be generated before the execution process of friendship social activities based on the algorithms described above. For example, part of the friendship relations of agents are generated among his/her neighbors and colleagues. The reasons that we choose to dynamically generate friendship social networks for the agents are twofold: first, it's too memory-consuming to store all friends lists for all agents (up to 100 friends for each agent), and second, the real human friendship social networks are dynamic and evolve over time, which influence a lot of social phenomena. To make this friendship relation generated by the stochastic method as stable as possible (most friends of an agent still remain the same over time), a reproducible random generator was designed using the agent id as the seed. Hence, every time when agents want to invite his/her friends to conduct a social activity in the simulation, the dynamically generated friendship relations will mostly remain the same although no static friends list are predefined, or need to be stored. The slight difference comes from the sequence of selecting candidates for friendship calculation from friends sources, which is on a first come, first served basis.

Another challenge is the consequences of the first challenge that the joint social activities are not pre-scheduled for all participants and only the organizer agent of the joint social activity foresees this activity in its schedule. Because there are no predefined friendship social networks, it's impossible to assign two consistent and semantically matched week patterns to two individual agents before the simulation starts while the two agents are modeled dynamically as friends during the simulation. This is solved through dynamically generating Functional Entities, Group Agents, to help execute the friendship social activities. When the originator/organizer agent tries to execute a social activity, a helping Group Agent is dynamically generated to take over the task to execute the social activity. First it will generate a social network and then invite the members in the network to attend this joint social activity. After a decision tree considering several rules and conditions (for example time and distance), each invitee can either decline or accept the invitation. After collecting all the response, the Group Agent will request all the participants to travel to the social location where agents can be late due to real travel delay which is caused by the transportation model. The major process of executing a social activity is presented in Figure 4.14.

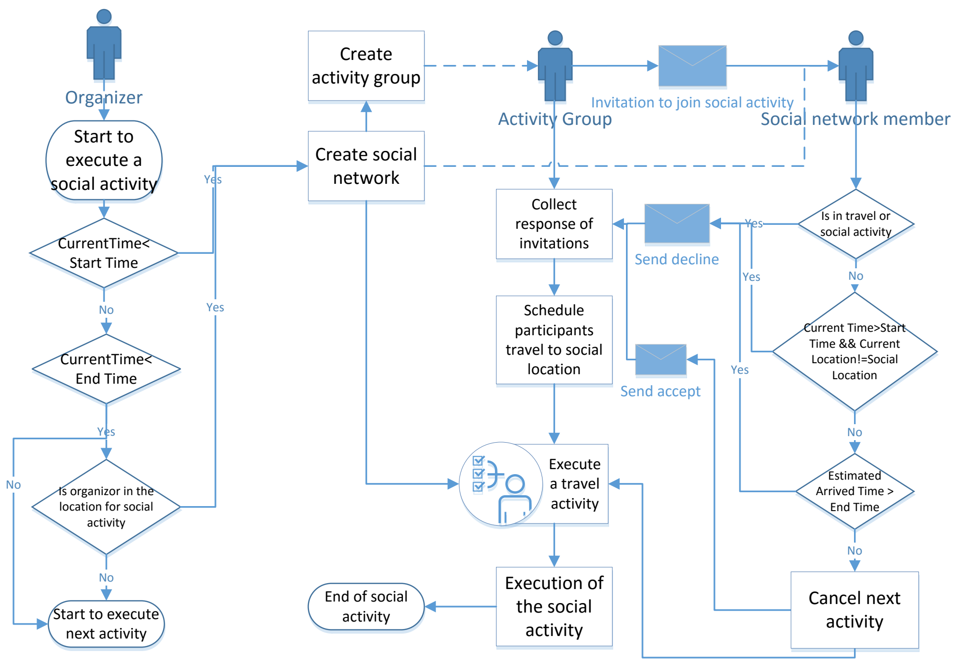


Figure 4.14: Execution process of a social activity

The detailed interaction procedure can be described as follows:

1. Before an agent starts to execute the current activity in the activity pattern, it will check the next activity to see if it's a joint social activity. If yes, check if the conditions are met for organizing it. Then a proposal of the joint social activity will be sent to all involved social networks members. It is worth noting that the friendship relations in social networks will only be generated in this step and the agent will only schedule a social activity within its current pattern.

2. Calculate the attendance possibility after receiving a social activity proposal for every agent I_i according to Equation 4.18, where N is the total number of agents involved in the planned social activity, I_o is the organizer of this activity, $S(I_i, I_j)$ calculates the link weight between the two agents based on a concept 'social similarity'. The considered variables include age a, social role s, family role f and the number of friends n. In this research, the social similarity is calculated as a weighted Euclidean distance, where μ represents the weight for different variables. By setting the weight coefficient $\{\mu_a, \mu_s, \mu_f, \mu_n\}$, the calculation result $S(I_i, I_j)$ will be constrained between 0 and 1. 1 means they are fully connected while 0 means no relations. $A(d, E)$ calculates the interest degree of the activity to the agent, where d is the distance between agent's current location and proposed activity location, E gives out the degree that the agent is interested in the activity and σ is a corrective coefficient for calibration.

$$
\begin{aligned}
P(i, o, N) &= e^{\sum_{j=1, j \neq o}^{N} S(I_i, I_j) - N} \times A(d, E) \\
S(I_i, I_j) &= 1 - \sqrt{\sum_{x=a,s,f,n} \mu_x (I_{i_x} - I_{j_x})^2} \\
A(d, E) &= \frac{\sigma \cdot E}{d}
\end{aligned}
\tag{4.18}
$$

3. For each agent, compare the attendance possibility with its own attendance threshold t. If it's negative, send a decline response to the activity organizer and continue its own schedule. Otherwise, start the second stage process for decision-making based on a decision tree (see Figure 4.15).

4. Two kinds of decisions can be made by the agents after the decision-tree based process, which are accept and decline. The decisions will be responded to the organizer immediately, and the organizer will make a decision on continuing the activity after collecting all responses.

5. Social activity organizers will only negotiate with other members for one time, which is necessary to avoid deadlocks.

6. When the final decision is made, the agents who are willing to join in the coming social activity will authorize a dynamically generated Functional Entity, 'Group Agent', to take the responsibility for state updating and moving agents back to their original schedule when the social activity is finished.

For social contacts among family members and colleagues, the execution process of their joint social activities is almost the same as the process in Figure 4.14. However, the difference with the friendship social contacts is that the social networks for family members and colleagues can be pre-defined in the initial data.

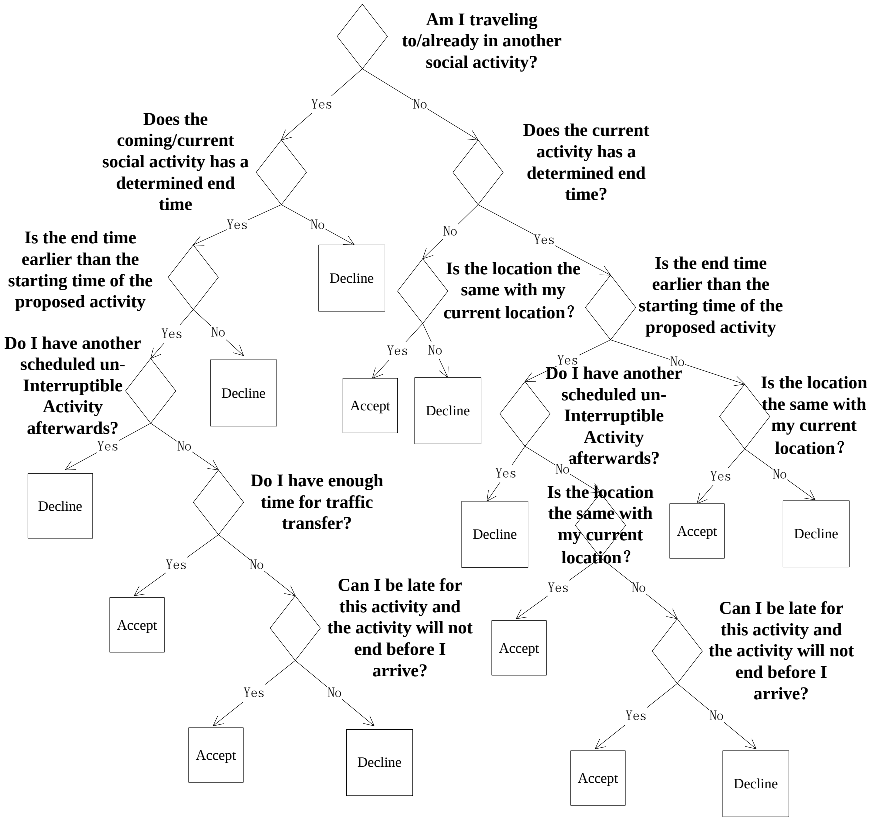


Figure 4.15: Decision tree for joining in social activities

4.5 Conclusion

This chapter introduces a new conceptual framework to build large-scale agent-based social simulation models, which separates the concept of agent environment into a physical container, a social regulation and a functional entity. Compared to the model concepts in the general ABM conceptual model, this separation overcomes the limitations on environmental completeness in other ABM models and provides flexibilities in simulating different system scenarios. In other words, the concepts are also more refined. For example, physical containers can be movable to represent transportation vehicles which are difficult to implement in general ABM platforms. Moreover, theories and concepts on social regulation from artificial intelligence can be easily implemented and integrated into an agent-based model while showing reasonable performance. Inspired by the concept of Functional Model in object-oriented conceptual models, the concept of functional entity is borrowed from the classical ABM conceptual model and introduced in our proposed conceptual model for large-scale ABSS. With this clear separation of concepts which is much easier for implementation using the object-oriented paradigm, experimental results in the following chapters show that models adopting this conceptual model are more efficient in terms of system performance and low agent-to-agent communication cost than a pure ABS model while keeping high model fidelity and the same agent capabilities.

With this chapter, research requirement RR1 is satisfied.

✓**RR1 Research Requirement - Model Architecture** A conceptual model is required for large-scale ABSS.

5

Case: Large-scale Agent-based Epidemic Prediction and Control

In order to test the proposed conceptual model architecture, we applied the large-scale agent-based social conceptual model in a large-scale epidemic model in this chapter. We constructed a large-scale artificial city Beijing with 19.6 million population and 8 million geo-referenced locations (like households, schools, offices, hospitals, stations, etc.), in which 200 million activities are executed per simulation day. In addition to regular contacts in locations like home and schools, agents can interact with each other based on the dynamic formation of social networks. Furthermore, we include a microscopic public transportation system in this city model to implement random travel contacts.

5.1 Model Preparation

As we described in Chapter 3, many large-scale agent-based epidemic prediction models have been developed in recent years due to the increasing threat from multiple epidemics. However, most of the models are in the context of the megacities or metropolitan areas either in the USA (STROUD and VALLE 2007, BISSET et al. 2009a, BISSET et al. 2014, BARRETT et al. 2008, GREFENSTETTE et al. 2013, MAO 2014) or in Europe (AJELLI et al. 2010, RAKOWSKI et al. 2010). None of these mentioned popular models have been applied in other areas, such as megacities in Asia.

On the other hand, there is a strong need to study epidemic models in the megacities in Asia. As of 2015, there are 35 megacities in existence in the world and 22 cities are in Asia of which 8 are in the top 10 [1].

As far as we can see, the scarcity of large-scale agent-based epidemic models in Asian megacities is caused by many difficulties. One intuitive difficulty is the lack of open data in demographics in Asian cities, such as population and environment. Large-scale real world data sets are expensive to collect and difficult to obtain high fidelity ground truth for (BERNSTEIN and O'BRIEN 2013). Thus, there is a trilemma of inadequate data from real-world datasets, statistical simulation models, and agent-based simulation models. This difficulty is reflected in the research in Beijing by the model of the spread of SARS in Beijing conducted by HUANG (2010).

Another difficulty is in applying the existing conceptual models for Asia megacities due to the differences in people's commuting patterns (VAN DE COEVERING and SCHWANEN 2006, BAUMAN et al. 2011) and social behavior (TRIANDIS 1989). Taking public transportation usage among workers as an example, the average rate of public transportation usage among workers in USA among a handful of the nation's large and densely populated regions was 5 percent from the 2009 ACS (MCKENZIE and RAPINO 2011), while Asian cities dominate the ranking of the world's biggest and busiest metro systems [2] (e.g., 40% workers use public transportation in Beijing [3]).

Beijing, as the capital of the People's Republic of China and the second largest Chinese city by urban population, is selected as the context of this case study. The population as of 2009 was 19.7 million.

The reason to choose Beijing as the context of the case study mainly comes from the availability of the raw data generated by an independent research by GE et al. (2014) for constructing an artificial city of Beijing. They adopted a mixing method which collect real data (statistical data and geographic information) and generate the other minimum required data by algorithms, which are the synthetic population and physical locations by utilizing the real data. More detailed information about the raw data on synthetic population and physical locations are as follows:

- The statistical population and location data were collected from the National Bureau of Statistics (NBS) at the city scale, and from the Municipal Bureau of Statistics (MBS) at the district scale, which include population, age-sex dis-

[1] Wikipedia, s.v. "Megacity", last modified on 27 April 2016, `https://en.wikipedia.org/wiki/Megacity`.

[2] Retrieved from `http://www.uitp.org/news/metro-ridership-ranking`

[3] Retrieved from `http://www.wtoutiao.com/p/Eb1ThP.html`

tribution, number of children distribution among families, family size distribution and geographic distribution of families among districts.

- With the algorithms in GE et al. (ibid.), each individual person is specified with the attributes of age, gender, family role, family index and social role to specify this individual's demographic characteristics. The family role can be defined as a set {grandparent, parent, child}. The social role is defined as a set {infant, student, worker, retired}. This design is based on findings from the China census data (available at http://www.stats.gov.cn) that households with more than three generations are a small proportion (less than 10%) of the total number of households.
- Besides individual persons, physical locations were generated where individuals can perform a variety of activities. Currently, there are 18 location types, and these location types are classified into 6 categories: houses, educational institutions, workplaces, consumption locations, entertainment locations, and medical institutions. Each location has a geographic reference and the distribution of these locations was generated according to both statistical data and the geographic distribution of the population (ibid.).
- The consistency between the individual person and the physical location was guaranteed. For example, a student of age 22 will be assigned a location which belongs to location type 'university' rather than 'primary school'.

The statistics of the synthetic population and physical locations are listed in Table 5.1.

Item	Description	Results
Population	Number of agents	19611800
Age	Scope of age	0-105
Location	Number of physical locations	8216011
Families	Number of families	8055324

Table 5.1: The statistics of the synthetic population and physical locations

The statistical results of the generated synthetic population are shown in Figure 5.1 in the form of an age distribution. According to the previous results, the standard deviation of errors between the generated age and the statistical data is 0.9823 (95% confidence interval (CI) from 0.7034 to 1.3510) (ibid.).

With the generated data, GE et al. (ibid.) constructed a large-scale agent-based virtual city model. Based on the same source of data, this research built a large-scale agent-based artificial city model for epidemic prediction in a new way. A key issue and challenge of utilizing the raw data to our model is the redundancy of the data, such as the agents' preferred location list for shopping, eating and entertainment. Together with the predefined social networks for agents in the data, the size of the data is initially around 130 Gb. Since the large-scale agent-based model implemented in this research is entirely different from GE et al. (ibid.)'s method

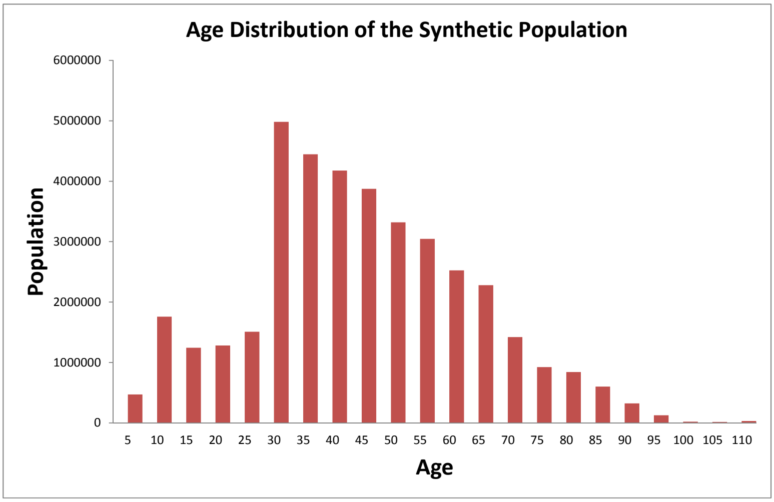


Figure 5.1: Age distribution of the Synthetic Population

which doesn't require the predefined location choices and social networks, we post-processed the raw data by extracting only the valuable fields of data items from the original database to fit our requirements. In addition, to speed up the initialization phase, we converted the data items to a compressed format (e.g., gzip) to reduce disk transfer time. With these post-processing steps, the time efficiency for loading the model could be improved by 65% in our case.

5.2 Model Implementation

5.2.1 System Architecture

There are many existing large-scale agent-based epidemic system architectures, such as FRED, FLUed and DISimS. FRED (A Framework for Reconstructing Epidemic Dynamics) is an open-source software system for modeling infectious diseases and control strategies using census-based populations (GREFENSTETTE et al. 2013). FLUed (A Four-Layer Model for Simulating Epidemic Dynamics) is a model that integrates complex daily commuting network data into multiple age-structured compartmental models for simulating the epidemic dynamics of emerging infectious diseases, assessing the potential efficacies of various intervention policies, and identifying the potential impacts of spatial-temporal epidemic trends on specific populations (HUANG et al. 2013). DISimS (Distributed Interactive Simulation System) is a flexible epidemiological modeling environment that combines high-resolution individual-based epidemic and intervention modeling environment with a web-based user-friendly analytics environment (BISSET and

MARATHE 2015).

As we stated in Chapter 1, all these large-scale agent-based epidemic models have gained reasonable performance while model fidelity is weaker considered in term of modeling agent behavior and a realistic set of contacts.

Based on the proposed conceptual model, the system architecture of the epidemic model of Beijing is shown in Figure 5.2. Components of physical containers, agents, social networks, functional entity and social regulation constitute the main parts of the model.

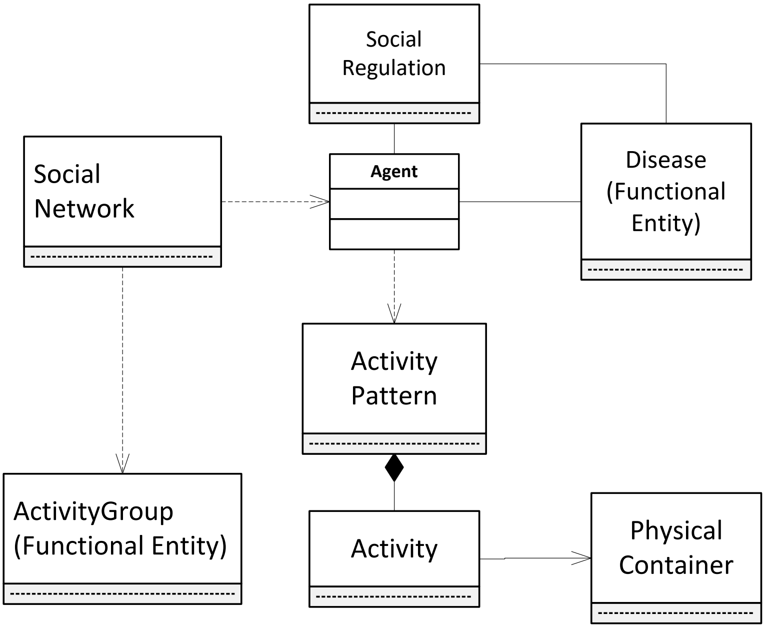


Figure 5.2: System architecture of the epidemic model of Beijing

Based on this system architecture, we built a model to study epidemic dynamics and effect of interventions based on the main class diagram in Figure 5.3. More details of this implementation are illustrated in the following subsections.

5.2.2 Agent-based Modeling

The typical way of implementing the daily behavior of agents in artificial cities is initializing an activity list (schedule) for each agent, which initially give out the activity location, duration and contact list explicitly. Instead, agents in this implementation are activity pattern based. Every agent is assigned with a reconfigurable index pointing to a week pattern which include seven day patterns. A day pattern consists of a linked list of executable activities which can repeatedly be

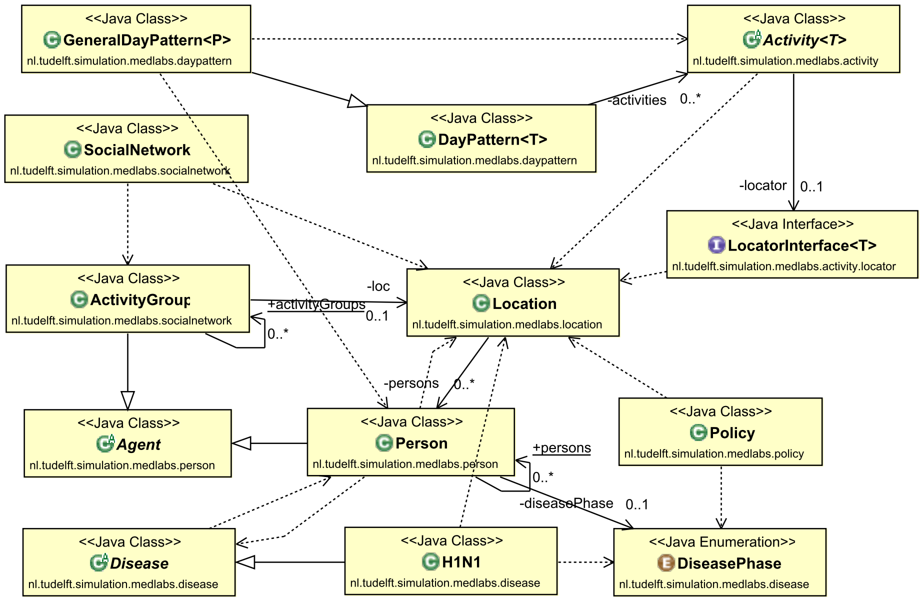


Figure 5.3: Class diagram of the epidemic model of Beijing

executed. Agents can be referenced to the same week pattern if they have similar schedules, but distinct information is provided to them during a simulation run, such as duration and location. That is, activities in the pattern don't offer any direct information regarding activity location and duration and social contacts. With the sequenced execution of the activities in the referenced pattern, the attributes of agents update with the simulation time.

In this research, we designed around 20 different day patterns for all social roles in the artificial city Beijing, which is based on other independent research conclusions. TA et al. (2015) distinguished the working people in the suburb area of Beijing into 5 types by recording the real GPS data and combining the difference in activity (work, eat and shop) distance and commuting frequency. In summary, they differentiated between 5 types of workers: (1) people who work at home and seldom go out; (2) people who work and do other activities nearby (within 3 km); (3) people who do activities in average distance of 7 km to home; (4) people who do activities in an average distance of 10 km to home; (5) people who do activities further than 15 km. Based on this research, firstly we merged type (3)(4) and (5), and then separated the resulting type into 2 new types by the way of commuting to work, which are commuting by public transportation and by private vehicles. The people of the type of commuting by private vehicles were separated into another 2 new types, which are those who need to carpool their children to school every school day and those who don't. For workers during weekend days, 4 types of day patterns were designed according to the conclusions made by the research in YUE, YANWEI et al. (2013), which are: (1) people who stay at home during weekend; (2) people who do activities nearby (within 3 km); (3) people who do activities further than 3 km by public transportation; (4) people who do activities further than 3 km by driving.

For people who are retired, TA et al. (2015) concluded that they behave mostly like Type (1) and (2) of workers. Thus, we designed 2 day patterns for them. The first type prefers to stay at home and the other prefers to do activities outside but nearby. Besides this design, there is also no difference for retired people between weekdays and weekends in this research. For students, due to the scarce data, 3 types of weekday patterns were designed for typical students according to the way they commute to school. For weekend days, 4 types of day patterns were designed similar to workers. Since the commuting ways for students are highly correlated to the distance to schools in the initial data set and the patterns of their parents (those who carpool their children to school or to other shopping and entertainment places), the proportion of assigning patterns to students were determined by the simulation model, both for weekdays and weekends. For babies, we assumed there is only one typical day pattern for them which is associated with their parents who work at home. Since this model is used to predict epidemics, a special day pattern for hospitalized people was designed as well.

A list of all designed day patterns are presented in Table 5.2. An algorithm was implemented to pick the proper weekday patterns and weekend patterns to form a week pattern, and to assign the resulting week pattern to agents during the initialization phase of the simulation.

To give a detailed impression of the designed typical day pattern, a weekday pattern example for workers who carpool their children to school in weekdays is presented in Table 5.3.

A day pattern example for workers who drive outside during weekends is presented in Table 5.4.

5.2.3 Physical Container

Every activity in the pattern that an agent owns is required to assign an activity location. An activity location was modeled as a physical container in this research. It means that the agent has a physical container associated at any time in the simulation, both when performing activities in static places (for example, eating in a restaurant), and during traveling (walking or riding on a bus). The physical container is locatable as well. Thus, every physical container or agent who stays in a place in this model has a geographic reference (longitude and latitude) assigned to it in order to locate it. This definition gives a strict requirement for the completeness and consistency of data required for modeling.

With the data generated by the statistical information, firstly we modeled each of the 8 million static physical containers in the city model of Beijing, which rep-

Table 5.2: Implemented day patterns according to social roles

Social role	Name of Day pattern	Proportion for the social role	Description for the typical day pattern
Infant	B_Pattern	100%	For all babies
Student	S_DayWalk	Based on initial data and model	For students who walk to school in weekdays
Student	S_DayPT	Based on initial data and model	For students who take public transportation in weekdays
Student	S_DayCarpool	Based on parents' pattern	For students who are sent by parents using cars in weekdays
Student	S_WeekendHome	Based on initial data and model	For students who stay at home during weekends
Student	S_WeekendNearby	Based on initial data and model	For students who do activities nearby (within 3 km) during weekends
Student	S_WeekendPT	Based on initial data and model	For students who do activities outside using public transportation during weekends
Student	S_WeekendDrive	Based on parents' pattern	For students who do activities outside with parents by driving during weekends
Worker	W_DayHome	12.9%	For workers who work at home
Worker	W_DayNearby	12.2%	For workers who work nearby (within 3 km)
Worker	W_DayPT	33.7%	For workers who take public transportation to work
Worker	W_DayDrive	19.2%	For workers who drive to work
Worker	W_DayCarpool	22%	For workers who drive but carpool child to school first
Worker	W_WeekendDayHome	20%	For workers who stay at home during weekends
Worker	W_WeekendDayNearby	20%	For workers who do activities nearby (within 3 km) during weekends
Worker	W_WeekendDayPT	30%	For workers who do activities by public transportation during weekends
Worker	W_WeekendDayDrive	30%	For workers who do activities by driving cars during weekends
Retired	R_DayHome	50%	For retired people who prefer staying at home
Retired	R_DayOut	50%	For retired people who prefer do activities outside
ALL	HospitalizedDay	Based on simulation	For hospitalized people

Table 5.3: A day pattern example for workers who carpool children to school in weekdays

No.	Activity Name	Activity Type	Duration
1	sleep	StochasticDurationActivity	Triangular(6.0, 7.0, 7.5)
2	carpool Child	CarpoolActivity	based on simulation
3	work	UntilFixedTimeActivity	until 12:00 am
4	lunch and rest	StochasticDurationActivity	Triangular(0.4, 0.6, 1.0)
5	work	StochasticDurationActivity	Uniform(4.0,7.0)
6	drive home	TravelActivityCar	based on simulation
7	walk to shop	TravelActivityWalk	based on simulation
8	shop	StochasticDurationActivity	Triangular(0.1, 0.3, 0.5)
9	walk home	TravelActivityWalk	based on simulation
10	family Dinner	FamilySynchronizedActivity	Fixed(20:00-21:00)
11	housework	StochasticDurationActivity	Uniform(1.0,2.0)
12	sleep till midnight	UntilFixedTimeActivity	until 24:00

Table 5.4: A day pattern example for workers who drive outside during weekends

No.	Activity Name	Activity Type	Duration
1	sleep	StochasticDurationActivity	Triangular(7.5, 8.5, 10.0)
2	housework	StochasticDurationActivity	Uniform(1.0,4.0)
3	drive to shop/relax	TravelActivityCar	based on simulation
4	shop/relax	StochasticDurationActivity	Uniform(2.0,10.0)
5	eat	StochasticDurationActivity	Triangular(0.4, 0.6, 1.0)
6	drive home	TravelActivityCar	based on simulation
7	housework	StochasticDurationActivity	Uniform(0.5,2.0)
8	sleep till midnight	UntilFixedTimeActivity	until 24:00

resent schools, restaurants, shops, hospitals, etc. The exact numbers are shown in Table 5.5.

From Table 5.5, we can find that currently there are 18 types of static physical containers which are categorized into 6 categories. Obviously, these can not cover all the types in Beijing, for example, small shops in 'Consumption locations' and cinemas in 'Entertainment locations' are missing in the current database. Further research should be conducted on generating or collecting real data for these missing types which are important for disease spread, as well.

Besides static physical containers, we modeled a transportation system including movable physical containers to execute travel activities, which helps commut-

Table 5.5: Static physical container statistics

Category	Type	Size
Houses	household	4.961 million
Consumption locations	restaurant	55257
Consumption locations	market	18686
Consumption locations	mall	547
Medical institutions	clinic	836
Medical institutions	community meds	1744
Medical institutions	hospital	569
Medical institutions	medservice	3335
Educational institutions	elementary	1090
Educational institutions	kindergarten	1305
Educational institutions	middle school	632
Educational institutions	middle university	91
Educational institutions	private university	79
Educational institutions	university	73
Entertainment locations	green	13983
Entertainment locations	playground	6151
Entertainment locations	garden	93
Other workplaces	workplace	11431

ing agents to determine a route and give out the travel duration. The public transportation system in the model is microscopic, where we modeled all lines and stops of metro and bus system in Beijing. The public transportation in Beijing contains 17 Metro lines, 227 Metro stations and nearly 1,000 public bus and trolleybus lines in the city, which makes it one of the largest public transportation systems in the world. A picture of Beijing metro system is present in Figure 5.4 from [4].

Modeled buses and metro trains will execute their schedules on these routes based certain timetables. The geographic information and routing data of the transportation infrastructure network are acquired from open source based OpenStreetMap [5] by using the Java library osmosis [6]. To show the topology of the whole public transportation network in Beijing, a graph is built in this model by the Java library jgrapht [7]. It models stops as nodes (static physical containers) and routes as links. For commuting vehicles (cars and taxis) on the road network, we don't model the real road networks, but calculate estimated travel duration according to the distance and historical statistical data on congestion.

[4] http://www.johomaps.com/as/china/beijing/beijingmetro.html
[5] http://www.openstreetmap.org
[6] http://wiki.openstreetmap.org/wiki/Osmosis
[7] http://jgrapht.org

Figure 5.4: Map of Beijing's metro system from JohoMaps

5.2.4 Social Regulation

The policies for interventions during disease outbreak are modeled as the main social regulations in this research. One of the policies modeled can be described as "When the number of 'infected' agents is observed to reach a threshold, agent who are in the phase of 'Symptomatic' are required to stay at home for isolation. If violated, agents will be forced to be isolated in hospitals".

In this social regulation, the *Type* is defined as "sanction". A *Monitor* is used to observe the number of 'infected' agents in the simulation. When the observed results are higher than a pre-defined threshold which is a *Standard* in the regulation, two steps of *Operations* are triggered.

The first step is specifying new behavior for agents who are under the regulation in the form of assigning new activity patterns. After agents receive the new assignments, the strategical level decision-making module of agents will decide whether they accept this change or not. If an agent decides to decline, the second step of the *Operations* is that this agent will be forced to assign another activity pattern.

5.2.5 Social Networks and Social Interaction

Three types of social networks are dynamically generated in the model, which are family, colleagues/ classmates and friendships. Family and colleagues/ classmates are intuitively generated by checking the other agents who are in the same living and working physical containers, while friendships are generated according to the algorithms specified in Section 4.4.1.

Besides regular interaction among agents, the agents in this model are able to communicate for scheduling joint social activities. This feature is realized through a dynamic construction of social networks and execution of social activities. When executing joint social activities, a functional entity called 'activity group' is generated to organize and manage the participants and social contact network emerges from the execution.

According to the fact that Dunbar's number (HILL and DUNBAR 2003) ranges from 100 to 250, the largest size of friends in this research is set to the lower boundary 100 to reduce the computational complexity. The skewness is set to 0.8, which is an example experiment setting in (HAMILL and GILBERT 2010).

5.2.6 Functional Entity: Modeling H1N1

Besides 'activity group', pandemic influenza is modeled as another functional entity in this model, which can change agents' health status 'diseasePhase'. The phase transitions are modeled according to the research of STROUD and VALLE (2007). In addition to their disease transition model, a phase called '*Vaccinated*' was added in this research, which can be used for policy modeling. The phase transitions and details about the transition time and probability are presented in Figure 5.5.

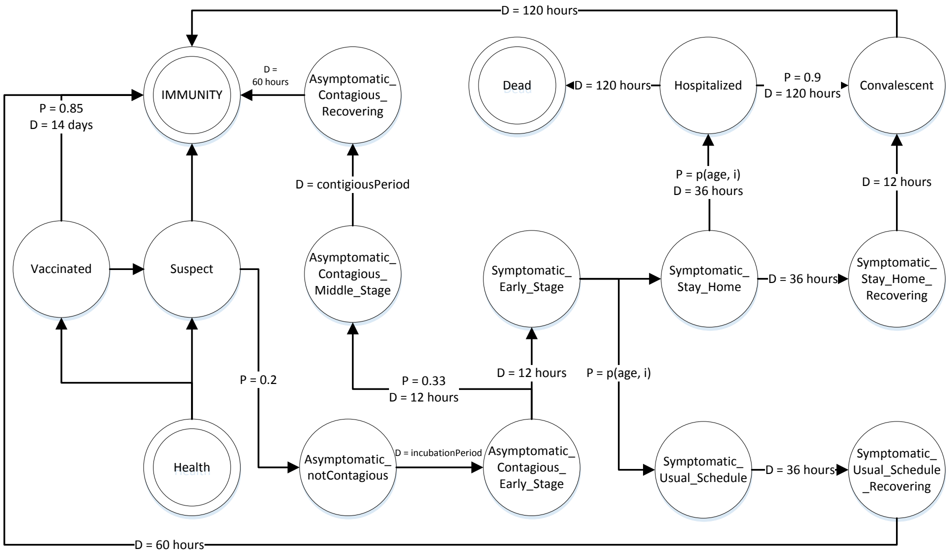


Figure 5.5: Phase transition of disease model (H1N1)

An infected agent is contagious as of the phase of 'Asymptomatic_Contagious_Early_Stage' until the phase of 'Convalescent' or the end phases of 'Dead' or 'IMMUNITY'. However, the contagious probability varies for different transition phases. The basic contagious rates in the phases are initialized in Table 5.6 based on the research of STROUD and VALLE (ibid.), which are considered as an initial experimental setting.

Table 5.6: Basic Transmission Probability(P) of All Contagious Phases

Phase Label	Phase	P
R_{ACE}	Asymptomatic_Contagious_Early_Stage	0.15
R_{ACM}	Asymptomatic_Contagious_Middle_Stage	0.5
R_{ACR}	Asymptomatic_Contagious_Recovering	0.125
R_{SES}	Symptomatic_Early_Stage	1.0
R_{SUS}	Symptomatic_Usual_Schedule	1.0
R_{SUSR}	Symptomatic_Usual_Schedule_Recovering	0.25
R_{SSH}	Symptomatic_Stay_Home	1.0
R_{SSHR}	Symptomatic_Stay_Home_Recovering	0.25
R_H	Hospitalized	0.25

Besides the basic contagious rates, the probability to infect a susceptible person is also highly related to factors such as the space of the sub-location, the number of infected persons in the same sub-location and the contact duration (BEGGS et al. 2003, HOUK et al. 1968). Because of this, we added more parameters in the disease progression model. The final contagious rate for a susceptible person i in a sub-location L containing N infected people can be calculated through Equation 5.1, where R_j can be found in Table 5.6, β is a corrective coefficient for the basic contagious probability, σ_L is a corrective coefficient for the sub-location, S_L is the space of the sub-location (in square meters) and t_ij is the contact duration between person i and j.

$$R(i,N,L) = \frac{(1 - e^{-\sum_{j=1}^{N} \beta \times R_j \times t_{ij}}) \times \sigma_L}{S_L} \tag{5.1}$$

In this research, the corrective coefficients β and σ_L in Equation 5.1 are both set to 1.0. This simplification is determined as one possible experimental setting and the sensitivity analysis of this set is not the research interest in this research.

With the above components implemented, the large-scale agent-based epidemic model is constructed. Before studying the epidemic dynamics under different interventions, a process of verification and validation for the underlined artificial city model is conducted.

5.3 Verification and Validation for the Artificial City Model

Two experiments are designed for verification and validation of the artificial model, in which the whole modeled population perform daily activities based on their weekday patterns and weekend patterns. There are ten replications in each experiment (Shapiro-Wilk test is conducted) and each replication runs for one simulation day (24 hours) based on the JAVA-based simulator DSOL (JACOBS et al. 2002). The warm-up period is set to 0 in the experiments as the simulation time starts at 0am and all agents are considered in their house locations as a hypothesis. In the simulation experiments, disease control interventions are excluded. The human spatial contacts in all the modeled locations (physical containers) are recorded for validation. The results are categorized into two types, indoor contacts in static locations and outdoor contacts during commuting. The indoor contacts are mainly used for verification purpose due to the available data while outdoor contacts are mainly used for validation.

5.3.1 Indoor Contacts

Spatial contacts are the main transmission measures for pandemic influenza to propagate among human beings. STROUD and VALLE (2007) studied the spatial dynamics of pandemic influenza in a massive artificial society. TOROCZKAI and GUCLU (2007) presented a framework to account for the dynamics of contacts in epidemic processes as well.

Spatial contacts emerge when individuals execute their daily activities in physical containers. For example, spatial contacts can emerge among students who are in the same school location. When executing a school activity in a student's day pattern, and another student is executing a school activity at the same location and the periods have overlap with each other, these two students are considered to have a spatial contact in this model.

Through the execution of the simulation experiment for weekday, firstly we show the hourly number of people in several typical types of locations in one replication in Figure 5.6, where the time of the day (0:00-24:00) goes on the x-axis. The 'others' item in the figure represents all the other location types according to Table 5.5.

From Figure 5.6, it can be found in this weekday replication that the largest part of the population during the day time is in their houses. The statistical results of the hourly number of people in the house location for ten replications are presented in Figure 5.7, where the 95% confidence interval is drawn in the sample point (each hour).

Since all the population in this research are modeled into four social roles (baby, worker, student and retired), the hourly results of agents with different role in the house location are presented in Figure 5.8 for the weekday experiment and in Figure 5.9 for the weekend experiment.

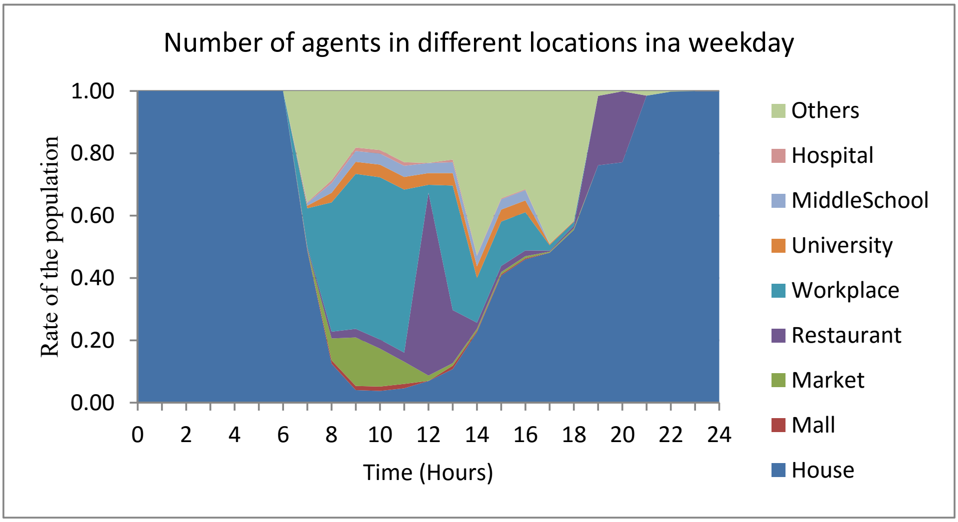


Figure 5.6: Number of people in physical containers

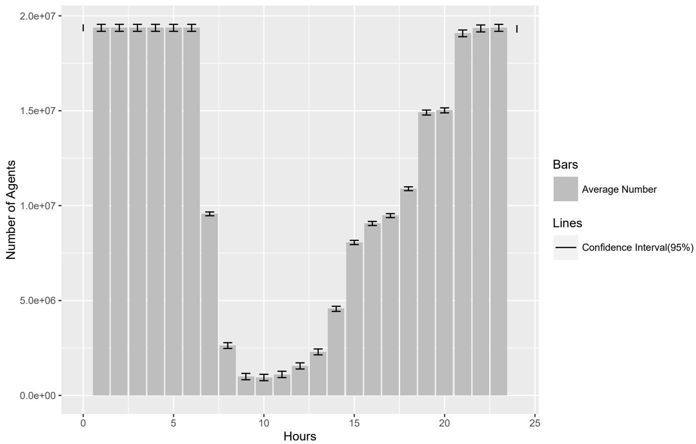
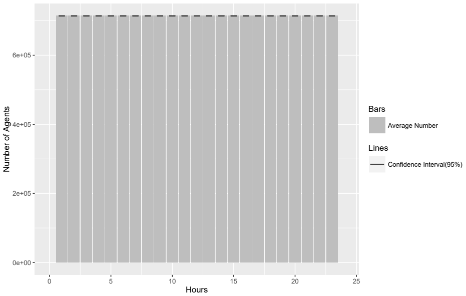


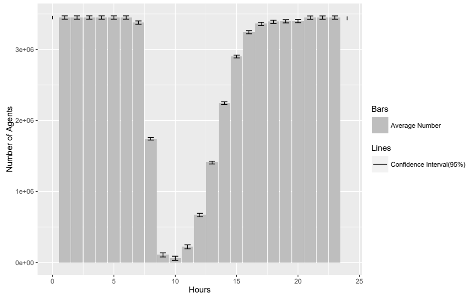
Figure 5.7: Number of total agents in the house location in a weekday (10 replications)

In Figure 5.8a, we can find that the baby agents stay at home for all 24 hours. This is the result of the design of the baby pattern, in which babies are modeled to execute all activities at home. Since the results for the baby agents are the same between the weekday and weekend experiments, no results are shown for babies in Figure 5.9.

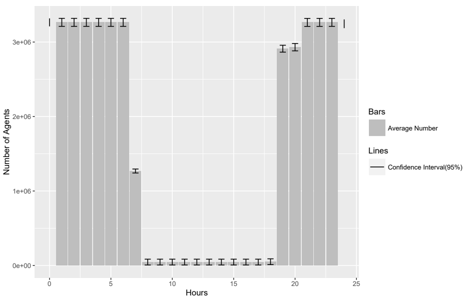
Besides the results in the house location, more results and evaluation analysis



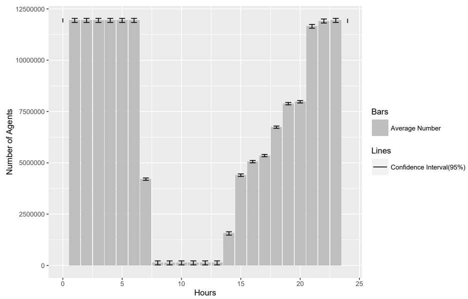
(a) Number of baby agents in the house location in a weekday



(b) Number of retired agents in the house location in a weekday



(c) Number of student agents in the house location in a weekday



(d) Number of worker agents in the house location in a weekday

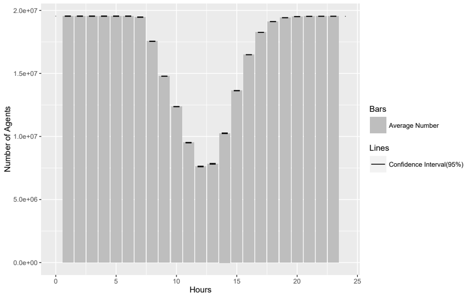
Figure 5.8: Statistics of the number of agents with social roles in the house locations in a weekday (10 replications)

are presented in Appendix A.

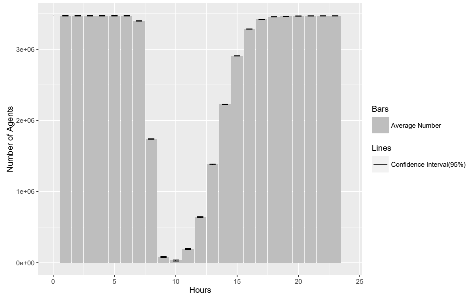
Due to the design of the activity in the pattern, the duration of staying in different types of locations varies among agents even when they use the same activity pattern. To verify this design, the average duration of agents staying in different locations in the weekday experiment is presented in Table 5.7.

From Table 5.7, we can find that the longest duration of stay occurs in households, followed by work or study places.

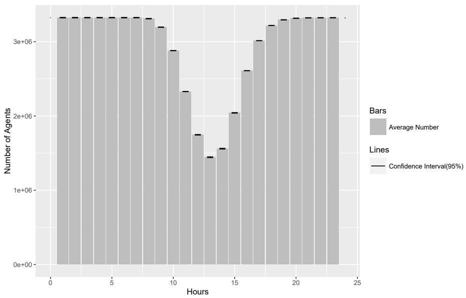
It's not difficult to find the causal relationship between the designed 20 day patterns in Table 5.2 for all the agents and the experiment results as a verification evaluation. To validate this design to some extent, the result of a survey by WANG et al. (2011) is used to compare with the experiment results. WANG et al. (ibid.) present the time-use patterns of the different neighborhood on a normal workday for workers. From this, two representative neighborhoods, TRA and CHC are



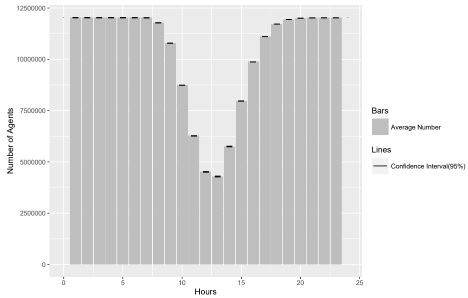
(a) Number of total agents in the house location in a weekend day



(b) Number of retired agents in the house location in a weekend day



(c) Number of student agents in the house location in a weekend day



(d) Number of worker agents in the house location in a weekend day

Figure 5.9: Statistics of the number of agents with social roles in the house locations in a weekend day (10 replications)

Table 5.7: Average duration by location types

Type	Average Duration	Standard Deviation	Confidence Interval (95%)
Household	10.2 hours	4.9 hours	[7.16, 13.24]
Mall	0.5 hours	0.2 hours	[0.38, 0.62]
Market	0.3 hours	0.1 hours	[0.24, 0.36]
Restaurant	1.3 hours	0.8 hours	[0.80, 1.80]
Workplace	4.2 hours	2.3 hours	[2.77, 5.63]
University	6.0 hours	3.6 hours	[3.77, 8.23]
Middle school	4.5 hours	2.1 hours	[3.20, 5.80]
Hospital	0.9 hours	0.4 hours	[0.65, 1.15]
Clinic	0.5 hours	0.2 hours	[0.38, 0.62]

chosen. Since there is only workers' result in the research by WANG et al. (2011) and the duration in different places are simply categorized into home, out-of-home and travel, we recorded the duration for workers in different locations separately and made a comparison in Table 5.8, where the duration in travel is excluded.

Table 5.8: Comparison of duration in home/out-of-home locations for workers in a weekday

Item		**Simulation results**	**TRA**	**CHC**
In-home	Mean	11.4 hours	14.5 hours	15.6 hours
	CI(95%)	[7.93, 14.87]	[12.14, 16.86]	[13.06, 18.14]
Out-of-home	Mean	9.1 hours	8.0 hours	6.9 hours
	CI(95%)	[7.05, 11.15]	[5.83, 10.17]	[4.54, 9.26]

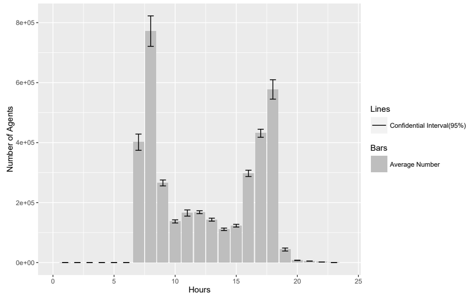
From Table 5.2, we can find that the relative error of the average duration of staying In-Home between TRA (equivalent to household in this research) and the experiment is relatively high (21.3%), compared to the average duration of staying Out-of-home between TRA (13.8%) and the experiment. This difference can be caused by many factors, such as the season of the survey, the monotonicity of the surveyed neighborhood and the incompleteness of our designed activity pattern. As our interest in this research is in a new agent-based modeling method, we accept this error while further surveys on human behavior patterns in Beijing is required in future research.

Due to the design of movable physical containers, agents can have spatial contacts outside which is considered as one contribution in this research on large-scale systems. Thus, the statistics on agents during commuting are discussed in the next section.

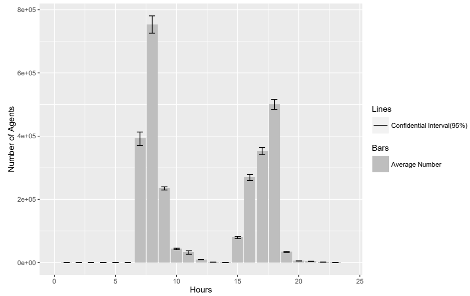
5.3.2 Outdoor Contacts

Travel contacts, also as spatial contacts in movable physical containers, emerge from the inclusion of the public transportation component in this model. We observed the information about the number of people in the public transportation infrastructure components, such as metro stops, metro trains, buses and bus stops during a working day. As an example, how the numbers of agents with different social roles in the bus location change in a weekday is shown in Figure 5.10. More results are shown in Appendix A. Through this transportation component, travel contacts emerge. In this research, stops or metro trains are divided into several sub-physical containers to represent platforms or train compartments, where agents can have travel contacts when they are in the same sub-physical containers at the same time.

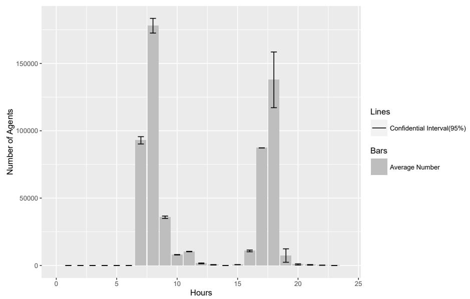
As we described before, the duration of a travel activity by bus/metro will be decided by the simulation model, and is dependent on several factors, such as the



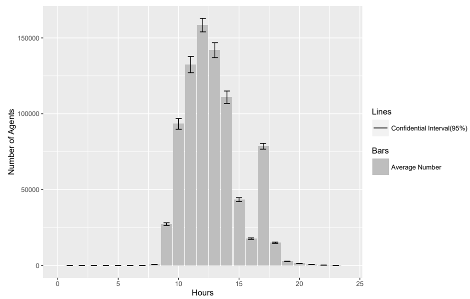
(a) Number of total agents in the bus location in a weekday



(b) Number of worker agents in the bus location in a weekday



(c) Number of student agents in the bus location in a weekday



(d) Number of retired agents in the bus location in a weekday

Figure 5.10: Statistics of the number of agents in the bus location in a weekday (10 replications)

travel distance, the path that the agent chooses (e.g. Dijkstra shortest path) and the waiting queue in the metro stops.

Validation of a model with a wide range of parameters would be very difficult (STOCKER et al. 2001). Thus, this simulation study shifts the focus to validation using several travel statistics. In order to validate the results in this public transportation component of the whole model, the average travel volume in a weekday by bus and by metro are compared to the historical traffic statistics report in 2011 (GUO and LI 2012) in Table 5.9. The reason for adopting the traffic statistics report in 2011 is to keep this research consistent as the generated population data is based on the census data of 2011 as well.

From the comparison in Table 5.9, we can find that the relative errors between simulation results and the historical traffic statistics are within 15%. Several factors are responsible for the differences and one of the crucial differences is that

Table 5.9: Comparison of daily travel volume by public transportation in a weekday (ibid.)

Item	Simulation results	Historical statistics
by bus	7.28 million	8.11 million
by metro	4.55 million	3.95 million

the data collected in the report (GUO and LI 2012) only covers part of Beijing city (within the 6th Ring Road). This difference will increase the total relative errors to 28% as the daily travel volume within the 6th Ring Road only accounts for 87% of the whole travel volume in Beijing.

Regarding the travel purpose, Table 5.10 shows the comparison of the main purposes of using public transportation in a weekday. The relative errors are less than 10%.

Table 5.10: Comparison of daily travel purpose in a weekday (ibid.)

Item	Simulation results	Historical statistics
For working and school	59.2%	54.5%
For shopping	8.1%	7.6%
For leisure	6.1%	6.5%

From Figure 5.10, it can be found that the rush hours for public traveling are from 7 am to 8 am and from 5 pm to 6 pm, which match the historical traffic statistics (ibid.).

Besides travel volume and travel purpose, travel duration is used to make a comparison as well. The data for comparison comes from the survey data used in the research by ZHAO et al. (2011) which present a survey data on commuting time (travel duration in this research) in a weekday conducted in a neighborhood in Beijing in 2001.

Table 5.11: Comparison of daily travel duration by public transportation in a weekday (ibid.)

Item	Simulation results (min)	Survey data (min)
Mean time	66.4	52.4

The relative errors between simulation results and the real data mainly come from the lack of certain activity patterns in the model, which results in the missing of a large amount of travel volume. For example, the model does not include patterns for business people and tourists who would use the public transportation multiple times in one day. These patterns were excluded in the model due to the lack of available data.

As a conclusion, we listed the missing components in the artificial city model that can be easily improved when the associated data becomes available.

- More refined activity patterns, such as worker pattern in night shift, tourist pattern, business people pattern.
- More rules in agents' architecture when making decisions. For example, people in reality would consider the choice of routes based on the price of tickets before traveling while agents in this research only consider the shortest path.
- More accurate distribution of the starting time, duration and ending time of activities. For example, the departure time to workplaces for workers who are employed by universities should be earlier than those who work in restaurants in general. For now, the departure time for workers with different type of jobs follows the same distribution in this thesis.

5.4 Disease Spread

5.4.1 Disease Dynamics

With the verified and validated artificial city model, statistics of disease spread can be used to validate the whole epidemic model in the artificial city. Thus, a simulation experiment with 10 replications for disease spread among the modeled 19 million agents was constructed.

Agents, who are in the disease (H1N1) outbreak in this research, have 16 potential phases which are presented in Figure 5.5. The contagious rate for a susceptible person in a sub-location containing multiple infected people can be calculated through Equation 5.1 which is presented in 5.2.6.

With these settings, we recorded the number of agents in different phases during the simulation run for 30 days in each replication. The initial condition for the disease model was that 1 in 2 million people in the population was in the '*Suspect*' phase.

In the disease spread model, the 16 potential phases are categorized into two types: end phases and transitional phases. Firstly, we present the number of agents in the end phases of 'IMMUNITY' in Figure 5.11. Other end phases such as 'HEALTHY' and 'DEAD' are presented in Appendix B ('IMMUNITY' is also presented as well).

One example of the transitional phases is 'Hospitalized', which is presented in Figure 5.12. Other major transitional phases are presented in Appendix B. It is worthy to mention that there is a drop in the trend-line in the phases such as 'Asymptomatic_Contagious_Early_Stage' in Simulation day 6. This is because all simulation studies in this thesis start from Monday and agents in Saturday (Simulation day 6) will have less contacts based on weekend patterns.

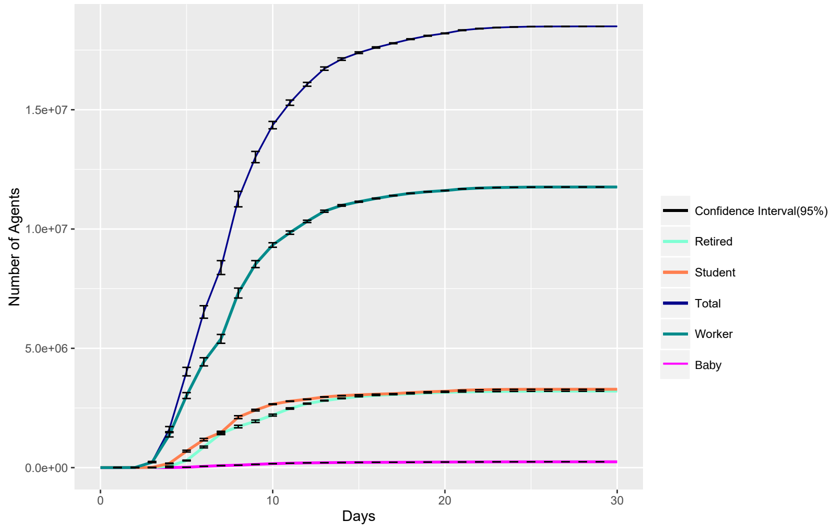


Figure 5.11: Number of agents in the phase 'IMMUNITY' (10 replications)

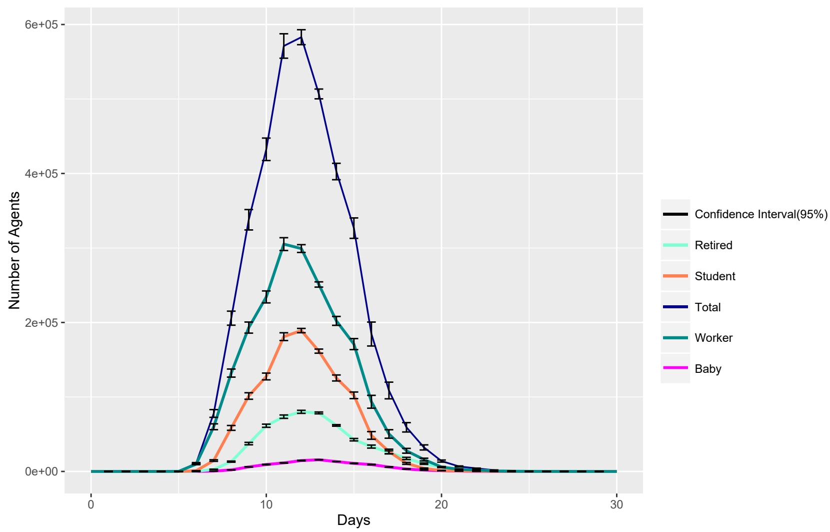


Figure 5.12: Number of agents in the phase 'Hospitalized' (10 replications)

5.4.2 Model Validation

As we stated before, the phase transitions are modeled mainly according to the research of STROUD and VALLE (2007). To validate the results of disease spread in this research, we made a comparison with the results of disease spread in (ibid.) through giving out the distribution of 'infected' agents (from Phase 'Asymptomatic_notContagious') by age group in Figure 5.13 and by infection location type in Figure 5.14.

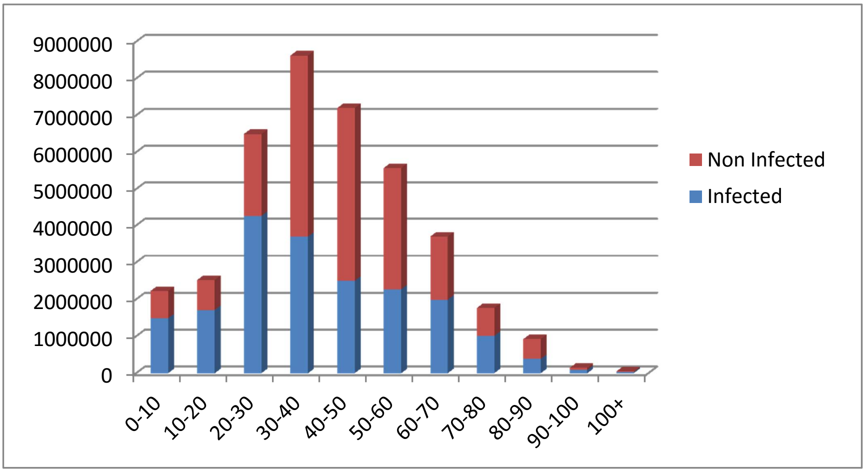


Figure 5.13: Distribution of infected agents' age

From Figure 5.13, we find find that the age group '10-20' (mainly consisting of students) got the highest infection rate which is aligned with the conclusion in (ibid.). CDC of the United States collected and analyzed the reported cases in 2009 and concluded that the infection rate is the highest among people in the 5 years to 24 years of age group [8].

From Figure 5.14, we can find that household (home) is the most possible location type for disease spread among the full population, followed by workplaces, schools and transportation. This result is also consistent with the conclusion in (ibid.). To give a detailed view, the distribution of infected agents with different social roles in different location categories is presented in Figure 5.15.

We also presented the distribution of infection sources for different social roles in Figure 5.16.

From Figure 5.16, we can find that the biggest part of the infections for a given social role are from the same social role type except for babies. It can be explained by the fact that students, workers and retired people stay with each other in most of their day time while babies always stay with their parents. Especially, workers get a higher infection possibility from their companions than the other social roles. It is caused by both the facts that workers are the biggest part of the population and workers are in more closed spaces during the day.

[8] http://www.cdc.gov/h1n1flu/surveillanceqa.htm

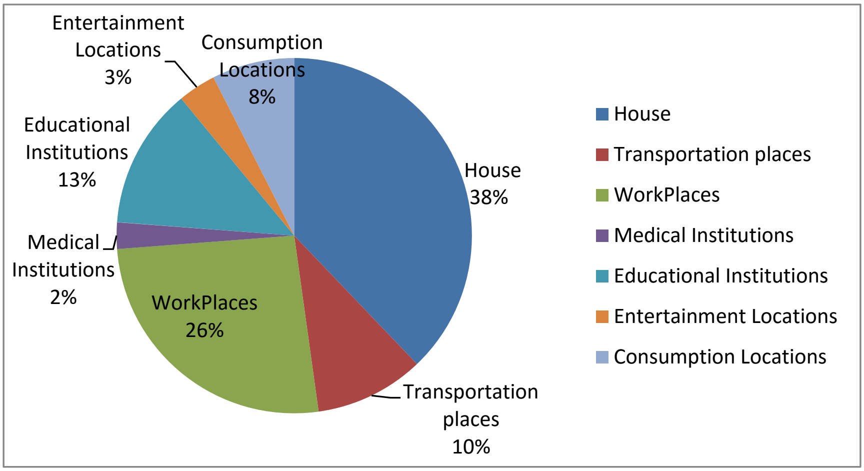
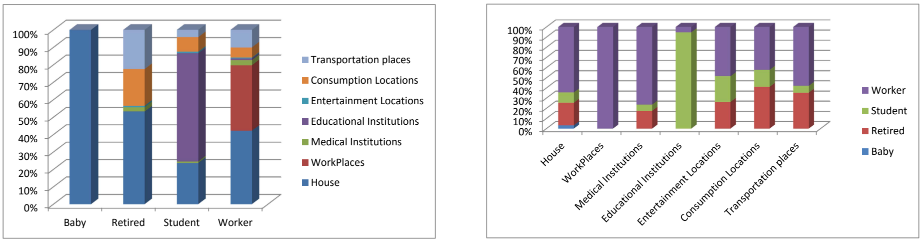


Figure 5.14: Distribution of location type for infection among full population

Although some basic disease dynamics reported in the figures above are consistent with STROUD and VALLE (2007), there are many difference in other specific indicators. For example, STROUD and VALLE (ibid.) reported that about 10% of the population will be symptomatic or convalescing at the pandemic peak after 30 days of stable and exponential growth, while we get a result of around 5% of the population that will be convalescing after 16 days. Furthermore, there are are also difference in the concrete numerical values in terms of the the breakout of cumulative infections by location type and the clinical attack rates by age groups. Since we believe these differences are correlated to the artificial city model and the underlying population data (China vs USA) are different, these indicators will not be validated in this research.

In reality, there was an H1N1 outbreak in Beijing in 2009 which lasted more than six months. However, the historical data, including the peak number of 'infection' and the time of the peak, will not be used for validation in this research due to many factors. First of all, the peak number of reported 'infection' was based on confirmed cases. These cases do not distinguish the disease phases and do not contain detailed personal information. Secondly, the reported peak time (day) lasted a rather long period as a series of interventions were conducted by different authorities in different part of the city among different social roles, from the first case in May 16 to the peak time on October 28, 2009. Therefore, the simulated results of the extreme situation in this research cannot be validated. Instead, an expert validation process is required as part of future research.



(a) Distribution of location type for infection by social roles

(b) Distribution of social roles with infection by location types

Figure 5.15: Distribution of location type and social role for infection

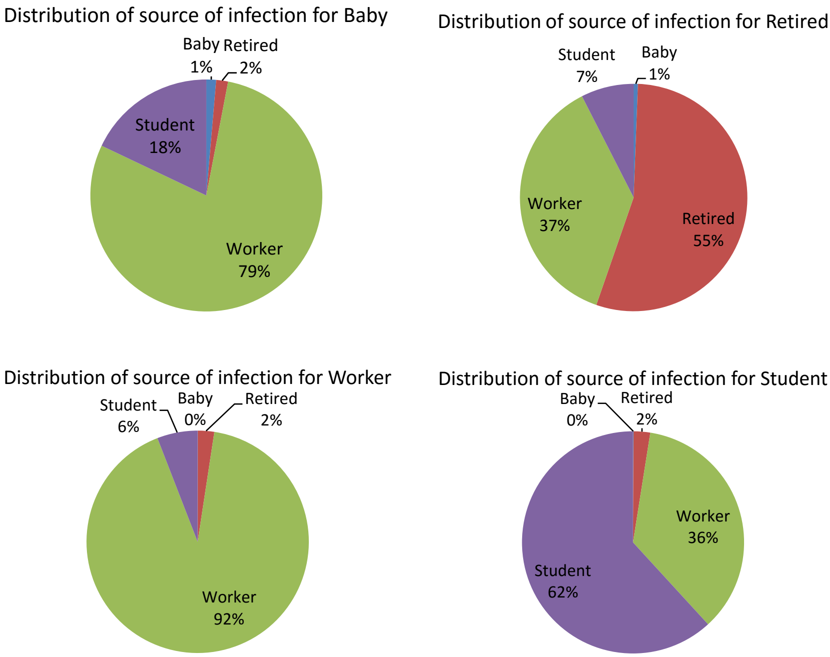


Figure 5.16: Distribution of the source of infection for different social roles

5.5 Comparison with Related Works

5.5.1 Model with Same Disease Transition Model

Since this research shares the same disease (H1N1) model with the research EpiSimS by STROUD and VALLE (ibid.) for prediction in southern California, it is necessary to compare the model by EpiSimS with the model in this research.

Besides the differences in data and parameters such as the basic contagious rate (R0) and the data source, the major differences are reflected in the choices in the design and implementation phases.

- Although sublocations are modeled in EpiSimS, the activity locations (physical containers) are organized in a more hierarchical way in this research.
- Weekdays and weekend days are averaged to get a representative day in EpiSimS while they are separately modeled in this research.
- EpiSimS does not capture disease transmission during travel while this research includes a public transportation component for commuting.
- Agents' behavior are based on fixed schedules in EpiSimS while both activity pattern can be replaced and specified activities (e.g., social activity) in the pattern can be rescheduled in this research.

5.5.2 Model with Same Data Source

Although the population and environmental data originates from GE et al. (2014)'s research, this research is independent and the way to design and implement the artificial city model and epidemic prediction model is different. To show how the research in this thesis is unique and innovative, we made a comparison between this research and the KD-ACP framework (CHEN et al. 2015) which was used to implement an epidemic model based on the same data.

- Agents implemented by KD-ACP behave according to fixed activity schedules in terms of the activity sequence, the activity locations (fixed choices) and duration. That is, agents in KD-ACP do not have decision-making capabilities. This thesis models agents in a different way by which agents own multi-level decision-making capabilities while still staying "simple" and "small" enough for computational efficiency.
- Social networks in KD-ACP are predefined in the initial data, thus, no unscheduled joint social activities can be executed in the simulation. This thesis generates social networks for agents dynamically by which agents can have complex social interactions in order to join in unscheduled joint social activities.
- Subway networks are modeled to represent the whole public transportation in KD-ACP. A lot of efforts are required to complete the public transportation networks. However, this thesis archives this task easily by proposing the concept of 'movable physical container'.
- The disease model are considered to be validated in KD-ACP in two indicators, the infection trend and the basic reproduction number. This thesis verifies and validates the model in both people's daily behavior and infection details, which include more model details.

5.6 Discussion

This chapter presented a model of a large-scale artificial city of Beijing, on which to test policies for controlling the spread of disease among the full population (19.6 million). This is used as a case study to test the proposed large-scale agent-based social conceptual model. Firstly, by combining diverse data sets, including generated census-based data, open source maps, activity patterns, an artificial city with a large population was constructed. In this artificial city, each of the 8 million physical locations and 19.6 million citizens was modeled. A microscopic public transport system (subways and buses) together with a predicted road traffic system are simulated in an artificial city and are well integrated with the daily activities of the population. With this model, spatial contact networks emerge and can be observed during the execution of the model. Thus, research requirements RR9 and RR10 are satisfied.

✓ **RR9 Research Requirement - Agent Environment in the Case Study** The agent environment in the large-scale epidemic prediction should model different kinds of physical spaces including movable spaces for disease transmission.

✓ **RR10 Research Requirement - Social Interaction in the Case Study** Large-scale agent-based epidemic prediction and control should have capabilities to model complex human social interactions.

Secondly, to investigate the effect of the emerging spatial contact network for epidemic prediction, a pandemic influenza disease progression model was implemented and several scenarios related to adding different contact types were tested.

The disease was modeled as a functional entity to change agents' healthy status. The concept of 'Functional Entity' was introduced to model those extra objects in the system that can influence or directly change attributes of either agents, physical containers or social regulations. For example, weather is modeled as a functional entity to change the temperature of physical containers. Thus, research requirement RR6 is satisfied.

✓ **RR6 Research Requirement - Disease in the Case Study** Large-scale epidemic prediction and control should have a flexible mechanism to model new diseases with different phase transition process.

6 Exploratory Model Studies

The popularity of large-scale agent-based social simulation comes from the fact that it can help to explore different outcomes for phenomena where people might not be able to view the results in real life. It can provide sociologically valuable information on the society, and provide decision makers the outcomes of a new policy on the society for policy evaluation. This chapter presents two kinds of simulation studies to test the quality and explore the possibility of the implemented large-scale agent-based model for epidemic prediction and control.

6.1 Simulation Study on Model Outcomes

Due to the proposed conceptual model and the implemented model, this simulation study on model outcomes covers several unique aspects that are not presented in other similar research. Two novel aspects among others in the model are the inclusion of public transportation and dynamic social interactions. In this section, two scenarios are used to test the effect of the inclusion of these two aspects in the model.

Before studying these scenarios, a baseline scenario without the inclusion of public transportation and dynamic social interactions was created to study the disease dynamics at first, which is considered as a traditional model for large-scale agent-based epidemic prediction. The initial condition for the model was that 1 in 2 million people in the population was in the '*Suspect*' phase. The number of persons in the '*Hospitalized*' phase was recorded during a simulation run of 30 days.

6.1.1 A Baseline Scenario without Public Transportation and Dynamic Social Interactions

The first scenario is a baseline model where agents will only have contacts in static physical containers such as home, schools and workplaces where movable physical containers such as buses and metro trains are excluded from the model. Agents will not go to attend social activities either, as no joint social activities will be included in agents' pattern. Furthermore, agents don't have decision-making capability during disease outbreak, which means no agents will decide to stay at home even when he/she has a symptom. No interventions are conducted during the simulation run either. This scenario contained an experiment with 5 replications, and the number of agents in the '*Hospitalized*' phase for each replication were recorded.

The statistical results are presented in Figure 6.1.

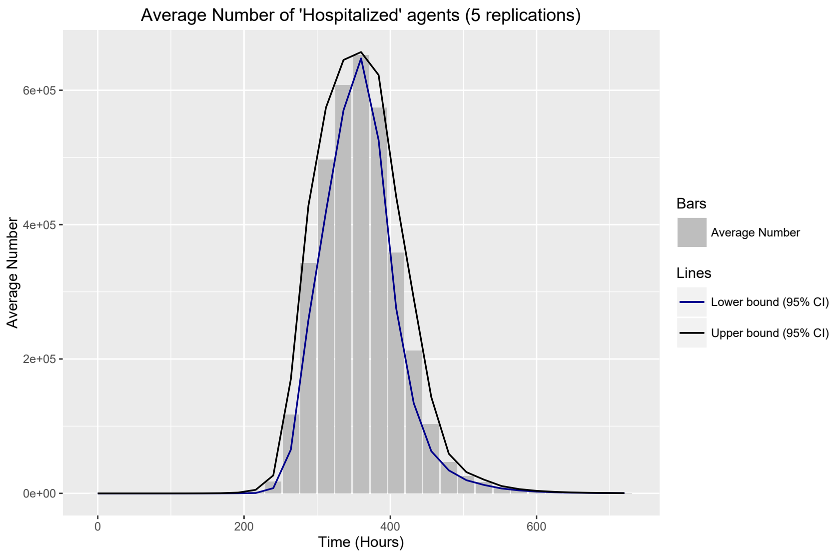


Figure 6.1: Number of '*Hospitalized*' agents in a baseline scenario

The major indicators and statistics are shown in Table 6.1 and the presented two KPIs will be compared with the results from the other scenarios as they can intuitively show the dynamics of disease transmission. The first indicator 'Average Peak Time' shows the average time point when the number of '*Hospitalized*' agents reaches the highest value and the second indicator 'Peak Number' presents the number of '*Hospitalized*' agents at the highest point. Besides the mean value and 95% confidence interval, the p value of the Shapiro-Wilk Test is also given to test the normality of indicators.

Table 6.1: Results of '*Hospitalized*' agents in a baseline scenario after 5 replications

Result	Mean	95% Confidence Interval	Shapiro-Wilk Test (P Value)
Average Peak Time (Hours)	362	[361.0, 363.0]	0.135
Peak Number	652,598	[647,989, 657,207]	0.9189

It can be found from Table 6.1 that both the 'Average Peak Time' and 'Peak Number' came from a normal distribution, which can be used to compare with other statistical results in other scenarios.

6.1.2 Scenario of Including Public Transportation

In order to test the quality of implemented physical containers, a scenario was designed to test the impact of including travel contacts in movable physical containers on disease spread, where the microscopic public transportation system was included in the model and agents can realistically 'travel' through the city and have travel contacts.

For this scenario, we also conducted an experiment with 5 replications. For each replication, the numbers of agents in the '*Hospitalized*' phase were recorded and the statistical results are presented in Figure 6.2.

Similar to the baseline scenario, the key statistics on indicators of 'Average Peak Time' and 'Peak Number' are shown in Table 6.2.

Table 6.2: Results of '*Hospitalized*' agents when including travel contacts after 5 replications

Result	Mean	95% Confidence Interval	Shapiro-Wilk Test (P Value)
Average Peak Time (Hours)	375	[371.4, 378.6]	0.7147
Peak Number	700,138	[691,524, 708,752]	0.4913

From Table 6.2, we can find both the 'Average Peak Time' and 'Peak Number' follow a normally distribution as well. To show the difference of the simulation results between the baseline scenario and the scenario including movable physical containers, we did statistical tests using SPSS and present the outcome in Table 6.3.

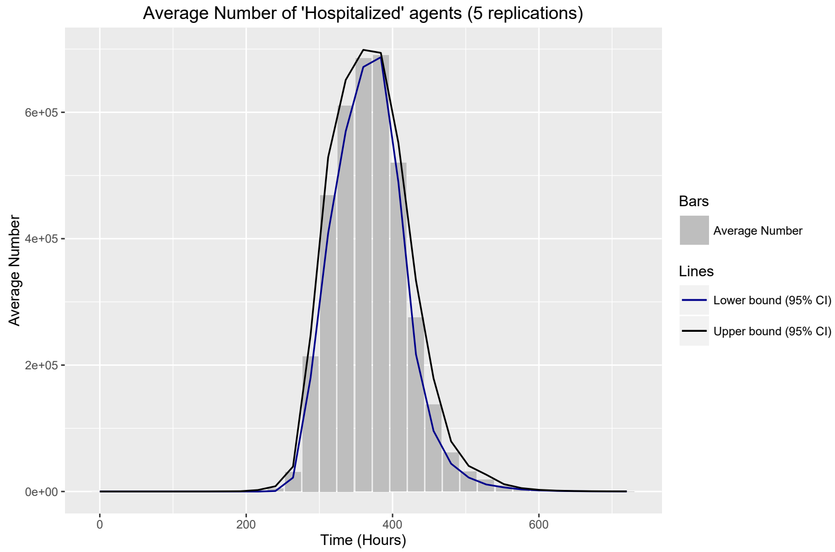


Figure 6.2: Number of '*Hospitalized*' agents in the scenario of including public transportation

Table 6.3: Tests for equality of means between baseline and including movable physical containers

		Levene's Test for Equality of Variances		**t-test for Equality of Means**			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Peak Time (Hours)	Equal variances not assumed	6.146	0.0382	-7.070	4.552	0.001	14
Peak Number	Equal variances assumed	0.837	0.387	-11.49	7.110	7.56e-06	47540

It can be seen from Table 6.3 that the p values for both the two indicators are are below 0.05 in the t-test. It means both the 'Average Peak Time' and 'Peak Number' in this scenario are different from the baseline scenario. That is, including movable physical containers in the model increases the peak number of infected agents (7.3%) in the experiment. In addition, it does shift the infecting peak to a later time (14 hours) according to the 'Mean Difference' for 'Peak Time'. This result confirms the results in the other research (COOLEY et al. 2011). The difference in 'Peak Number' is minor, which means travel contacts in movable physical containers are not the dominating factor for the disease to transmit among people in the model. Figure 5.14 proved this cause where infection rate occurred in public transportation is around 8%. Furthermore, in the baseline scenario where movable physical containers are excluded from the model, people are designed to have longer contact with their colleagues, families or social networks for the saved travel time in the model. These longer contacts cause the 'Peak Time' of disease outbreak to be slightly ahead (14 hours in the experiment) of the scenario of 'including movable physical containers'.

Studying the disease transmission through travel contacts in movable physical containers in large-scale epidemic models is not a new idea. However, most of the works in the literature uses simplified or estimated models to represent travel contacts in daily transportation. The research in GREFENSTETTE et al. (2013), PARKER and EPSTEIN (2011) uses gravity models with simplified assumptions to model the travel patterns in order to create random contacts during travel. RAKOWSKI et al. (2010) considered only the intermediate breakpoints (transfer cities) between end points (the origin and the target cities) to determine the number of co-travelers for each traveling agent during its travel. PEREZ and DRAGICEVIC (2009) modeled a transportation network to represent the movement path as a trajectory in space for disease propagation, while the disease does not propagate during the transportation. All these research works claim that these methods are sufficient to model disease spread through travel contacts to some extent, however, none of their models offers the ability to represent accurate travel contacts in movable physical containers which are important to test some specific interventions on travel controlling (e.g. shut down one metro line), while the interactions of people using public transportation in large metropolitan areas actually help spread an influenza epidemic (COOLEY et al. 2011).

According to our findings in the literature, only COOLEY et al. (ibid.) developed an agent-based model of New York city that incorporates subway ridership which simulates the actual travel contacts of subway riders and examines the impact that a severe influenza epidemic would have on NYC and the potential effects of different hypothetical subway-related disease control measures. Compared to their research, the novelty and contribution of this research are reflected in the following aspects:

- We consider the metro system and bus systems as an indivisible whole and investigate the role of the entire system on epidemic control, while only subways are studied in Cooley's model (ibid.). With this contribution, we can test more types of interventions in our model, such as closing the whole public transportation system or avoiding public transportation.
- We modeled the diversity of traveling purpose in public transportation,

while COOLEY et al. (2011) didn't account for children commuting to school on the metro system. Thus, interventions related to traveling purpose such as forbidding or discouraging students to take public transportation to schools can be tested in this model.

- There is a much more refined disease model in this research where 16 potential phases of the disease are included. This can lead to test interventions related to people in different phases. For example, the intervention of preventing the symptomatic population to board the public transport system, e.g., can be easily studied.

The study showed the relatively limited effect on epidemic outbreak, and the results comply with other research findings (ibid.). As a matter of fact, it can be seen from the scenario in this section that the model presented in this research offers more flexible opportunities for policy makers to test different interventions regarding travel, such as forbidding students to take public transportation to schools or preventing symptomatic passengers from using public transport.

6.1.3 Scenario of Including Social Networks and Social Interactions

In order to test the quality of implemented social networks and social interactions, a scenario was designed to test the impact of including dynamic social contacts during unscheduled social interactions on disease spread. In this scenario, social contacts emerge when agents participate in social activities based on dynamically generated social networks.

For this scenario, we also conducted an experiment with 5 replications. For each replication, the numbers of agents in the '*Hospitalized*' phase were recorded and the statistical results are presented in Figure 6.3.

Similar to the scenario of including public transportation, the key statistics on indicators of 'Average Peak Time' and 'Peak Number' are shown in Table 6.4.

Table 6.4: Results of '*Hospitalized*' agents when including social interactions after 5 replications

Result	Mean	95% Confidence Interval	Shapiro-Wilk Test (P Value)
Average Peak Time (Hours)	260	[252.5, 267.7]	0.5765
Peak Number	694,458	[689,520, 699,396]	0.8103

From Table 6.4, we can find both the 'Average Peak Time' and 'Peak Number' follow a normally distribution as well. To show the difference of the simulation results between the baseline scenario and the scenario including social networks and social interactions, we also did statistical tests using SPSS and present the results in Table 6.5.

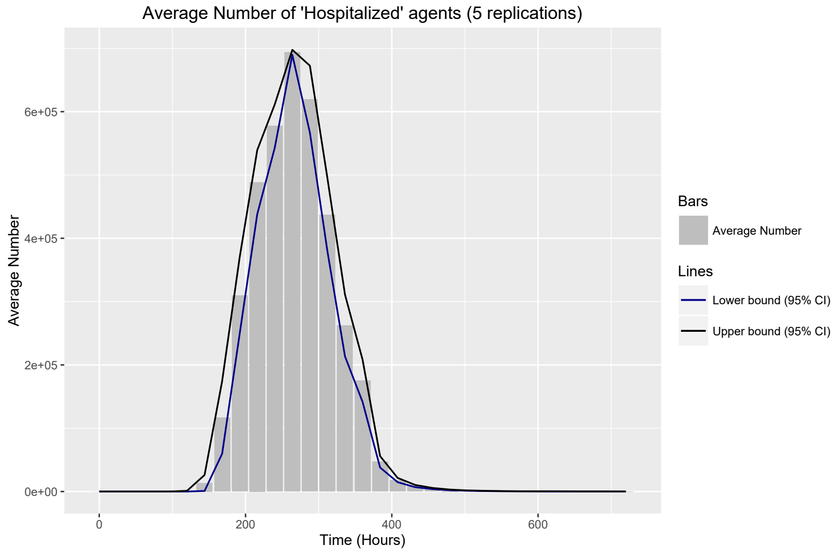


Figure 6.3: Number of '*Hospitalized*' agents in the scenario of including social interactions

Table 6.5: Tests for equality of means between baseline and including social interactions

		Levene's Test for Equality of Variances		**t-test for Equality of Means**			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Peak Time (Hours)	Equal variances not assumed	9.184	0.016	26.306	4.132	9.308e-06	102
Peak Number	Equal variances assumed	0.041	0.845	-12.146	7.962	2.036e-06	41860

From Table 6.5, we can see that the p values for both the two indicators are below 0.05 in the t-test. It means both the peak number and the total number of '*Hospitalized*' agents have a significant difference between the two scenarios. According to the 'Mean Difference' in the table, the 'Average Peak Time' is increased by 6.4% and the peak time is shifted to an earlier time (102 hours). Although the difference in the number of infections is minor, the result proves that including social interactions does have a great impact on disease outbreak in terms of peak time.

Compared to other large-scale epidemic models, the role of social interactions in disease progression is not widely explored. The most proximate research is a test of the effects of interventions such as self-isolation (Bisset et al. 2009b, Chen et al. 2010), which eliminates all activities including social activities. The difficulty comes from the design of these models where the social networks are treated as fixed attributes of individuals.

This study is also quite novel as testing the effect of interventions such as eliminating social interactions from part of the population can be easily realized due to dynamic generation of social networks and the decision capability of agents to join non-predefined social activities. This research is inspired by Ronald et al. (2012a), but there are big differences. For example, the agent in this research is pattern based which results in the capability of modeling interactions among millions of agents while agents in Ronald et al. (ibid.) are activity based and the system scale is hundreds of agents.

With the way to generate social networks and execute social activities in this research, interventions such as eliminating social interactions from part of the population or blocking a percentage of the social interactions can be easily realized.

6.2 Simulation Study on Epidemic Interventions

The most commonly-used policies for H1N1 epidemic interventions in current research are for instance, vaccination, antiviral treatment and school closure (Cao et al. 2014). The scarcity of the types interventions that can be studied relies on the existing incomplete and unrealistic models.

With the epidemic model of this research, three policies to control disease outbreak will be introduced and tested to explore the model, which are closing schools, prohibiting social activities and avoiding public transportation.

6.2.1 Simulation Scenarios

The intervention of closing the whole school system (including universities) is that all students and employees in the schools need to change their behavior when the number of agents in the city who are in the 'Hospitalized' phase reaches a threshold. A monitor in the model is created to give alerts. When the trigger is fired, a notification is fired to all targets. Agents who receive this notification either switch their activity pattern to a new preferred one or will be forced to use one special pattern specified by the social regulation (policy).

The intervention of prohibiting social activities is assumed to have a low effectiveness, as it's not compulsory. We assume that 50% of the agents will follow this regulation. The intervention is triggered by a monitor in the model as well. Agents who decide to follow the intervention will skip social activities and stay at home.

The intervention of avoiding public transportation means that the occupancy of the public transportation will decrease. When the 'hospitalized' agents exceed a certain number, the whole population will get a warning of the danger taking public transport. However, the agents can choose either to find other transportation methods or to travel as usual.

The initial condition at t=0 for all the scenarios is that 0.01% of the population (100,000 agents) are in the 'Suspect' phase. The number of people in the 'Hospitalized' phase is recorded during the simulated 30 days within the three scenarios and the results are compared to a baseline model where no interventions are introduced. Table 6.6 describes the other settings of the applied interventions.

Table 6.6: Settings of applied policies in the simulation scenario

Item	Policy	Mandatory degree	Trigger
1	Close school	1.0	0.1% population hospitalized
2	Prohibit social activity	0.5	0.1% population hospitalized
3	Avoid Public Transportation	0.5	0.1% population hospitalized

Based on this initial setting, we ran the simulations and recorded the number of agents in the phase of 'Hospitalized'.

Before conducting the experiments for the three interventions, firstly an experiment with 5 replications without interventions was conducted as a baseline, and the numbers of agents in the 'Hospitalized' phase for each replications were recorded.

6.2.2 Simulation Results and Analysis

The first tested intervention is closing schools when the number of students in the 'Hospitalized' phase is monitored to reach a predefined threshold. If the involved schools are shut down, the related students will be using another week pattern. However, the students can return to their original week pattern when the alert is lifted. This scenario puts forward a challenge to the family members of the students. For example, a parent's week pattern is required to be replaced as well when the parent is supposed to carpool the student to school. Thus, a detection and replacement for parent's week pattern is necessary after replacing the students' week pattern during the simulation.

The second tested intervention is prohibiting social activities for controlling disease spread. When this intervention is carried out, agents can decide whether to still go to the social activities jointly with friends or not.

The third tested intervention is avoiding public transportation. When the number of people in the city who are in the phase of 'Hospitalized' reach a threshold, a policy called "avoiding public transportation" is triggered. Similar to the implementation of the intervention "prohibiting social activities", a monitor is modeled to give alerts. Agents who are going to execute public transportation activities have certain degrees of preference to follow the policy or obey it. When an agent decided to follow this policy, he/she will take a TAXI (assume unlimited) to go to the destination (a duration of travel is scheduled), otherwise the agent will take the public transportation as usual.

For all these scenario, we conducted an experiment with 5 replications for each scenario. For each replication, the numbers of agents in the phase of 'Hospitalized' are recorded. The main objective of this simulation study is to study and provide an effective comparison on consequence of disease progression between the scenario of no intervention and the scenarios with interventions.

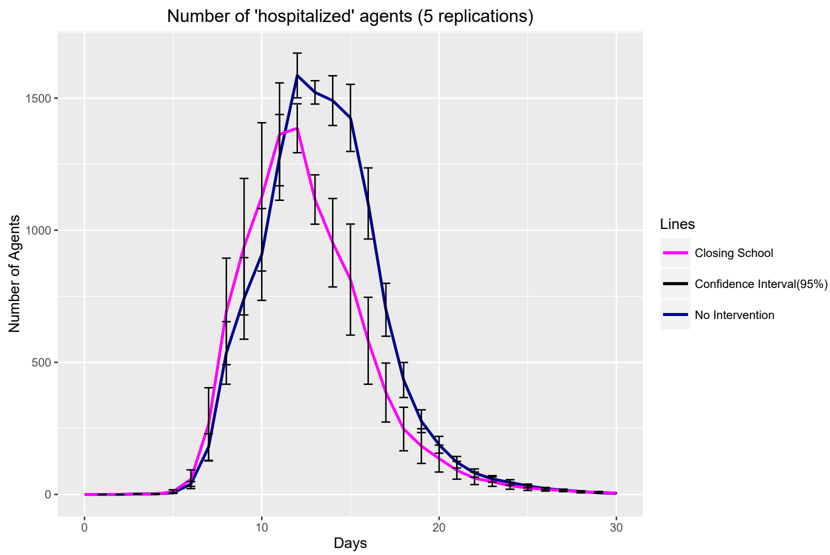


Figure 6.4: Comparison of Number of '*Hospitalized*' agents between 'closing school' and 'no intervention'

For the first intervention closing schools, the simulation results showed that the trigger was fired at simulation day 4 and the policy took effect at day 5. From Figure 6.4, we can see that the peak when using the school closure intervention is lower (13.98% in average). The simulation results also showed that the average age of infected agents is 6 years older. It means applying school closure to reduce regular contacts can effectively suppress the disease outbreak.

There are many large-scale epidemic models that can offer the ability for testing interventions of school closure (BISSET et al. 2009b, 2014, MNISZEWSKI et al.

2008, GREFENSTETTE et al. 2013, PALESHI and EVANS 2011). However, there is a fundamental difference in our model. The model in this research uses a more flexible and memory-efficient way to change the original schedule/itinerary of an agent which is induced by policies such as school closures. The models mentioned above either replace the involved activity in the schedule with another activity (for example, anyone whose activity schedule would have taken them to one of the closed locations will go home during that time instead, and they will follow their other scheduled activities as usual (MNISZEWSKI et al. 2008)), or pre-compute and keep the alternate schedules for agents in order to enable the alternate schedules to work in a consistent way on locations and duration (BISSET et al. 2014), or use a simple alternate schedule (for example, household and community contact durations of the students whose schools are closed are considered to be equal to their weekend values (PALESHI and EVANS 2011)). The model in this research is more flexible in the implementation of changing schedules if necessary due to the fact the agents are activity pattern based. Every agent will only keep an index pointing to a week/day pattern, while the behavior details such as activity location and duration are not kept in the pattern, but will be computed during the simulation runs. This design benefits on not only that agents can have many alternate schedules without much memory consumption, but also that modeler can easily test different scenarios through creating new activity patterns.

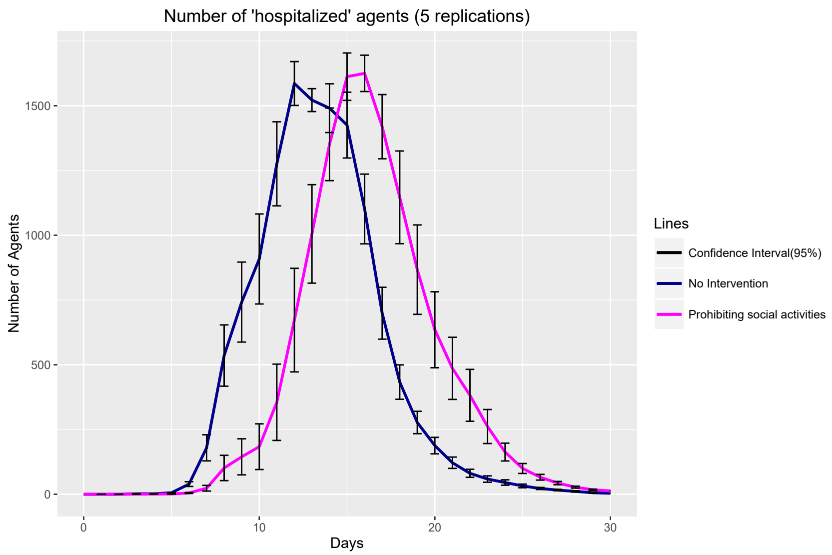


Figure 6.5: Comparison of Number of '*Hospitalized*' agents between 'prohibiting social activities' and 'no intervention'

For the intervention of prohibiting social activities, the day for the implementing the policy (day 5) is the same with other interventions as all the thresholds are set to be the same. We can find in Figure 6.5 that prohibition of social activities with friends in the model can't reduce the number of the infected agents in the long term but does shift the infecting peak to a later date (4 days in average).

Compared to the mandatory interventions such as school closure, the effect of advisory interventions such as prohibiting social activities for controlling the disease progression is not widely explored in current research on agent-based epidemic models. The most proximate intervention is self isolation (BISSET et al. 2009b, CHEN et al. 2010), which eliminates all activities. This lack of similar research is caused by the design in the models that the social networks are treated as fixed attributes of individuals, which results in the difficulties of implementing interventions of changing social interactions dynamically.

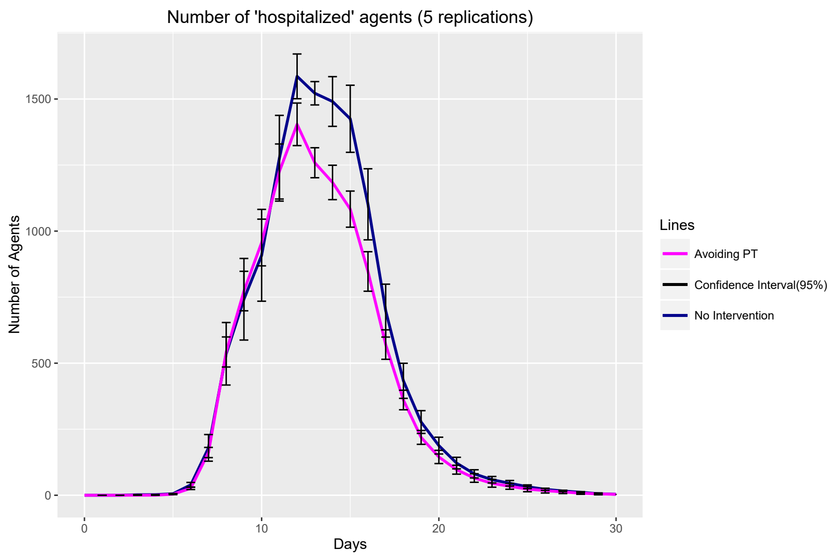


Figure 6.6: Comparison of Number of '*Hospitalized*' agents between 'avoiding public transportation' and 'no intervention'

An observation in Figure 6.6 is that the number of infected agents in the scenario of avoiding public transportation has a lower peak (11.94% in average) but without delay. The lower peak shows that fewer public transportation would decrease the infection possibility caused by random contacts in travel. The fact that the time of the peak didn't shift means travel contacts are not the dominating factor for the disease to transmit among people. This is because the disease (H1N1) we modeled in this research will spread effectively based on stable and long-time contacts while the environment and the contacts in public transportation frequently change.

To summarize, these experiments show the difference of the epidemic outbreak using different interventions through presenting the statistical results of

'Hospitalized' agents. We found that all policies introduced in the model have a positive influence to control disease spread.

The current trend in researches on influenza pandemic control has been shifting to study parameter changes or interventions one-at-a-time, with many studies focusing on a single intervention, such as school closures, mass vaccination, or the use of face masks. The results of the study by BEELER et al. (2012) with a five-factor designed experiment suggest that such an approach to pandemic modeling is fundamentally limited. The effects of pandemic control strategies may vary considerably in the presence of other control strategies, or be contingent on the epidemiological characteristics of the particular strain causing the pandemic. There may be a risk of misrepresenting or oversimplifying the effectiveness, and cost-effectiveness, of pandemic control strategies if these strategies are not evaluated in a way that considers the presence of other policies or variable pandemic characteristics. However, the study presented in this thesis performs well in terms of testing different scenarios and combining scenarios is extremely easy. In addition, shifting behavior of agents as a result of scenarios or changing compliance with government measures, e.g., decreasing compliance over time, or increasing compliance with an increasing number of infections, can be easily modeled as well.

6.3 Discussion

Two types of simulation studies were conducted in this chapter and the simulation results clearly showed the characteristics of epidemic dynamics which reflect model outcomes and model capabilities. The simulation study on model outcomes used different scenarios to show the effect of the implemented model components. The simulation study on epidemic interventions showed the impact of different interventions on controlling disease spread.

Epidemic interventions in this research are modeled as social regulations. The concept of 'Social Regulation' is designed to model norms and institutions that can guide and influence agent behavior globally. In a social regulation, a *monitor* is used to observe agents' behavior and status, analyze the results and compare with *standards*. Based on the comparison, social regulations can trigger various *Operations* to the agent society in order to regulate agents' behavior. One of the referenced implementation on *Operations* in a large-scale agent-based social simulation is to switch agents' activity patterns when the objective agents don't comply with any of the standards. With this process, agents can respond to different situations during a simulation run. For example, regulating agents' behavior during a disease outbreak is an indispensable part of a large-scale agent-based epidemic simulation. How agents would respond to a disease outbreak is a lightweight strategical level decision-making process in our reference implementation as it would have a big impact on the agent's behavior. With the implemented interventions (policies) in this chapter, research requirement RR7 is satisfied.

✓**RR7 Research Requirement - Policy in the Case Study** Large-scale epidemic prediction and control should have a flexible mechanism to model new policies with different settings.

7

Reflection on Large-Scale Agent-based Social Simulation

Chapter 6 provided several simulation scenarios to show the effect of the proposed conceptual model in the case of epidemic prediction and control. In this chapter, we will use several tests to show that several novel designs during the practical phase of the case study can also be used as general designs in large-scale agent-based social simulation studies in addition to the general conceptual model.

7.1 Agent with the 'Right' Amount of Intelligence

This thesis presents a new way to model large-scale agents with complex behaviors. An agent consists of three main parts: (1) agent object, (2) activity pattern, and (3) multi-level decision-making module. An agent object, as part of the agent architecture, is the body which is responsible for updating the agent status as the carrier, receiving, processing and forwarding input messages to the corresponding decision-making module, and enabling agents to behave according to activity patterns. For a given agent, an activity pattern specifies which behavior schedule will be conducted. Based on the activity pattern, agents will mainly perform their activities according to the initial predefined sequences of activities. However, this schedule does not specify how long, when, where and with whom these activities are to be, which is decided by the decision-making module. The decision-making module serves as the 'brain' of the agent architecture which is specially designed for a number of proposed decision-making problems. It is considered as a supplement to the agent's behavior pattern. With this design, the agent can carry out a lot of complex activities and show diverse behavior, such as traveling around, or joining non-predefined social activities.

In order to check whether the implemented agents have the 'right' amount of intelligence that can guarantee both system performance and model complexities, two simulation tests were conducted.

7.1.1 System Performance

One of the challenges to model agents with decision-making capabilities for large-scale agent-based social simulation is to guarantee the system performance, as the countless communication messages for complex reasoning and decision-making processes and the agents' large amount of local argumentation knowledge will greatly decrease the simulation performance.

The typical solution for the simulation phase in current research on large-scale ABSS is partitioning the model into parts and running different parts in a parallel and distributed setting (SIŠLÁK et al. 2009) through introducing corresponding communication middlewares (such as MPI, or RTI) and high-performance computing (BISSET et al. 2012).

As a different tryout, this research tried to gain reasonable system performance from the consideration of model design and implementation, which adopted a pattern-based design for agents' daily behavior and a multi-level mechanism for decision-making. In order to test the performance of this design, we implemented three sets of simulation experiments and recorded the execution time of the experiments when the agent population increases in different situations. The results are shown in Figure 7.1.

From Figure 7.1, we can find that the execution time is basically linear if agents only include operational level decision-making process. When the higher level decision-making processes are added to agents in the simulation, the execution time increases slightly. When all agents (19.6 million) with complete decision-making capabilities are simulated which is the worst case, it will take around 40 hours for one simulation run. It still has acceptable performance as this simulation runs for a simulation period of 30 days.

Compared to other agent architectures presented in Table 2.1, the agent architecture designed and implemented in this research satisfied all the requirements. It can be easily decomposable into components, has its own decision-making capability, supports a mechanism to sense and reflect the surrounding environment,

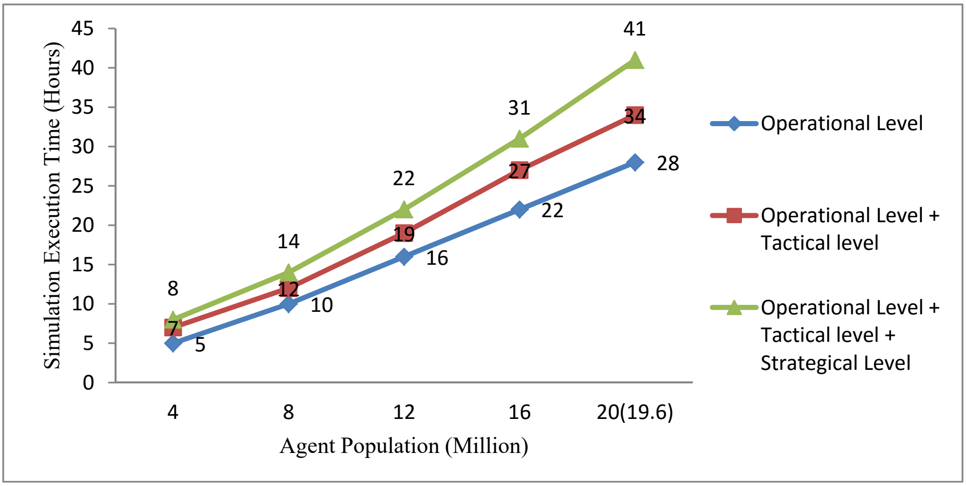


Figure 7.1: Simulation performance comparison for agents with different capabilities

has dynamical social interaction capability, has the ability to response to emergent events, and is "simple" and "small" enough.

7.1.2 Agent Capability

Section 7.1.1 presented a test to show the performance of the multi-level decision-making agents. In this section, a test is conducted to show the effect of the decision-making capability in the context of disease spread.

The capability that agents with the role of 'Student' have is to decide to stay at home when the total number of agents in the phase of 'Hospitalized' reaches a threshold. However, the preference of making this decision is various in different settings. We presented the results in Figure 7.2 to show the dynamics of the number of infected agents in the role of 'Student', where the setting 'Preference = 0' means no student agents will make the decision to stay at home, 'Preference = 0.2' and 'Preference = 0.4' mean 20% and 40% of the student agents will decide to stay at home. The population of agents is set as 10,000 in this test.

From Figure 7.2, we can find that the capability of making decisions of students does have positive impact on the spread of disease among the students themselves. To show the impact to the whole population, the results about the dynamics of the number of infected agents among the the full population are presented in Figure 7.3.

This test shows the effect of the decision-making capability for agents in the context of disease spread.

Together with the case study in Chapter 5, these two tests prove that research requirements RR2 and RR8 are satisfied.

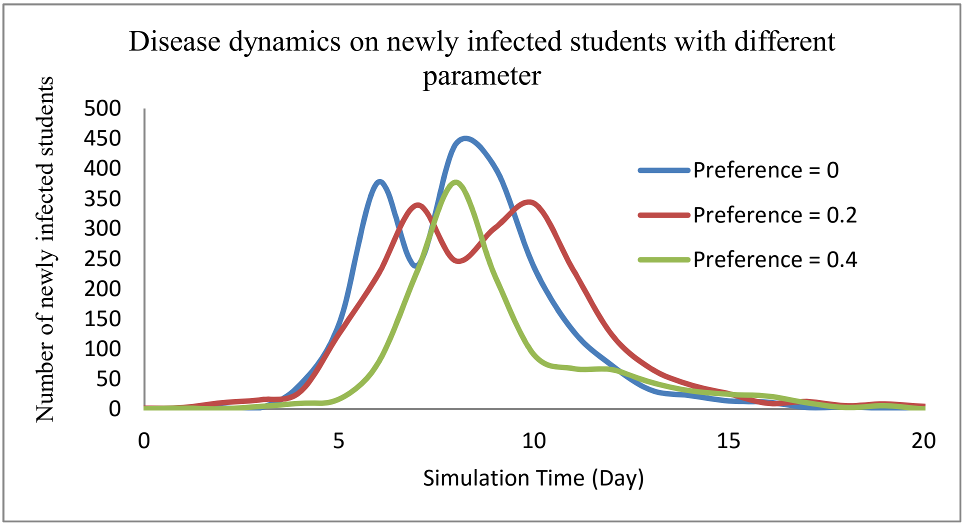


Figure 7.2: Number of newly infected students in the cases of different 'Preference'

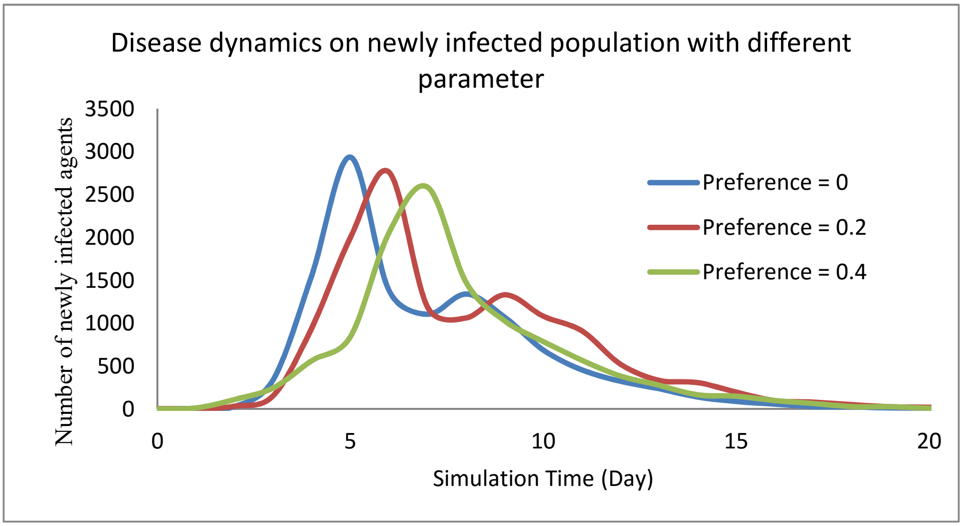


Figure 7.3: Number of newly infected agents in the cases of different 'Preference'

✓**RR2 Research Requirement - Agent Architecture** The architecture for agents in the target conceptual model for large-scale ABSS should satisfy the following requirements.

RR2.1 It should be easily decomposable into components.

RR2.2 It should have its own decision-making capability.

RR2.3 The architecture should support a mechanism to sense and reflect the surrounding environment.

RR2.4 It should have dynamic social interaction capability.

RR2.5 It should have the ability to respond to emergent events, rather than having to blindly follow a prearranged plan.

RR2.6 The agents should be "simple" and "small".

✓**RR8 Research Requirement - Agents in the Case Study** Agents in the large-scale epidemic prediction should have capabilities to make complex decisions to some extent during disease outbreaks.

7.2 Agent Environment

The agent environment in the proposed conceptual model is separated into three concepts: physical container, social regulation and functional entity. In the case of epidemic prediction and control, social regulations are implemented as interventions to control disease spread, physical containers are implemented as activity locations and the disease is considered as a functional entity.

In this section, we will run several tests to show that not only these concepts can be generalized for other social systems, but also some of the used algorithms can contribute to other systems.

7.2.1 Three-level cache system of Physical Containers

This thesis introduces a new concept of a hierarchical 'Physical Container' to represent the physical environment where agents stay. Typical physical containers are school, classroom, office, bedroom, and train, etc. Physical containers are organized hierarchically. Each physical container can be partitioned into sub-physical containers. Examples are classrooms in a school, stores in a shopping mall, or offices in a working place. Agents can have different forms of contacts when they are in the different level of physical contain hierarchy. In addition, physical containers show "behavior" just like agents. For example, an important "behavior" 'Calculate Distance' should be implemented in the physical container, which calculates the distance between two physical containers based on the GIS coordinate information (e.g., latitude and longitude).

Besides the large number of agents, the challenge of large-scale agent-based social simulation also arises from the large number of physical containers. In the context of the case study in this research, there are 8 million static locations in the city of Beijing. A three-level cache system was proposed in Chapter 4 (pp 61-62) to organize the millions of locations during a simulation run.

To show the effect of this cache system on the simulation, we made a comparison between two experiments. The first experiment is to measure the simulation execution time for a simulation period of one week (7 days) when the cache system is adopted, while the other experiment excluded the cache system. Other initial conditions are set to be the same, for example, the agents are the same in all properties and decision-making capabilities in both experiments. The results for different agent populations (from 1000 to 5000) are shown in Figure 7.4.

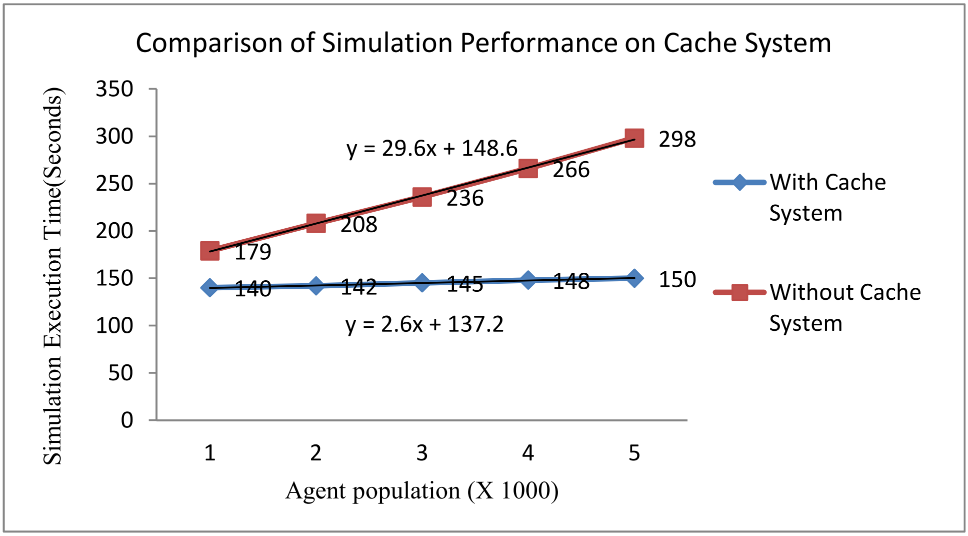


Figure 7.4: Performance comparison in simulation with and without the use of cache system for a week

From Figure 7.4, we can find that the execution time increases linearly in both cases. However, the execution time in the situation without cache system is ten times longer than the situation with cache system in all the 5 settings, where we exclude the initial set-up time of the simulation run. Since no epidemics are involved in this test, this cache system is supposed to benefit other large-scale social system research, as well.

7.2.2 '*GridZones*' Algorithm

Besides an effective mechanism to organize physical containers in large-scale systems, the concept to separate physical containers from the general agent environment concept makes it much easier to include a transportation component in a social simulation model. This is achieved by considering vehicles as movable physical containers in the model.

With the consideration of transporters as movable physical containers, this research included a public transportation component into an artificial city model, where 19.6 million agents can realistically travel through this component for daily commuting purpose.

An algorithm called '*GridZones*' was proposed in Section 4.3.2.1 in order to optimize agents' routing processes. To evaluate this algorithm, a test is conducted by measuring the simulation performance between two situations when this algorithm is adopted or not. The simulation execution time is recorded for both situations when the number of agents increase from 1000 to 5000 for a simulation period of a week (7 days). The comparison is shown in Figure 7.5.

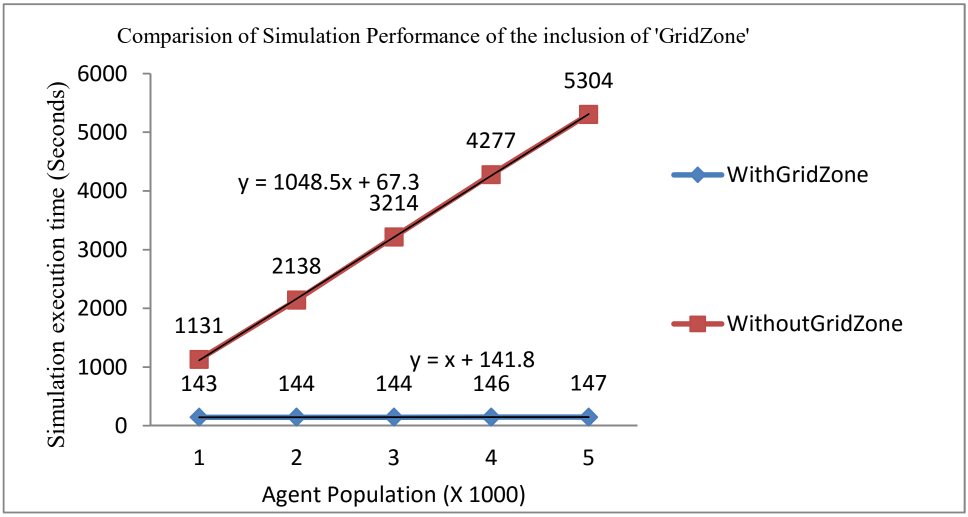


Figure 7.5: Performance comparison for in simulation with and without the use of 'gridzone' algorithm for a week

From Figure 7.5, we can find that the simulation execution time increases linearly in both of the two situations, and the speedup of the situation with the '*GridZones*' algorithm is enormous when the agent population increases. In addition to the performance optimization, the other benefit are reflected on the efficiency of memory usage. With the inclusion of '*GridZones*', all the connecting information for the bus/metro stops inside the zone are gathered together and stored in this '*GridZones*'. This design avoids the redundancy when the connecting information is separated into different stops, and ensures the updating efficiency when one of the stops in the traffic graph is closed.

7.2.3 Including Weather as Functional Entity

In the case of epidemic prediction and control, the disease is modeled as a functional entity which can influence the behavior of agents directly. Considering the generalization of the concept of functional entity, more objects that can influence either the behavior of agents or other entities in the social system can be modeled as functional entities. This feature is shown by designing a test where the weather condition is included in the epidemic prediction system.

This test models the daily temperature as a component in the weather condition in Beijing as a functional entity to show how the fluctuating of temperature would influence the disease dynamics. The test designs a situation where the fluctuating temperature in Beijing are applied to a disease spread model. The result of the number of new infected agents in this situation is compared to another situation where the temperature is constant. The comparison is shown in Figure 7.6.

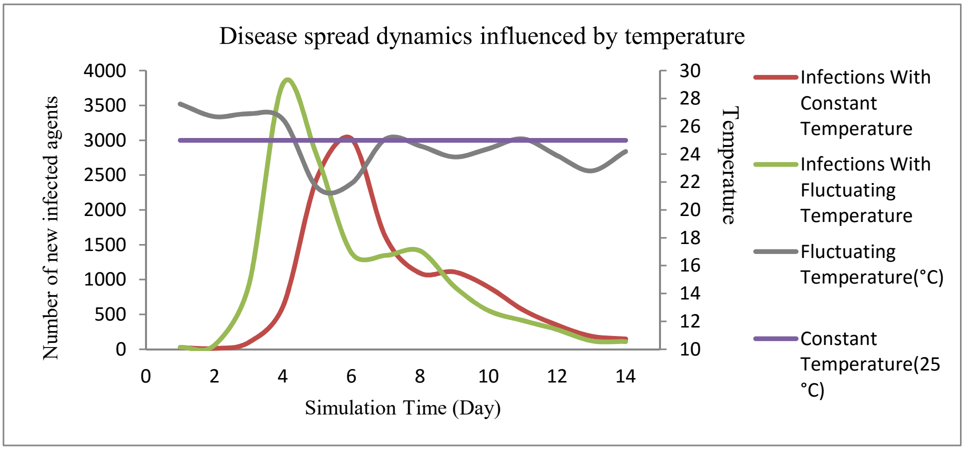


Figure 7.6: Disease spread dynamics influenced by the temperature

From Figure 7.6, we can find that the temperature has a positive effect for disease transmission. The result of the effect will not be validated as the aim of this test is to show the easy use of the concept of functional entity in a general social simulation.

With these three tests, research requirement RR5 is satisfied.

✓**RR5 Research Requirement - Agent Environment** The agent environment in the conceptual model for large-scale ABSS should be separated into different concepts to represent spatial information together with other independent artifacts and social concepts that can influence agent behavior.

7.3 Dynamic Social Networks and Social Interaction

One of the innovations for agents realized in this research is the capability to respond reasonably to unscheduled social activities based on dynamically generated social networks including family, classmates/colleagues and friends.

Since the family and classmates/colleagues are intuitively generated by checking the other agents who stay in the same living and studying/working places, defined in the input source data, we only focus on the generation of friends and interaction with friends.

The friends in the dynamically generated social networks are based on the proposed concept 'social similarity' which is inspired from the concept of social reach in the social circle model and the aggregation utility between two agents (NAVARRO et al. 2011).

To show the effect of this design for general social systems, we constructed a simulation for the large-scale artificial city model (without disease outbreak) and ran it for a simulation period of 30 days.

The parameters in this experiment are initialized using the data from Table 7.1. Since the four factors (age, social role, family role and the number of friends) are considered to be equally weighted to generate a friendship link, the corresponding weight coefficients (μ_a,μ_s, μ_f, μ_n) are calculated according to boundary conditions, which is to enable the result$S(I_i, I_j)$ to be constrained between 0 and 1. 1 means they are fully connected while 0 means that they have no relations. The other parameters are initialized as one possible experimental setting and the sensitivity of them will not be discussed in this research as the aim of this test is to show how the dynamic social networks work in the context of a large-scale artificial city model.

Table 7.1: Parameter initialization for social activity participation analysis

Item	Value	Description
μ_a	2.268×10^{-5}	weight coefficient
μ_s	1.563×10^{-2}	weight coefficient
μ_f	0.0625	weight coefficient
μ_n	2.5×10^{-5}	weight coefficient
$A(d,E)$	1	interest degree
t	0.25	attendance threshold

Based on this initial setting, agents' friends can be generated when 'FriendsSynchronizedActivity' is scheduled during a simulation run. The number of agents' friends is assigned to agents by the algorithm in Section 4.4.1 which follows a power-law distribution (HAMILL and GILBERT 2010) with an average number of 13.

Together with the family and the classmates/colleagues network, agents' social networks are formed. However, agents will only generate their social networks when they need execute social activities.

Agents, who receive invitations from their friends for attending social activities which are unscheduled in their activity patterns, can make interactions with the organizing agents in order to make a final decision.

Table 7.2 shows the average distribution of agents' decisions on a new family social activity after executing the processes explained in Section 4.4.2. The equation-based process and decision tree-based process are the processes after which agents receive an activity proposal.

Table 7.2: Distribution of agents' decisions on family social activities

Decisions	Equation based Process	Decision Tree based Process
Accept	0.78	0.67
Decline	0.22	0.33

From Table 7.2, it can be found that 33% of agents decide to decline the invitation after the decision tree process.

Similar to Table 7.2, Table 7.3 shows the average distribution of agents' decisions on a new colleague/classmate social activity. The biggest difference between the figures is that more agents are willing to participate in a colleague/-classmate social activity than in a family social activity. This is because colleague/-classmate social activities are often scheduled during the time when there are no conflicts in the agents' schedule.

Table 7.3: Distribution of agents' decisions on colleague/classmate social activities

Decisions	Equation based Process	Decision Tree based Process
Accept	0.88	0.75
Decline	0.12	0.25

Table 7.4 shows the average distribution of agents' decisions on a new social activity after executing the planning processes. Compared with the other two figures, the unusual aspect of the figure is that fewer agents accept the new proposal. This demonstrates that the composition of members in a friendship network can be heterogeneous in terms of daily schedules.

Table 7.4: Distribution of agents' decisions on friends social activities

Decisions	Equation based Process	Decision Tree based Process
Accept	0.67	0.54
Decline	0.33	0.46

Figure 7.7 shows how the number of agents in different social activities evolved during a typical week day.

For the family social activity, there are three peaks in Figure 7.7 and it reaches the highest point in the evening. This indicates that people are more willing to

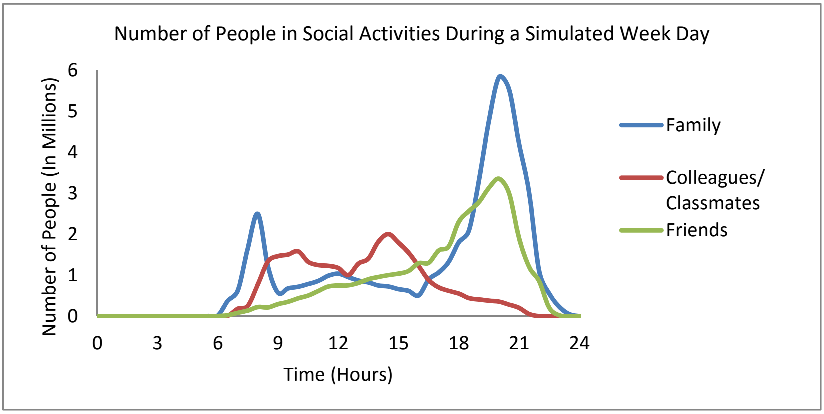


Figure 7.7: Number of people in social activities

plan activities with their families. For colleagues or students, there are two peaks in the morning and the afternoon. This is caused by the day patterns where working people have to attend meetings in the morning and afternoon and students attend joint sport activities in the afternoon. For the friend activity, it seems that most friends will only meet in the evening, to have dinner, go shopping or go to cinemas together. This phenomenon can be verified by the fact that Chinese people are more wiling to have joint dinner as a social interaction (HORIZONKEY 2007). However, these results can't be well validated in this research as no independent data exists at this moment.

As a conclusion, this thesis provides a new method to generate social networks dynamically for interactions among a group of agents on a large scale during a simulation run. Three types of social networks are dynamically generated in the model, which are family, colleagues/ classmates and friendships. Family and colleagues/ classmates are intuitively generated by checking the other agents who are in the same living and working physical containers, while friendships are generated dynamically. This thesis learns from the concept of 'social reach' in a social circle model, and proposes the concept of 'social similarity' to generate the special type of social networks-friendship. Based on the entire social network, agents in this model are able to communicate for scheduling joint social activities. When executing joint social activities, a functional entity called 'activity group' is generated to organize and manage the participants and the social contact network emerges from the execution.

With the simulation test in this section, research requirements RR3 and RR4 are satisfied.

✓**RR3 Research Requirement - Social Networks** The conceptual model for large-scale ABSS should support modelers to generate dynamic social networks.

✓**RR4 Research Requirement - Social Interaction** The conceptual model for large-scale ABSS should support agents to join in different types of social activities through complex social interactions.

7.4 The Hybrid Conceptual Model and Event Scheduling Mechanism

As we mentioned in Chapter 4, the proposed conceptual model in this research borrows concepts from both general agent-based models and discrete event simulation models. Thus, it is worth to mention that the simulation studies in this research are based on the commonly used event scheduling mechanism in discrete event system simulation (ZEIGLER et al. 2000, NANCE 1996), which is contrary to the commonly used time-step mechanism in agent-based modeling. This choice is mainly based on the consideration of system performance. The traffic component is included in the model in this research which makes using fixed time steps difficult for scalability. Small steps (e.g. seconds) are usually required for a microscopic transportation model (SCHWERDTFEGER 1984) which would generate a large number of events in the model. Suppose the time step of an epidemic prediction model with the time-step mechanism can be set as 1 minute, the resulted number of events for each agent is at least 1440 per simulation day, while the event scheduling mechanism only generates around 10 events per day for each agent. Since the number of events for each agent in the event scheduling mechanism depends on the number of activities in its activity pattern, this mechanism is more efficient than the time step mechanism in this case.

However, parallel simulation (multiprocessors and distributed memory) (FUJIMOTO 1999) is used to increase system performance in large-scale agent-based simulation with the time step mechanism (COLLIER and NORTH 2012). The difficulty in parallel simulation is to preserve causality constraints among events (LIU et al. 2001), while it can be easily solved in the time step driven simulation as the synchronization messages with time stamps can strictly prohibit out-of-order execution of events (CHANDY and MISRA 1979).

Compared to the time step driven simulation, it is more difficult for event driven simulation (event scheduling mechanism) to execute the model in parallel. The difficulty mainly comes from model partitioning and time synchronization (CELIK et al. 2012). Furthermore, it often turns out that parallel programs run surprisingly slow as the model is tightly-coupled (SCHWARZ 1995). The model implemented in this research can be considered as tightly-coupled due to the dynamic social interaction and the microscopic public transportation. Even without the dynamic social interaction, the parallel execution in the independent research with same original data by CHEN et al. (2014) has not been applied in the further research (CHEN et al. 2015) as the system performance did not improve in their experiments.

7.5 Conclusion

This Chapter presented several tests and results to show the effect of the design and algorithms for the case study for large-scale epidemic prediction and control, and how other large-scale social systems can benefit from this design and these algorithms. As a conclusion, we list the key design components and algorithms in this research that can contribute to other large-scale agent-based social simulation systems.

- Activity pattern for agent behavior and multi-level decision-making for agent architecture. The activity pattern-based design made the agents 'simple' and 'small' enough, which causes the whole system to be scalable. On the other hand, the multi-level decision-making agent architecture enable agents with the same activity patterns to show heterogeneous behavior.
- Three-level cache system for physical container management. A key to gain good performance for large-scale agent-based systems is optimizing the method to store and to retrieve objects in the system. For objects such as physical containers in this research, a three-level cache system was designed to manage all physical containers. This algorithm shows great performance and can serve in many similar systems.
- 'GridZone' algorithm for shortest path calculation. With the inclusion of a public transportation component, the imposed challenge by the large number of commuting request was solved by the 'GridZone' algorithm. It can be applied for other social systems containing a travel component.
- Functional entity for changing agent behavior or system state globally. The separation of the functional entity from the agent environment was used to model other artifacts in the social systems that can influence agents' behavior or the system state globally.
- Dynamic generation of social networks for efficient memory consumption and realistic friendship evolvement. A fixed social network is sometime a bottleneck for large-scale social systems. Thus, the 'social similarity' concept was proposed to generate social networks dynamically.
- The difficulty caused by the dynamic generation of social networks is the execution of unscheduled joint social activities. The 'ActivityGroup' concept was proposed to enable the complex social interactions during the execution process.
- The event scheduling mechanism for reducing unnecessary synchronization messages, which increased the simulation performance if the model runs on a single processor.

8
Epilogue

Large-scale agent-based social simulation is gradually proving to be a versatile methodological approach for studying human societies, which could make contributions ranging from policy making in social science, to distributed artificial intelligence and agent technology in computer science, and to theory and modeling practice in computer simulation systems. Simultaneously, the focused application areas of large-scale agent-based social simulation vary a lot as well, from daily transportation in a city/country level, to large-scale emergency response and to prediction of social change and analysis of social structure.

However, large-scale agent-based social simulation is facing difficulties in balancing model complexity and simulation performance. The widely use of distributed/parallel mechanism in current large-scale agent-based social simulation has proven to be an efficient solution to achieve system performance and scalability. The consequence, on the other hand, is usually a severe simplification of the model including agent behavior, agent environment and the social networks and interactions, which are proven to be important to understand complex social systems.

Based on the existing challenges, this thesis introduces a novel conceptual framework for large-scale agent-based social simulation model development, provided the reference implementation of the proposed model components, and presents a simulation study of the case of epidemic prediction and control in the city of Beijing.

8.1 Research Findings

The main contributions of this thesis are given below. With these contributions, the research requirements presented in Chapter 2 and 3 are met.

1. This thesis introduces a new conceptual framework to build large-scale agent-based social simulation models, which separates the concept of agent environment into a physical container, a social regulation and a functional entity. Compared to the model concepts in the general ABM conceptual model, this separation overcomes the limitations of environmental completeness in other ABM models and provides flexibility in simulating different system scenarios. In other words, the concepts are also refined. For example, physical containers can be movable to represent transportation vehicles which are difficult to implement in general ABM platforms. Moreover, theories and concepts on social regulation from artificial intelligence can be easily implemented and integrated into an agent-based model while showing reasonable performance. Inspired by the concept of Functional Model in object-oriented conceptual models, the concept of functional entity is separated from the classical ABM conceptual model and introduced in our proposed conceptual model for large-scale ABSS.
2. This thesis presents a new way to model large-scale agents with complex behavior. An agent consists of three main parts: (1) agent object, (2) activity pattern, and (3) multi-level decision-making module. The agent object, as part of the agent architecture, is the body which is responsible for updating the agent status as the carrier, receiving, processing and forwarding input messages to corresponding decision-making module, and enabling agents to behave according to activity patterns. For a given agent, an activity pattern specifies which behavior schedule will be conducted. Based on the activity pattern, agents will mainly perform their activities according to the initial predefined sequences of activities. However, this schedule does not specify how long, when, where and with whom these activities take place, which is decided by the decision-making module. The decision-making module serves as the 'brain' of the agent architecture which is specially designed for the proposed decision-making problems. It is considered as a supplement to the agent's behavior pattern. With this design, the agent can carry out many complex activities and show diverse behavior, such as traveling around, or joining non-predefined social activities.
3. This thesis presents a new method to generate social networks dynamically for interactions among a group of agents on a large scale during a simulation run. Three types of social networks are dynamically generated in the model, which are family, colleagues/ classmates and friendships. Family and colleagues/ classmates are intuitively generated by checking the other agents who are in the same living and working physical containers, while friendships are generated dynamically. This thesis borrows from the concept of

'social reach' in a social circle model, and proposes the concept of 'social similarity' to generate the special type of social networks-friendship. Based on the entire social network, agents in this model are able to communicate for scheduling joint social activities. When executing joint social activities, a functional entity called 'activity group' is generated to organize and manage the participants and the social contact network emerges from the execution.

4. This thesis introduces a new concept of 'Physical Container' to represent the physical environment where agents stay. Typical physical containers are school, classroom, office, bedroom, train, etc. Physical containers are organized hierarchically. Each physical container can be partitioned into sub-physical containers. Examples are classrooms in a school, stores in a shopping mall, or offices in a working place. Agents can have different forms of contacts when they are in a different level of the physical container hierarchy. In addition, physical containers show "behavior" just like agents. For example, an important "behavior" 'Calculate Distance' should be implemented in the physical container, which calculates the distance between two physical containers based on the GIS coordinate information (e.g., latitude and longitude). Besides an effective mechanism to organize physical containers in large-scale systems, the fact that physical container is separated from the general agent environment concept makes it much easier to include a transportation component in a social simulation model. This is achieved by considering vehicles as movable physical containers in the model.
5. This thesis introduces a new concept of 'Social Regulation' which is designed to model norms and institutions that can guide and influence agent behavior globally. In a social regulation, A *monitor* is used to observe agents' behavior and status, analyze the results and compare with *standards*. Based on the comparison, social regulations can trigger various *Operations* to the agent society in order to regulate agents' behavior. One of the reference implementation on *Operations* in a large-scale agent-based social simulation is to switch agents' activity patterns when the agents don't comply with any of the standards. With this process, agents can respond to different situations during a simulation run. For example, regulating agents' behavior during a disease outbreak is an indispensable part of a large-scale agent-based epidemic simulation. How agents would respond to a disease outbreak is a lightweight strategical level decision-making process as it would have a big impact on the agent's behavior.
6. This thesis introduces a new concept of 'Functional Entity'. Functional entities are those extra objects in the system that can influence or directly change attributes of either agents, physical containers or social regulations. For example, a disease is modeled as a functional entity to change the agents' healthy status. A storm is modeled as a functional entity to change the temperature of a room (physical container).
   With the research findings 1-6, Research Question 1 is answered.

✓**Research Question 1** What conceptual model and model concepts can support large-scale agent-based social simulation?

7. This thesis presents a model of a large-scale artificial city of Beijing, on which to test policies for controlling the spread of disease among the full population (19.6 million). This is used as a case study to test the proposed large-scale agent-based social conceptual model. Firstly, by combining diverse data sets, including generated census-based data, open source maps, activity patterns, an artificial city with a large population was constructed. In this artificial city, each of the 8 million physical locations and 19.6 million citizens was modeled. A microscopic public transport system (subways and buses) together with a predicted road traffic system are simulated in the artificial city and are well integrated with the daily activities of the population. With this model, spatial contact networks emerge and can be observed during the execution of the model. Secondly, to investigate the effect of the emerging spatial contact network for epidemic prediction, a pandemic influenza disease progression model was implemented. At last, to test the quality and explore the possibility of the implemented large-scale agent-based model for epidemic prediction and control, different simulation scenarios were conducted and simulation results were analyzed.
   With this research finding, Research Question 2 and 3 are answered.

   ✓**Research Question 2** How can the components of this conceptual model be implemented in the case of epidemic prediction and control?

   ✓**Research Question 3** How can the case of epidemic prediction and control benefit from the proposed conceptual model regarding of model outcomes and system performance?

8. This thesis can be considered as a proof of concept which exemplifies how large-scale social systems with complex human behaviors and social interactions can be modeled with the help of the proposed conceptual framework, it also indicates potential use in other social science areas, such as microscopic transportation systems and city level evacuation planning.
   With this research finding, Research Question 4 is answered.

   ✓**Research Question 4** How can large-scale agent-based social simulation benefit from the case of epidemic prediction and control in this research?

8.2 Future Research

As for future research, many extra efforts are required to refine the proposed conceptual model and the referenced implementation of model components. A number of first targets for additional research in the case of epidemic prediction and control are:

1. *Activity pattern generation.* The design of a full set of activity patterns is important to model agent behavior in the proposed implementation for the agent component. The current method to generate activity pattern requires a large effort on surveying people's behavior and abstracting them into a set of patterns. Thus, an activity pattern generator can be an interesting direction to explore which considers factors such as culture, history, law, etc. Another direction is to build a repository of activity patterns for different cities, countries and areas.
2. *Decision-making capability.* Three levels of decision-making capability for agents are defined to solve different levels of behavior dynamics problems when processing different types of events. This thesis implemented several specific decision-making problems, mainly in the lower level (e.g., operational level), while more situations and solutions should be studied to progressively fulfill the higher level, such as strategical level.
3. *Realistic environmental data.* Currently there are 18 types of static physical containers that can not cover all the types in Beijing, for example, small shops and cinemas are missing. Further research should be conducted on generating or collecting real data for these missing types.
4. *Expert validation.* The simulated results of the disease spread scenario in this thesis have not been well validated due to the missing of available date. An expert validation process is required as part of future research.

Regardless of the case in this thesis, there are several aspects that should be researched further with the proposed conceptual model for large-scale agent-based social simulation:

1. *Parallel execution.* This thesis presents a way to guarantee the scalability issue from the aspect of model design. However, the traditional way to model large-scale models that adopts a parallel/distributed execution mechanism should also be studied in the proposed architecture, which can make current model execution much faster (or slower).
2. *More application areas.* This thesis uses the case of large-scale epidemic prediction and control to test the proposed conceptual model for large-scale agent-based social simulation model development. More cases need to be studied in other application areas to improve the versatility and applicability of the proposed conceptual model.

Appendices

A

Contact Dynamics in Different Location Types

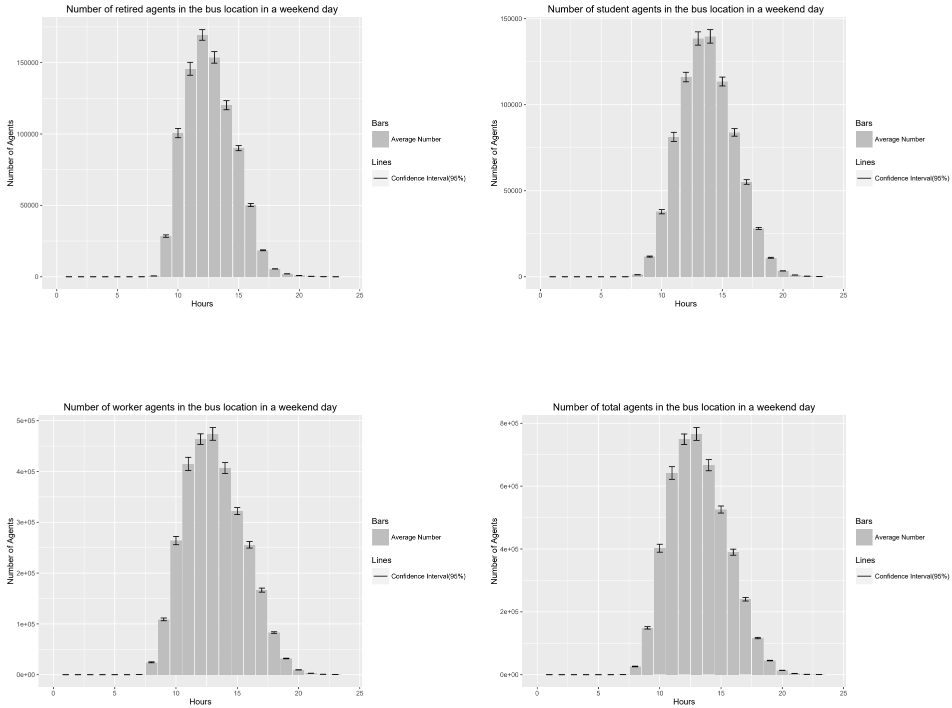


Figure A.1: Statistics of the number of agents in the bus location in a weekend day (10 replications)

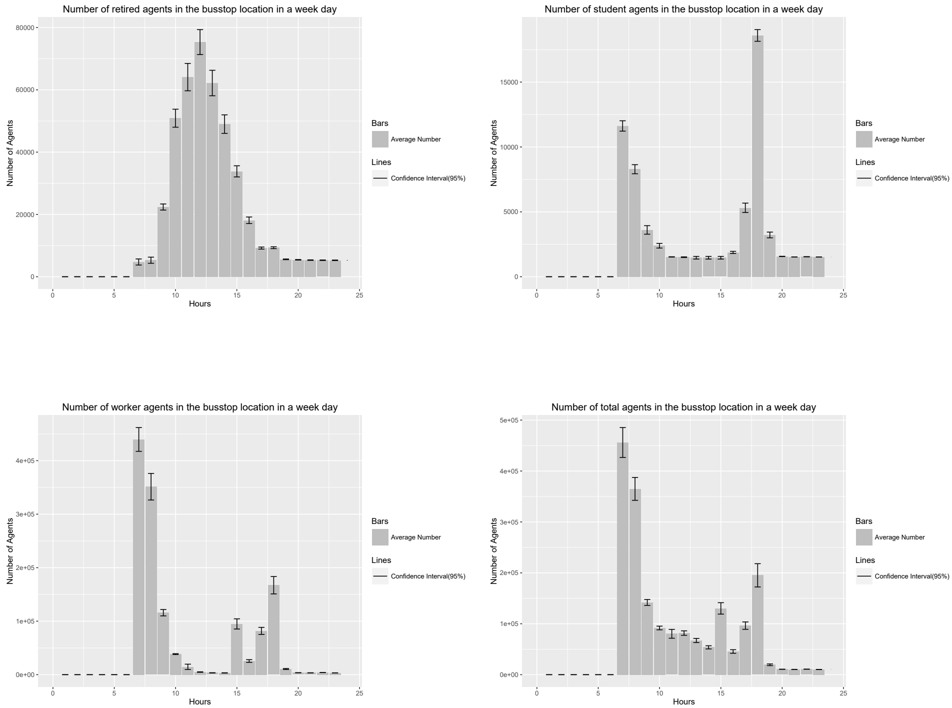


Figure A.2: Statistics of the number of agents in the bustop location in a weekday (10 replications)

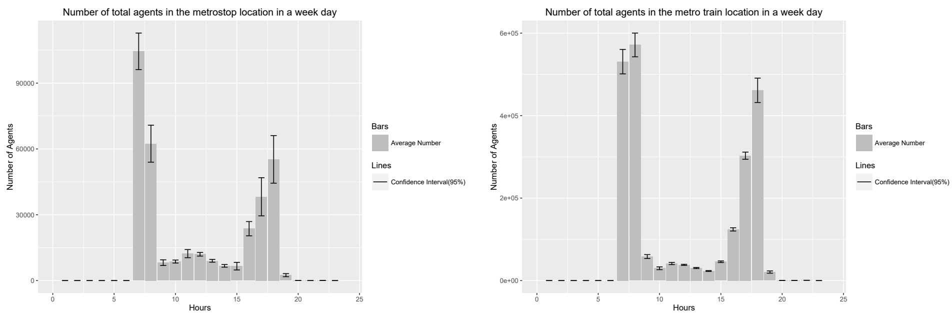


Figure A.3: Statistics of the number of agents in the metro system location in a weekday (10 replications)

Figure A.4: Statistics of the number of agents in the car location in a weekday (10 replications)

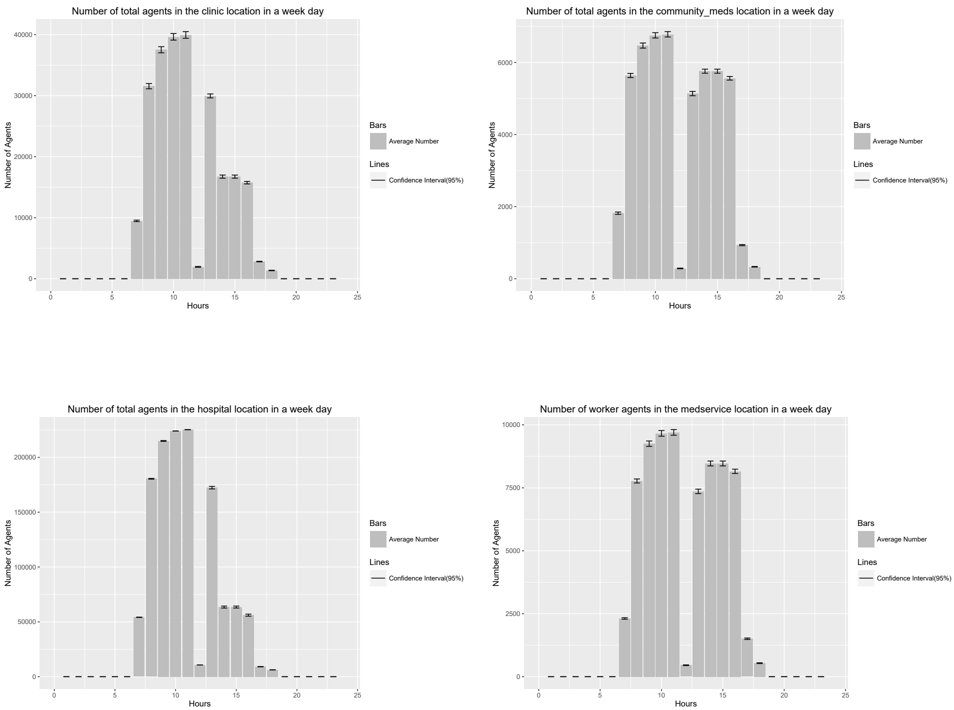


Figure A.5: Statistics of the number of agents in the medical locations in a weekday (10 replications)

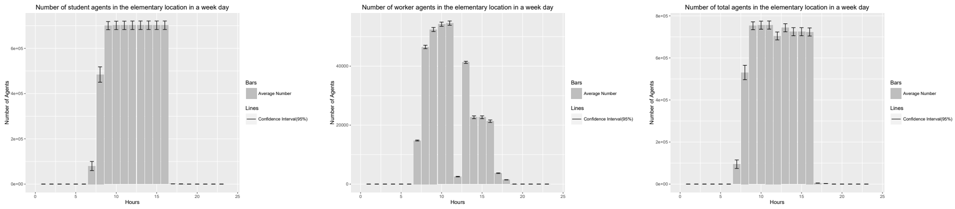


Figure A.6: Statistics of the number of agents in the elementary location in a weekday (10 replications)

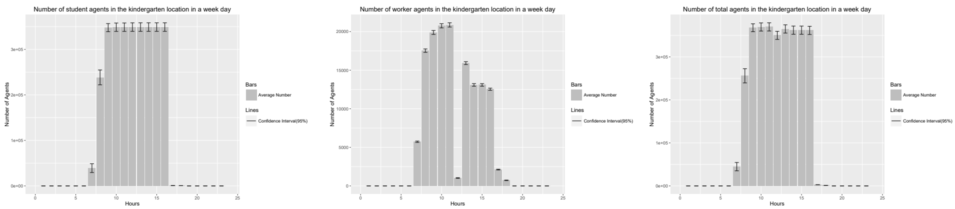


Figure A.7: Statistics of the number of agents in the kindergarten location in a weekday (10 replications)

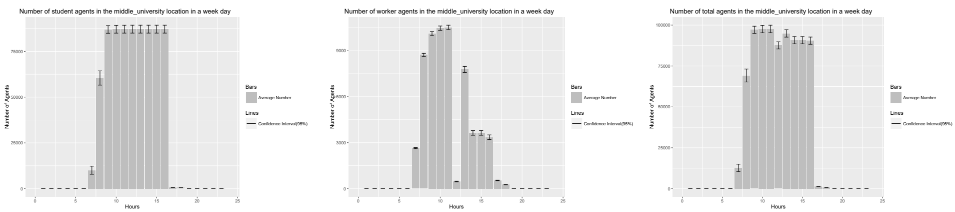


Figure A.8: Statistics of the number of agents in the middle_university location in a weekday (10 replications)

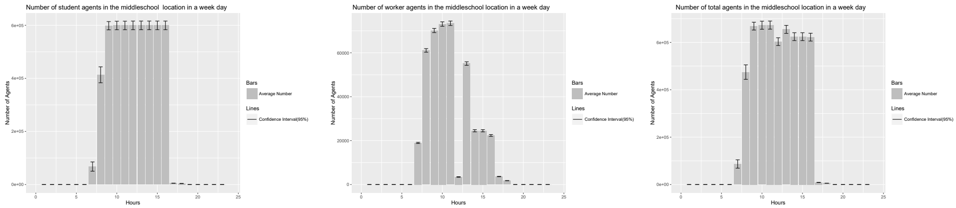


Figure A.9: Statistics of the number of agents in the middleschool location in a weekday (10 replications)

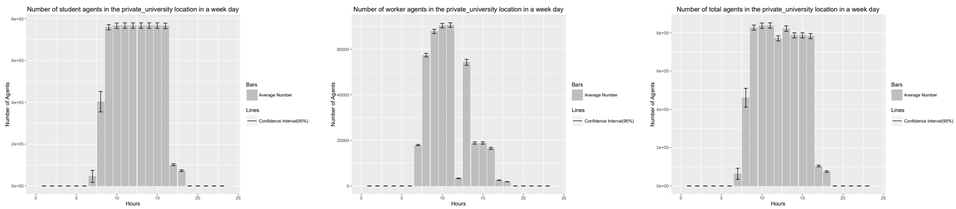


Figure A.10: Statistics of the number of agents in the private_university location in a weekday (10 replications)

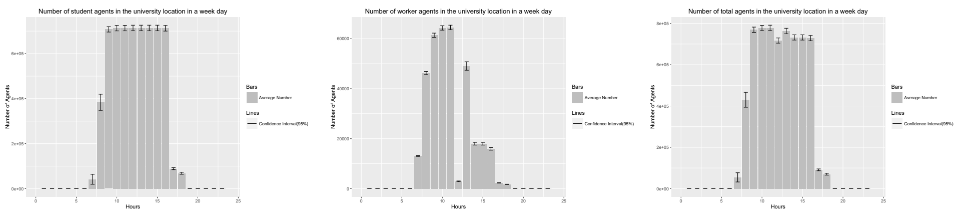


Figure A.11: Statistics of the number of agents in the university location in a weekday (10 replications)

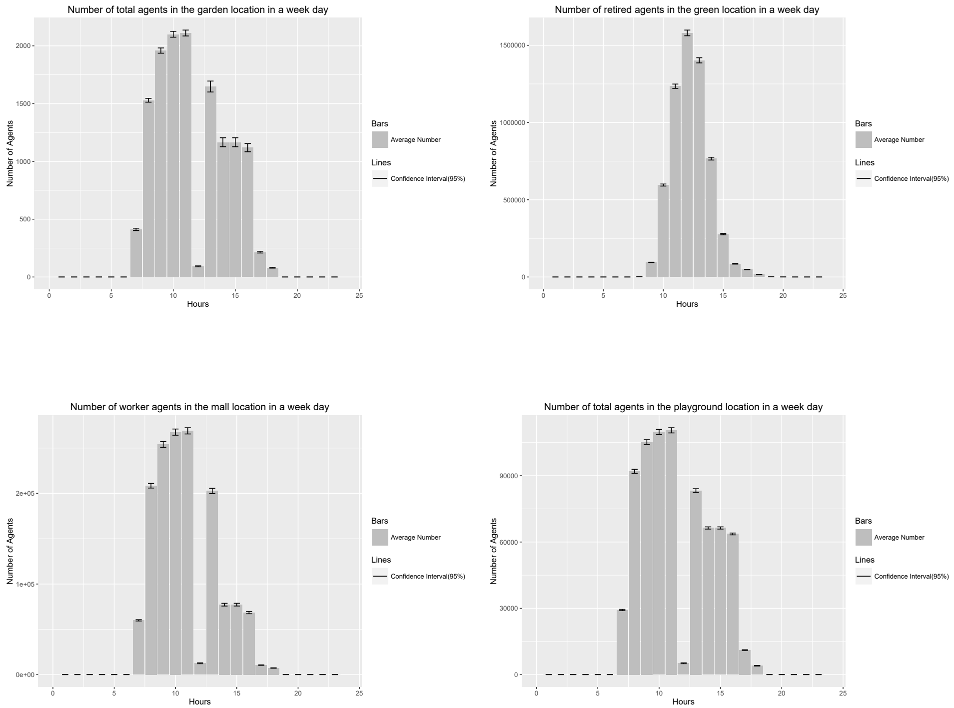


Figure A.12: Statistics of the number of agents in the entertainment locations in a weekday (10 replications)

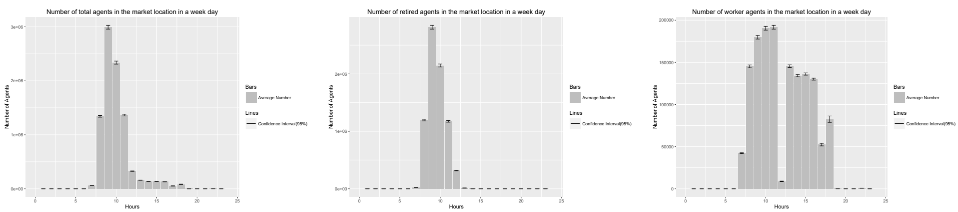


Figure A.13: Statistics of the number of agents in the market location in a weekday (10 replications)

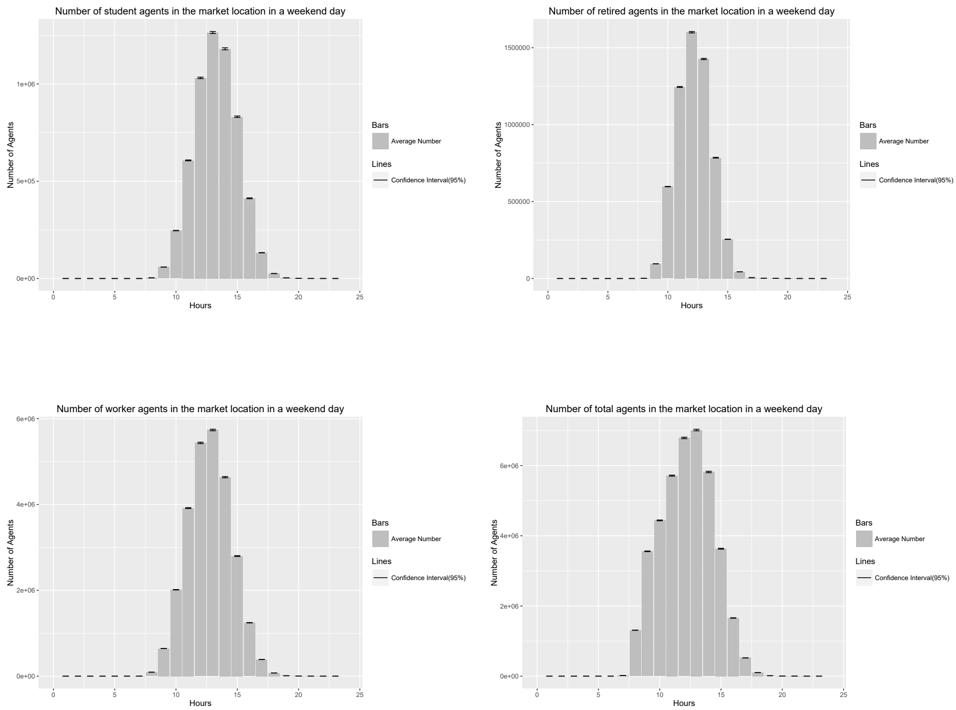


Figure A.14: Statistics of the number of agents in the market location in a weekend day (10 replications)

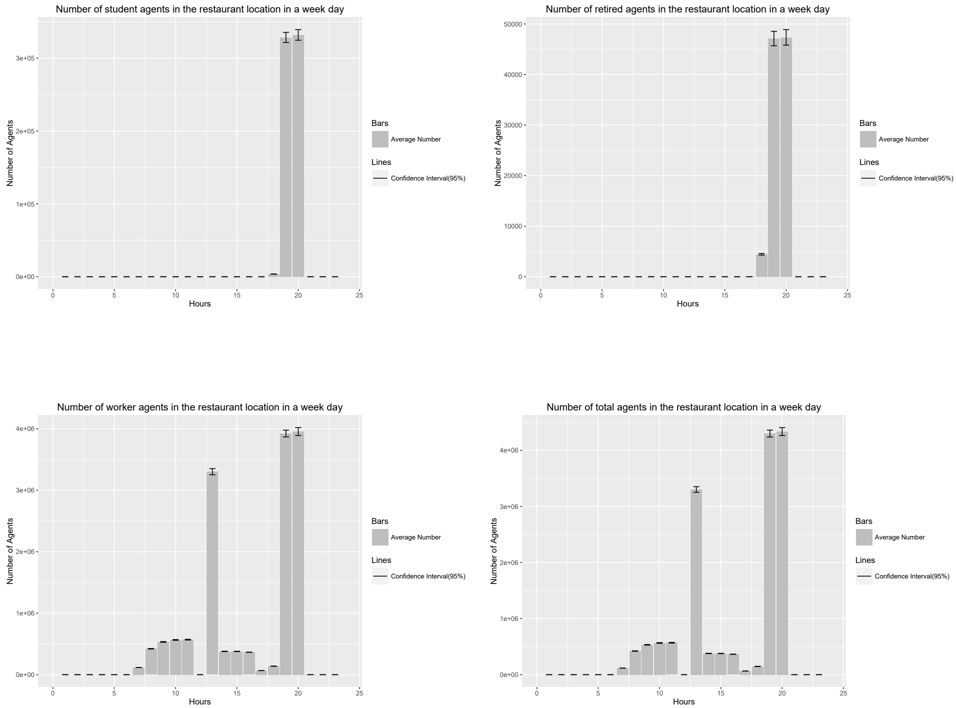


Figure A.15: Statistics of the number of agents in the restaurant location in a weekday (10 replications)

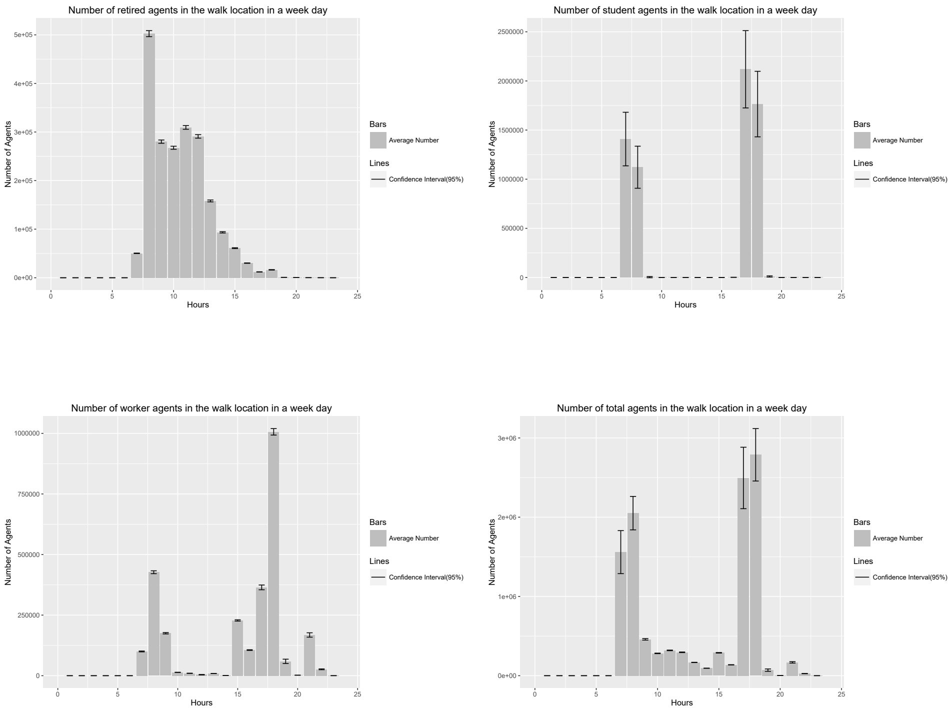


Figure A.16: Statistics of the number of agents in the 'walk' location in a weekday (10 replications)

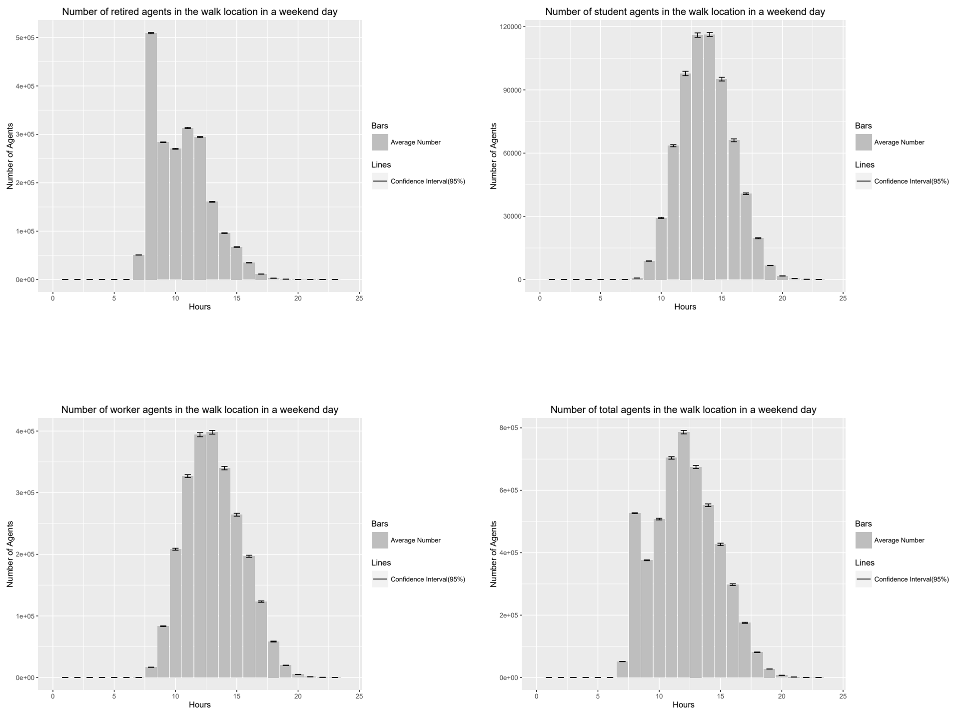


Figure A.17: Statistics of the number of agents in the 'walk' location in a weekend day (10 replications)

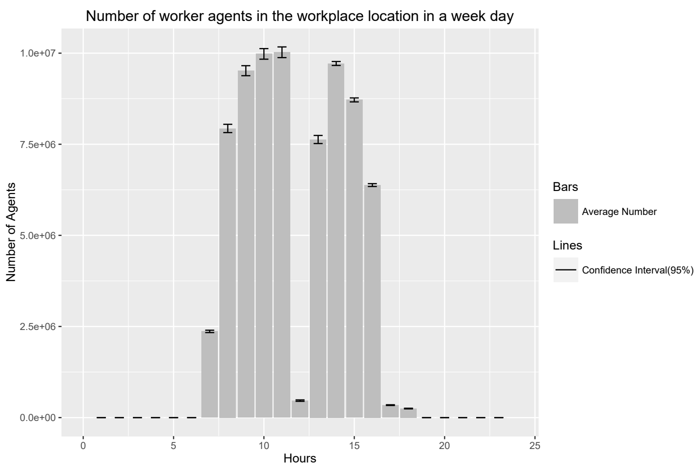


Figure A.18: Number of total agents in the workplace location in a weekday (10 replications)

B

Disease Dynamics

B.1 End Phases

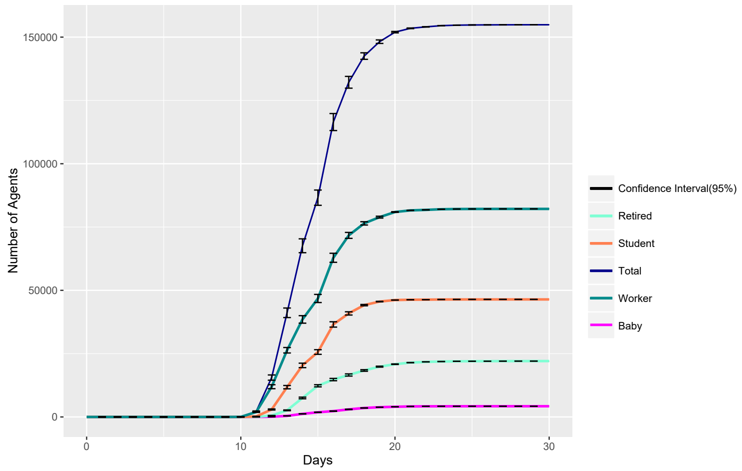


Figure B.1: Number of agents in the phase 'Dead' (10 replications)

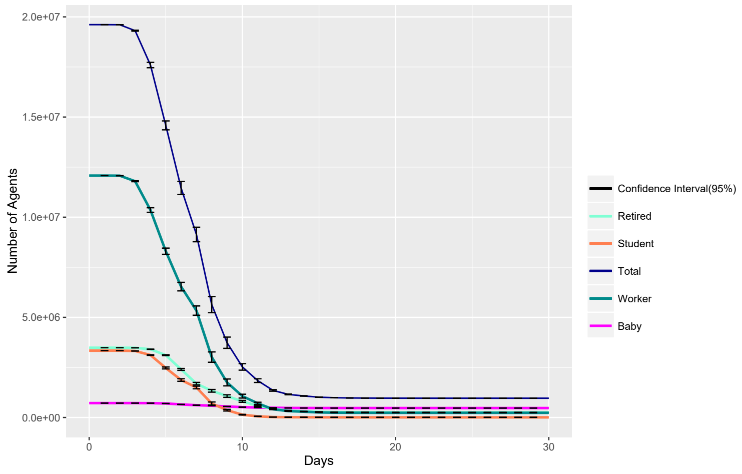


Figure B.2: Number of agents in the phase 'HEALTHY' (10 replications)

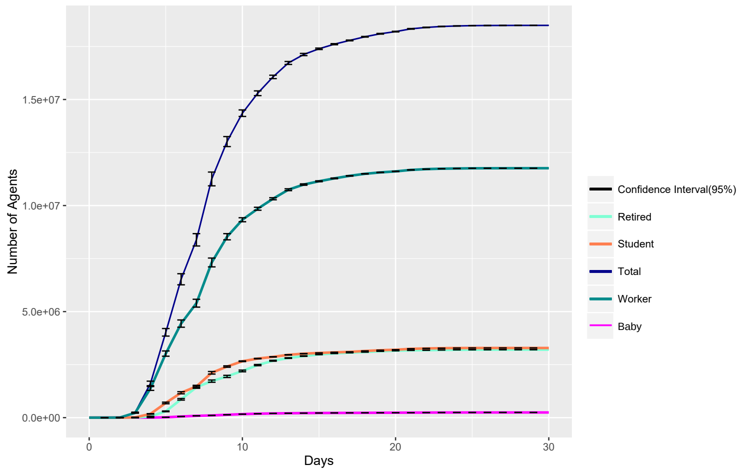


Figure B.3: Number of agents in the phase 'IMMUNITY' (10 replications)

B.2 Transitional Phases

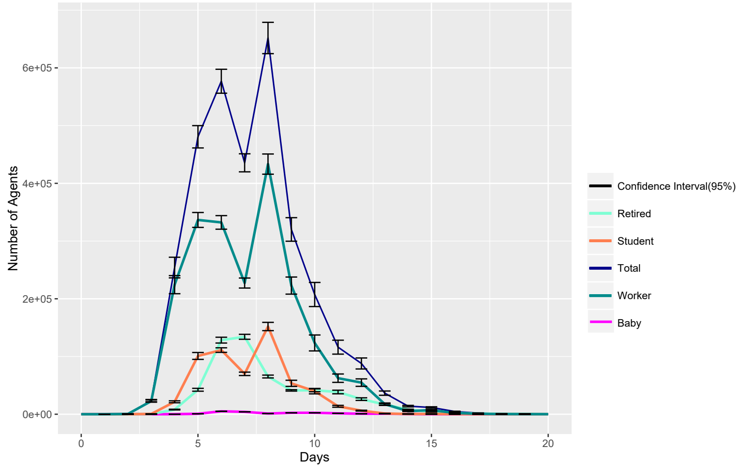


Figure B.4: Number of agents in the phase 'Asymptomatic_Contagious_Early_Stage' (10 replications)

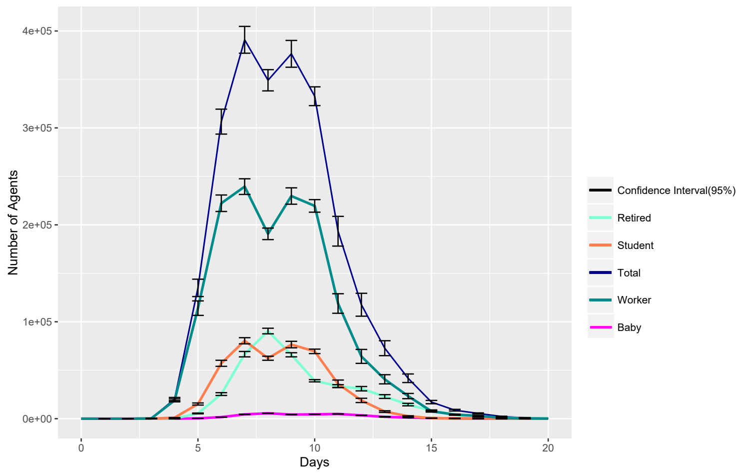


Figure B.5: Number of agents in the phase 'Asymptomatic_Contagious_Middle_Stage' (10 replications)

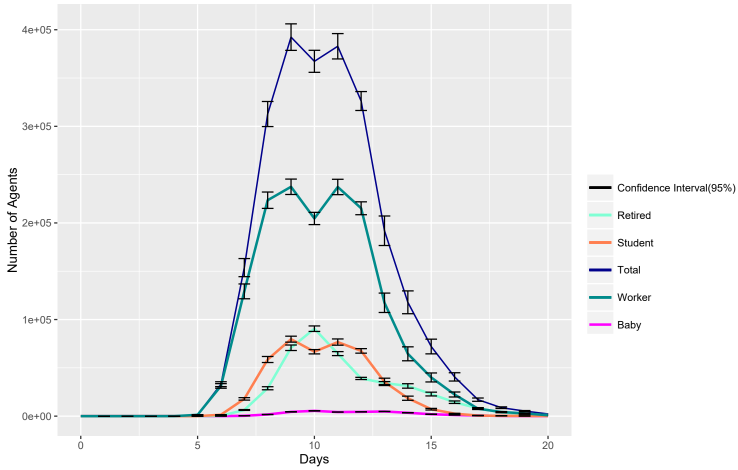


Figure B.6: Number of agents in the phase 'Asymptomatic_Contagious_Recovering' (10 replications)

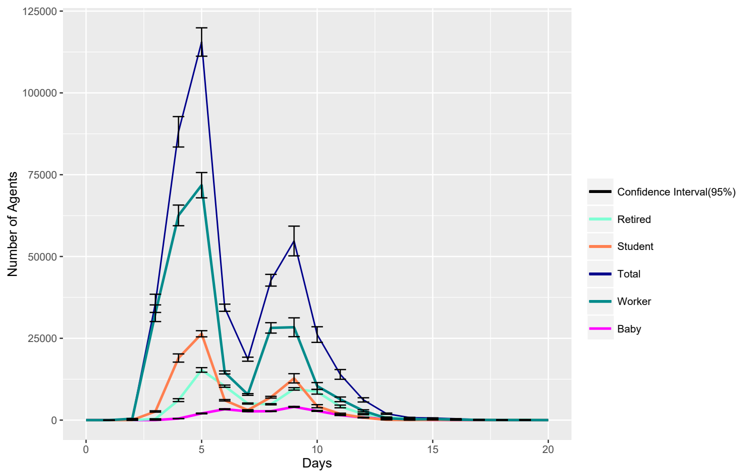


Figure B.7: Number of agents in the phase 'Asymptomatic_notContagious' (10 replications)

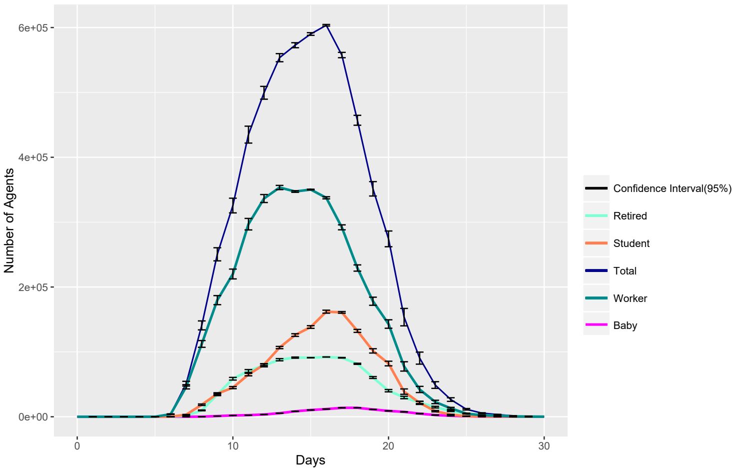


Figure B.8: Number of agents in the phase 'Convalescent' (10 replications)

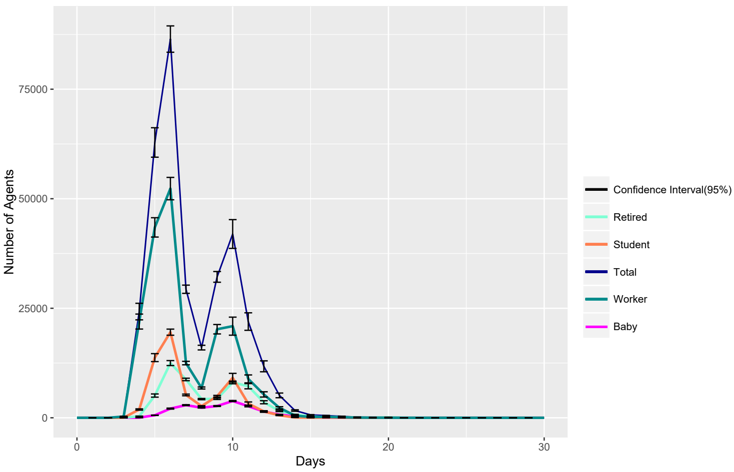


Figure B.9: Number of agents in the phase 'Symptomatic_Early_Stage' (10 replications)

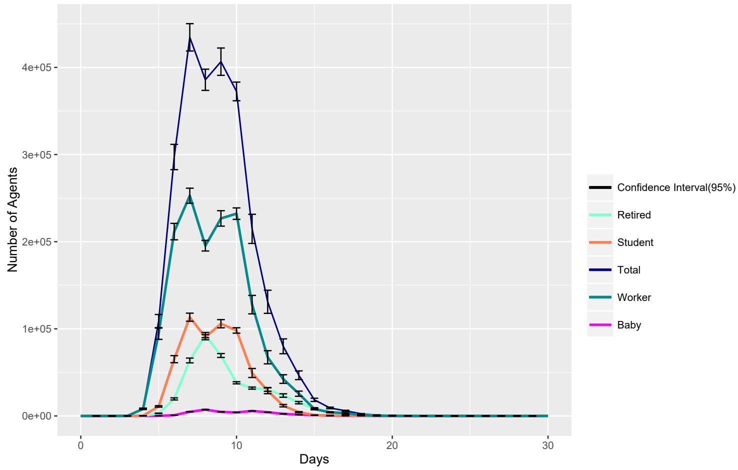


Figure B.10: Number of agents in the phase 'Symptomatic_Stay_Home' (10 replications)

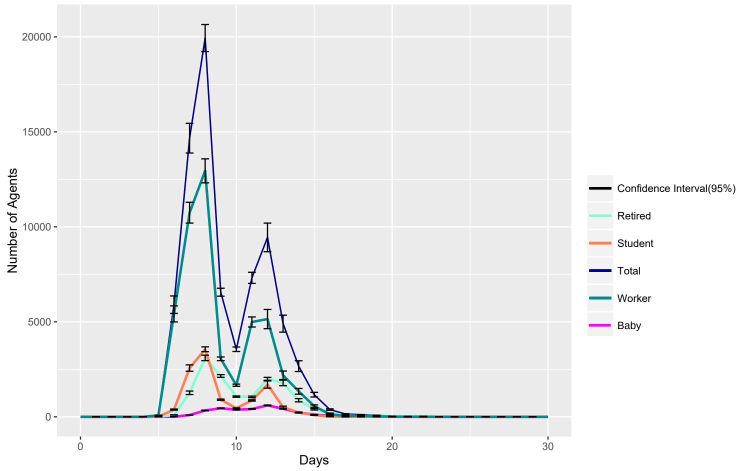


Figure B.11: Number of agents in the phase 'Symptomatic_Stay_Home_Recovering' (10 replications)

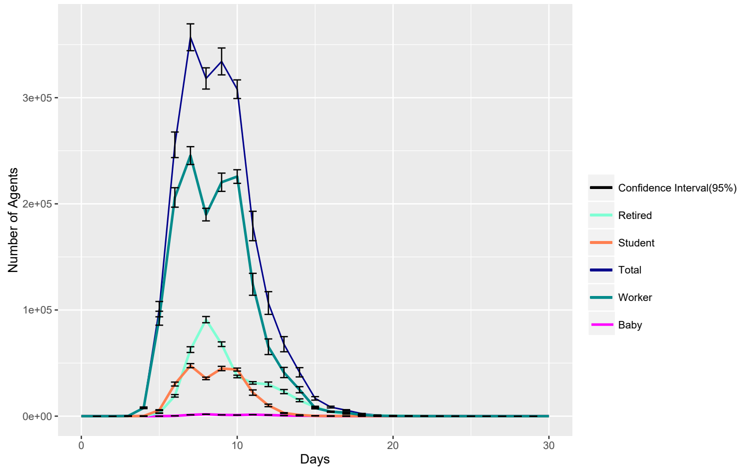


Figure B.12: Number of agents in the phase 'Symptomatic_Usual_Schedule' (10 replications)

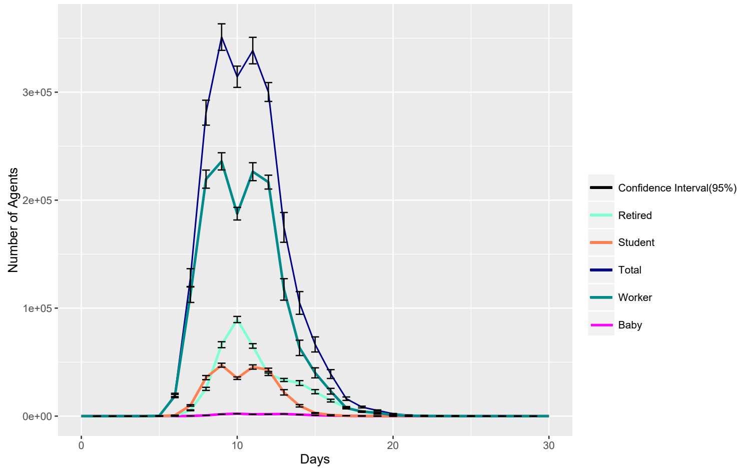


Figure B.13: Number of agents in the phase 'Symptomatic_Usual_Schedule_Recovering' (10 replications)

References

Adamatti, D. (2014). *Interdisciplinary Applications of Agent-Based Social Simulation and Modeling*. Advances in Human and Social Aspects of Technology: IGI Global (cit. on pp. 2, 31).

Ajelli, M., B. Gonçalves, D. Balcan, V. Colizza, H. Hu, J. J. Ramasco, S. Merler and A. Vespignani (Jan. 2010). Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC infectious diseases*, 10(190) (cit. on pp. 7, 37, 39, 43, 44, 78).

Alam, S. J. and A. Geller (2012). Networks in agent-based social simulation. *Agent-based models of geographical systems*. Ed. by A. J. Heppenstall, A. T. Crooks, L. M. See and M. Batty. Springer, pp. 199–216 (cit. on p. 27).

Aldewereld, H. (2007). Autonomy vs. Conformity: An Institutional Perspective on Norms and Protocols. PhD thesis. Utrecht University, pp. 0–225 (cit. on pp. 24, 32).

Allan, R. J. (2009). *Survey of agent based modelling and simulation tools*. Tech. rep. STFC Daresbury Lab (cit. on pp. 18, 19).

Allen, L. J. (1994). Some discrete-time SI, SIR, and SIS epidemic models. *Mathematical Biosciences*, 124(1), pp. 83–105 (cit. on p. 37).

Amiguet, M, J. P. Muller, J. A. Baez-Barranco and A Nagy (2002). The MOCA platform - Simulating the dynamics of social networks. *Proceedings of the 3rd International Conference on Multi-Agent-Based Simulation II*. Vol. 2581. Springer-Verlag, pp. 70–88 (cit. on pp. 29, 71).

An, G. (2008). Introduction of an agent-based multi-scale modular architecture for dynamic knowledge representation of acute inflammation. *Theoretical Biology and Medical Modelling*, 5(1), pp. 0–11 (cit. on p. 7).

Anderson, J. R., D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere and Y. Qin (Oct. 2004). An integrated theory of the mind. *Psychological Review*, 111(4), pp. 1036–1060 (cit. on p. 22).

Anderson, R. M., R. M. May and B Anderson (1992). *Infectious Diseases of Humans: Dynamics and Control*. Vol. 28. Wiley Online Library (cit. on p. 37).

Andrews, J. R., C. Morrow and R. Wood (2013). Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. *American Journal of Epidemiology*, 177(6), pp. 556–561 (cit. on pp. 41, 42, 64).

Apolloni, A., C. Poletto and V. Colizza (2013). Age-specific contacts and travel patterns in the spatial spread of 2009 H1N1 influenza pandemic. *BMC Infectious Diseases*, 13(176) (cit. on p. 51).

Arentze, T. and H. Timmermans (2000). *Albatross: a learning based transportation oriented simulation system*. Citeseer (cit. on p. 48).

ARENTZE, T., C. PELIZARO and H. TIMMERMANS (2010). An agent-based micro-simulation framework for modelling of dynamic activity-travel rescheduling decisions. *International Journal of Geographical Information Science*, 24(8), pp. 1149–1170 (cit. on p. 30).

ATKIN, M. S., G. W. KING, D. L. WESTBROOK, B. HEERINGA and P. R. COHEN (2001). Hierarchical agent control: A framework for defining agent behavior. *Proceedings of the Fifth International Conference on Autonomous Agents*. ACM, pp. 425–432 (cit. on p. 40).

AXELROD, R. and L. TESFATSION (2006). A Guide for Newcomers to Agent-Based Modeling in the Social Sciences. *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*. Ed. by L. Tesfatsion and K. L. Judd. Vol. 2. North Holland (cit. on pp. 2, 3).

BADHAM, J and R STOCKER (2010). A spatial approach to network generation for three properties: degree distribution, clustering coefficient and degree assortativity. *Journal of Artificial Societies and Social Simulation*, 13(1), pp. 0–11 (cit. on p. 29).

BARABÁSI, A.-L., N. GULBAHCE and J. LOSCALZO (2011). Network Medicine: A Network-based Approach to Human Disease. *Nature Reviews Genetics*, 12(1), pp. 56–68 (cit. on p. 67).

BARRETT, C. L., K. R. BISSET, S. G. EUBANK, X. FENG and M. V. MARATHE (2008). EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. *2008 SC - International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2008*. IEEE, pp. 1–12 (cit. on pp. 7, 39, 48, 78).

BAUMAN, A., G. MA, F. CUEVAS, Z. OMAR, T. WAQANIVALU, P. PHONGSAVAN, K. KEKE and A. BHUSHAN (2011). Cross-national comparisons of socioeconomic differences in the prevalence of leisure-time and occupational physical activity, and active commuting in six Asia-Pacific countries. *Journal of Epidemiology and Community Health*, 65(1), pp. 35–43 (cit. on p. 78).

BECKER, J. and B. NIEHAVES (2007). Epistemological perspectives on IS research: a framework for analysing and systematizing epistemological assumptions. *Information Systems Journal*, 17(2), pp. 197–214 (cit. on p. 11).

BEELER, M. F., D. M. ALEMAN and M. W. CARTER (2012). A large simulation experiment to test influenza pandemic behavior. *Proceedings of the 2012 Winter Simulation Conference*. IEEE, pp. 828–834 (cit. on p. 117).

BEGGS, C., C. NOAKES, P. SLEIGH, L. FLETCHER and K SIDDIQI (2003). The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. *The International Journal of Tuberculosis and Lung Disease*, 7(11), pp. 1015–1026 (cit. on p. 89).

BELTRAN, F. S., L. SALAS and V. QUERA (2006). Spatial behavior in groups: an agent-based approach. *Journal of Artificial Societies and Social Simulation*, 9(3), pp. 1–16 (cit. on p. 70).

BEN-ZION, Y., Y. COHEN and N. M. SHNERB (2010). Modeling epidemics dynamics on heterogenous networks. *Journal of Theoretical Biology*, 264(2), pp. 197–204 (cit. on pp. 39, 42).

BERNOULLI, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir. *Histoire de l'Acad. Roy. Sci.(Paris) avec Mém. des Math. et Phys. and Mém*, pp. 1–45 (cit. on p. 37).

BERNSTEIN, G. and K. O'BRIEN (2013). Stochastic agent-based simulations of social networks. *Proceedings of the 46th Annual Simulation Symposium*. Society for Computer Simulation International, pp. 33–40 (cit. on p. 78).

BHADURI, B., E. BRIGHT and A. ROSE (2014). Data driven approach for high resolution population distribution and dynamics models. *Proceedings of the 2014 Winter Simulation Conference*. IEEE, pp. 842–850 (cit. on p. 17).

BHAT, C., J. GUO, S. SRINIVASAN and A. SIVAKUMAR (2004). Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, 1894(1), pp. 57–66 (cit. on p. 48).

BICCHIERI, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms.* Cambridge University Press (cit. on p. 31).

BISSET, K. and M. MARATHE (2015). An interactive epidemiological modeling environment using high performance computing for enhanced user-productivity. *ACM Transactions on Management Information Systems*, 5(212), pp. 1–19 (cit. on pp. 6, 16, 17, 80).

BISSET, K., J. CHEN and X FENG (2009a). EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems. *23rd International Conference on Supercomputing*, pp. 430–439 (cit. on pp. 7, 39, 43, 44, 78).

BISSET, K. R., X. FENG, M. MARATHE and S. YARDI (2009b). Modeling interaction between individuals, social networks and public policy to support public health epidemiology. *Proceedings of the 2009 Winter Simulation Conference (WSC)*. IEEE, pp. 2020–2031 (cit. on pp. 9, 43, 112, 114, 116).

BISSET, K. R., A. M. AJI, M. V. MARATHE and W.-c. FENG (2012). High-performance biocomputing for simulating the spread of contagion over large contact networks. *BMC Genomics*, 13(Suppl 2), pp. 1–12 (cit. on p. 120).

BISSET, K. R., J. CHEN, S. DEODHAR, X. FENG, Y. MA and M. V. MARATHE (Jan. 2014). Indemics: an interactive high-performance computing framework for data-intensive epidemic modeling. *ACM Transactions on Modeling and Computer Simulation*, 24(1), 4:1–4:32 (cit. on pp. 7, 39, 43, 78, 114, 115).

BLIKSTEIN, P., D. ABRAHAMSON and U. WILENSKY (2005). Netlogo: Where we are, where we're going. *Proceedings of the annual meeting of Interaction Design and Children.* Press (cit. on p. 19).

BOBASHEV, G. V., D. M. GOEDECKE, F. YU and J. M. EPSTEIN (2007). A hybrid epidemic model: Combining the advantages of agent-based and equation-based approaches. *Proceedings of the 2007 Winter Simulation Conference*. IEEE, pp. 1532–1537 (cit. on p. 37).

BOELLA, G. and L. D. VAN TORRE (2006). Organizations in artificial social systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 3913 LNAI. Springer Berlin Heidelberg, pp. 198–210 (cit. on p. 28).

BOERO, R. (2006). *The Spatial Dimension and Social Simulations: A Review of Three Books* (cit. on p. 31).

BOMBLIES, A. and E. a. B. ELTAHIR (2009). Assessment of the impact of climate shifts on malaria transmission in the Sahel. *EcoHealth*, 6(3), pp. 426–437 (cit. on p. 69).

BONABEAU, E. (2002a). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3), pp. 7280–7287 (cit. on pp. 2, 16, 20, 21).

BONABEAU, E. (2002b). Predicting the unpredictable. *Harvard Business Review*, 80(3), pp. 109–116 (cit. on p. 1).

BORDINI, R. H., J. F. HÜBNER and M. J. WOOLDRIDGE (2007). *Programming Multi-Agent Systems in AgentSpeak Using Jason.* John Wiley & Sons, Ltd (cit. on p. 4).

BORGATTI, S. P., A. MEHRA, D. J. BRASS and G. LABIANCA (2009). Network analysis in the social sciences. *Science*, 323(5916), pp. 892–895 (cit. on p. 70).
BORSHCHEV, A., Y. KARPOV and V. KHARITONOV (2002). Distributed simulation of hybrid systems with AnyLogic and HLA. *Future Generation Computer Systems*, 18(6), pp. 829–839 (cit. on p. 6).
BRATMAN, M. (Mar. 1999). *Intention, Plans, and Practical Reason.* Stanford, CA 94305-4101: Cambridge University Press (cit. on pp. 22, 23).
BROWN, D. G., R. RIOLO, D. T. ROBINSON, M. NORTH and W. RAND (2005). Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems*, 7(1), pp. 25–47 (cit. on p. 31).
BROWN, J. J. and P. H. REINGEN (1987). Social ties and word-of-mouth referral behavior. English. *Journal of Consumer Research*, 14(3), pp. 350–362 (cit. on p. 29).
BRUZZONE, A. G., A. TREMORI, F. TARONE and F. MADEO (2011). Intelligent agents driving computer generated forces for simulating human behaviour in urban riots. *International Journal of Simulation and Process Modelling*, 6(4), pp. 308–316 (cit. on p. 16).
BRUZZONE, A., M. MASSEI, F. LONGO, S. POGGI, M. AGRESTA, C. BARTOLUCCI and L. NICOLETTI (2014). Human behavior simulation for complex scenarios based on intelligent agents. *Proceedings of the 2014 Annual Simulation Symposium.* Society for Computer Simulation International, p. 10 (cit. on p. 16).
BRYMAN, A. (2012). *Social Research Methods.* Great Clarendon Street, Oxford OX2 6DP: Oxford university press (cit. on p. 10).
BUSEMEYER, J. R. and A DIEDERICH (2010). *Cognitive Modeling.* SAGE Publications (cit. on p. 22).
CAMILLEN, F., S. CAPRÌ, C. GAROFALO, M. IGNACCOLO, G. INTURRI, A. PLUCHINO, A. RAPISARDA and S. TUDISCO (2009). Multi agent simulation of pedestrian behavior in closed spatial environments. *Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference.* Toronto, ON: IEEE, pp. 375–380 (cit. on pp. 6, 16, 17).
CAO, J, F YANG, Z GENG and X SHI (2014). Evaluating the impacts of vaccination, antiviral treatment and school closure on H1N1 influenza epidemic. *Proceedings of the 2014 Winter Simulation Conference.* Savanah, GA: IEEE, pp. 1457–1468 (cit. on p. 112).
CARRASCO, J. A., B. HOGAN, B. WELLMAN and E. J. MILLER (2008). Collecting social network data to study social activity-travel behavior: an egocentric approach. *Environment and Planning B: Planning and Design*, 35(6), pp. 961–980 (cit. on p. 28).
CASTELFRANCHI, C., F. DIGNUM, C. M. JONKER and J. TREUR (2000). Deliberative normative agents: Principles and architecture. *Intelligent Agents VI. Agent Theories, Architectures, and Languages.* Springer, pp. 364–378 (cit. on p. 24).
CASTELLANO, C., S. FORTUNATO and V. LORETO (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2), pp. 1–58 (cit. on p. 2).
CASTIGLIONE, F. (2006). Agent based modeling. *Scholarpedia*, 1(10) (cit. on pp. 16, 20).
CAZELLES, B., M. CHAVEZ, G. C. DE MAGNY, J.-F. GUÉGAN and S. HALES (2007). Time-dependent spectral analysis of epidemiological time-series with wavelets. *Journal of the Royal Society Interface*, 4(15), pp. 625–636 (cit. on p. 8).
CECEZ-KECMANOVIC, D. (2001). Doing critical IS research: The question of methodology. *Qualitative Research in IS: Issues and Trends.* IGI Global, pp. 141–162 (cit. on p. 11).
CELIK, N., E. MAZHARI, J. CANBY, O. KAZEMI, P. SARFARE, M. S. AL-OTAIBI and Y.-J. SON (2012). Automatic partitioning of large scale simulation in grid computing for run time re-

duction. *Innovations in Information Systems for Business Functionality and Operations Management*. Vol. 380. IGI Global, pp. 225–252 (cit. on p. 130).

CHANDY, K. M. and J. MISRA (1979). Distributed simulation: A case study in design and verification of distributed programs. *IEEE Transactions on Software Engineering*, (5), pp. 440–452 (cit. on p. 130).

CHEN, B., Y. GE, L. ZHANG, Y. ZHANG, Z. ZHONG and X. LIU (2014). A modeling and experiment framework for the emergency management in AHC transmission. *Computational and Mathematical Methods in Medicine*, 2014, pp. 1–18 (cit. on pp. 39, 43, 130).

CHEN, B., L. ZHANG, G. GUO and X. QIU (2015). KD-ACP: a software framework for social computing in emergency management. *Mathematical Problems in Engineering*, 2015 (cit. on pp. 102, 130).

CHEN, J., A MARATHE and M. MARATHE (2010). Coevolution of epidemics, social networks, and individual behavior: A case study. *Third International Conference on Social Computing, Behavioral Modeling, and Prediction*. Springer Berlin Heidelberg, pp. 218–227 (cit. on pp. 112, 116).

CHEN, S.-H. (2008). Software-agent designs in economics: an interdisciplinary. *Computational Intelligence Magazine, IEEE*, 3(4), pp. 18–22 (cit. on p. 22).

CICIRELLI, F., A. GIORDANO and L. NIGRO (2015). Efficient environment management for distributed simulation of large-scale situated multi-agent systems. *Concurrency and Computation: Practice and Experience*, 27(3), pp. 610–632 (cit. on pp. 4, 8, 31).

COLLIER, N. and M. NORTH (2012). Parallel agent-based simulation with Repast for High Performance Computing. *Simulation*, 89(10), pp. 1215–1235 (cit. on p. 130).

CONDON, B. J. and T. SINHA (2010). Who is that masked person: The use of face masks on Mexico City public transportation during the Influenza A (H1N1) outbreak. *Health Policy*, 95(1), pp. 50–56 (cit. on p. 41).

COOKE, K. L. and P VAN DEN DRIESSCHE (1996). Analysis of an SEIRS epidemic model with two delays. *Journal of Mathematical Biology*, 35(2), pp. 240–260 (cit. on p. 38).

COOLEY, P., S. BROWN, J. CAJKA, B. CHASTEEN, L. GANAPATHI, J. GREFENSTETTE, C. R. HOLLINGSWORTH, B. Y. LEE, B. LEVINE, W. D. WHEATON and D. K. WAGENER (2011). The role of subway travel in an influenza epidemic: a New York City simulation. *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, 88(5), pp. 982–995 (cit. on pp. 42, 109, 110).

CRAWFORD, S. E. and E. OSTROM (1995). A grammar of institutions. *American Political Science Review*, 89(03), pp. 582–600 (cit. on p. 32).

CROOKS, A. (2006). Exploring cities using agent-based models and GIS. *Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects*, pp. 125–132 (cit. on pp. 6, 16).

CROSSAN, F. (2003). Research philosophy: towards an understanding. *Nurse researcher*, 11(1), pp. 46–55 (cit. on p. 11).

CUI, J., Y. TAKEUCHI and Y. SAITO (2006). Spreading disease with transport-related infection. *Journal of Theoretical Biology*, 239, pp. 376–390 (cit. on p. 42).

D'ONOFRIO, A. (2002). Stability properties of pulse vaccination strategy in SEIR epidemic model. *Mathematical Biosciences*, 179(1), pp. 57–72 (cit. on p. 38).

DARDENNE, A., A. VAN LAMSWEERDE and S. FICKAS (1993). *Goal-directed requirements acquisition* (cit. on p. 28).

DASTANI, M. (2008). 2APL: a practical agent programming language. *Autonomous Agents and Multi-Agent Systems*, 16(3), pp. 214–248 (cit. on p. 4).

DAVILA, J, E GOMEZ, K LAFFAILLE, K. TUCCI and M. UZCATEGUI (2005). Multiagent distributed simulation with GALATEA. *Ninth IEEE International Symposium on Distributed Simulation and Real-Time Applications*. IEEE, pp. 165–170 (cit. on p. 5).

DE GIACOMO, G., Y. LESPÉRANCE and H. J. LEVESQUE (2000). ConGolog, a concurrent programming language based on the situation calculus. *Artificial Intelligence*, 121(1), pp. 109–169 (cit. on p. 4).

DE ESPÍNDOLA, A. L., C. T. BAUCH, B. C. TROCA CABELLA and A. S. MARTINEZ (2011). An agent-based computational model of the spread of tuberculosis. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(05), pp. 1–17 (cit. on pp. 6, 39).

DEADMAN, P., D. ROBINSON, E. MORAN and E. BRONDIZIO (2004). Colonist household decisionmaking and land-use change in the Amazon Rainforest: an agent-based simulation. *Environment and Planning B*, 31, pp. 693–710 (cit. on p. 16).

DECROP, A. (2006). *Vacation Decision Making*. CABI (cit. on p. 11).

DEL VALLE, S., J. HYMAN, H. HETHCOTE and S. EUBANK (2007). Mixing patterns between age groups in social networks. *Social Networks*, 29(4), pp. 539–554 (cit. on pp. 38, 42).

DELL'AVERSANA, P. (2013). *Cognition in Geosciences: The Feeding Loop between Geodisciplines, Cognitive Sciences and Epistemology*. Elsevier Science (cit. on p. 10).

DENPHEDTNONG, A., S. CHINVIRIYASIT and W. CHINVIRIYASIT (2013). On the dynamics of SEIRS epidemic model with transport-related infection. *Mathematical Biosciences*, 245(2), pp. 188–205 (cit. on p. 42).

DOOLEY, K. and K. DOOLEY (2001). *Social Research Methods*. 4 th ed. Upper Saddle River, N. J.: Prentice Hall (cit. on p. 13).

DUAN, W., Z. FAN, P. ZHANG, G. GUO and X. QIU (2015). Mathematical and computational approaches to epidemic modeling: a comprehensive review. *Frontiers of Computer Science*, 9(5), pp. 1–21 (cit. on pp. 38, 39).

DUBOZ, R., D. VERSMISSE, G. QUESNEL, A. MUZY and E. RAMAT (2006). Specification of dynamic structure discrete event multiagent systems. *Agent Directed Simulation (Spring Simulation Multiconference)*. Huntsville, AL, USA: Society for Computer Simulation, p. 103 (cit. on p. 28).

DUNIN-KEPLICZ, B. and R. VERBRUGGE (2002). Collective intentions. *Fundamenta Informaticae*, 51(3), pp. 271–295 (cit. on p. 22).

EASLEY, D. and J. KLEINBERG (2010). *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press (cit. on p. 2).

EDMONDS, B. (2002). Towards an ideal social simulation language. *Proceedings of the 3rd International Conference on Multi-agent-based Simulation II*. Springer-Verlag, pp. 50–53 (cit. on p. 29).

EPSTEIN, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(suppl 3), pp. 7243–7250 (cit. on p. 16).

EPSTEIN, J. M. and R. L. AXTELL (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. The Brookings Institution (cit. on pp. 2, 17).

ESTRADA-PEÑA, A., D. CARREÓN, C. ALMAZÁN and J. DE LA FUENTE (2014). Modeling the impact of climate and landscape on the efficacy of white tailed deer vaccination for cattle tick control in northeastern Mexico. *PLoS ONE*, 9(7), pp. 1–8 (cit. on p. 69).

FAN, M., M. Y. LI and K. WANG (2001). Global stability of an SEIS epidemic model with recruitment and a varying total population size. *Mathematical Biosciences*, 170(2), pp. 199–208 (cit. on p. 38).

FELDKAMP, N and S. STRASSBURGER (2014). Automatic generation of route networks for microscopic traffic simulations. *Proceedings of the 2014 Winter Simulation Conference.* Savanah, GA: IEEE, pp. 2848–2859 (cit. on p. 64).

FENICHEL, E. P., C. CASTILLO-CHAVEZ, M. CEDDIA, G. CHOWELL, P. A. G. PARRA, G. J. HICKLING, G. HOLLOWAY, R. HORAN, B. MORIN, C. PERRINGS et al. (2011). Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences*, 108(15), pp. 6306–6311 (cit. on p. 8).

FERBER, J. and O. GUTKNECHT (1998). A meta-model for the analysis and design of organizations in multi-agent systems. *Proceedings of the 3rd International Conference on Multi Agent Systems.* IEEE Computer Society, pp. 128–135 (cit. on pp. 28, 32).

FERBER, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence.* 1st ed. Boston, MA, USA: Addison-Wesley Reading (cit. on p. 21).

FERGUSON, I. A. (1992). TouringMachines: an architecture for dynamic, rational, mobile agents. PhD thesis. University of Cambridge, pp. 1–206 (cit. on p. 22).

FILHO, H. B., F. NETO and W. FUSCO (2013). Migration, communication and social networks - an agent-based social simulation. *Complex Networks*, 424, pp. 67–74 (cit. on pp. 6, 18).

FOO, N. and P. PEPPAS (2005). Systems theory: melding the AI and simulation perspectives. *Lecture Notes in Computer Science : Artificial Intelligence and Simulation.* Ed. by T. G. Kim. Vol. 3397. Springer Berlin Heidelberg, pp. 14–23 (cit. on p. 31).

FRANKLIN, S. and A. GRAESSER (1997). Is it an agent, or just a program?: A taxonomy for autonomous agents. *Intelligent Agents III Agent Theories, Architectures, and Languages.* Ed. by J. P. Muller, M. J. Wooldridge and N. R. Jennings. Vol. 1193. Springer, pp. 21–35 (cit. on p. 21).

FRIAS-MARTINEZ, E (2011). An agent-based model of epidemic spread using human mobility and social network information. *2011 IEEE International Conference on Privacy, Security, Risk, and Trust.* Boston, MA: IEEE, pp. 57–64 (cit. on p. 39).

FUJIMOTO, R. M. (1999). Parallel and distributed simulation. *Proceedings of the 31st conference on Winter simulation.* ACM. Phoenix, AZ, pp. 122–131 (cit. on pp. 4, 8, 130).

GABBRIELLINI, S. and P. TORRONI (2014). A new framework for ABMs based on argumentative reasoning. *Advances in Social Simulation.* Ed. by B. Kaminiski and G. Koloch. Vol. 229. Advances in Intelligent Systems and Computing. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 25–36 (cit. on pp. 28, 30).

GANZHA, M. and L. JAIN (2012). *Multiagent Systems and Applications: Volume 1:Practice and Experience.* Intelligent Systems Reference Library. Springer Berlin Heidelberg (cit. on p. 6).

GARGIULO, F., M. LENORMAND, S. HUET and O. BAQUEIRO ESPINOSA (2012). Commuting Network Models : Getting the Essentials. *Journal of Artificial Societies and Social Simulation*, 15(2), pp. 1–17 (cit. on p. 64).

GASSER, L and K. KAKUGAWA (2002). MACE3J: fast flexible distributed simulation of large, large-grain multi-agent systems. *Proceeding of The First International Joint Conference on Autonomous Agents & Multiagent Systems.* Bologna, Italy: ACM, pp. 745–752 (cit. on p. 4).

GATTI, M, P. CAVALIN and S. NETO (2014). Large-scale multi-agent-based modeling and simulation of microblogging-based online social network. *14th International Workshop on Multi-Agent-Based Simulation.* Ed. by S. J. Alam and H. V. D. Parunak. Vol. 8235. Springer Berlin Heidelberg, pp. 17–33 (cit. on pp. 28, 70).

GE, Y., L. LIU, X. QIU, H. SONG, Y. WANG and K. HUANG (2013). A framework of multilayer social networks for communication behavior with agent-based modeling. *Simulation*, 89(7), pp. 810–828 (cit. on pp. 7, 43, 44).

GE, Y., R. MENG, Z. CAO, X. QIU and K. HUANG (2014). Virtual city: An individual-based digital environment for human mobility and interactive behavior. *Simulation*, 90(8), pp. 917–935 (cit. on pp. 48, 78, 79, 102).

GEERTS, G. L. (2011). A design science research methodology and its application to accounting information systems research. *International Journal of Accounting Information Systems*, 12(2), pp. 142–151 (cit. on p. 12).

GETCHELL, A. (2008). Agent-based modeling. *Physics*, 22(6), pp. 757–767 (cit. on pp. 16, 20).

GILBERT, N. (2004). Agent-based social simulation: dealing with complexity. *The Complex Systems Network of Excellence*, 9, pp. 1–14 (cit. on p. 2).

GILBERT, N. and P. TERNA (2000). How to build and use agent-based models in social science. *Mind & Society*, 1(1), pp. 57–72 (cit. on pp. 4–6).

GILBERT, N. and K. G. TROITZSCH (2005). *Simulation for the Social Scientist*. Open University Press, pp. 1–291 (cit. on pp. 2, 3).

GONG, X. and R. XIAO (2007). Research on multi-agent simulation of epidemic news spread characteristics. *Journal of Artificial Societies and Social Simulation*, 10(3), pp. 1–17 (cit. on pp. 6, 17).

GRANT, J., S. KRAUS and D. PERLIS (2005). A logic-based model of intention formation and action for multi-agent subcontracting. *Artificial Intelligence*, 163, pp. 163–201 (cit. on p. 22).

GREEN, M. S., T. SWARTZ, E. MAYSHAR, B. LEV, A. LEVENTHAL, P. E. SLATER and J. SHEMER (2002). When is an epidemic an epidemic? *The Israel Medical Association journal: IMAJ*, 4(1), pp. 3–6 (cit. on p. 35).

GREFENSTETTE, J. J., S. T. BROWN, R. ROSENFELD, J. DEPASSE, N. T. STONE, P. C. COOLEY, W. D. WHEATON, A. FYSHE, D. D. GALLOWAY, A. SRIRAM, H. GUCLU, T. ABRAHAM and D. S. BURKE (2013). FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. *BMC Public Health*, 13(1), pp. 1–14 (cit. on pp. 39, 40, 42, 78, 80, 109, 115).

GRIMALDO, F., M. LOZANO, F. BARBER and A. GUERRA-HERN (2011). A J-MADeM agent-based social simulation to model urban mobility. *9th International Conference on Practical Applications of Agents and Multi-Agent Systems*. Springer-Verlag Berlin Heidelberg, pp. 1–11 (cit. on p. 5).

GRONER, M., R COX, G GETTINBY and C. REVIE (2013). Use of agent-based modelling to predict benefits of cleaner fish in controlling sea lice, Lepeophtheirus salmonis, infestations on farmed Atlantic salmon, Salmo salar L. *Journal of Fish Diseases*, 36(3), pp. 195–208 (cit. on p. 16).

GROSSI, D., F. DIGNUM, M. DASTANI and L. ROYAKKERS (2005). Foundations of organizational structures in multiagent systems. *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*. New York, NY, USA: ACM Press, pp. 690–697 (cit. on p. 28).

GRUNE-YANOFF, T. (Sept. 2010). Agent-based models as policy decision tools: The case of smallpox vaccination. *Simulation & Gaming*, 42(2), pp. 225–242 (cit. on pp. 6, 39).

GUBA, E. and Y. LINCOLN (1994). Competing paradigms in qualitative research. *Handbook of Qualitative Research*, 2, pp. 163–194 (cit. on p. 10).

GUO, D., K. C. LI, T. R. PETERS, B. M. SNIVELY, K. A. POEHLING and X. ZHOU (2015). Multi-scale modeling for the transmission of influenza and the evaluation of interventions toward it. *Scientific Reports*, 5(8090), pp. 1–9 (cit. on pp. 8, 39, 69).

GUO, J. and X. LI (2012). *The 2012 Beijing Traffic Development Report*. Tech. rep. Beijing Traffic Research Center, pp. 1–90 (cit. on pp. 95, 96).

GUO, Z. and J. C. TAY (2008). Multi-timescale event-scheduling in multi-agent immune simulation models. *BioSystems*, 91, pp. 126–145 (cit. on p. 56).

HALL, C. B. (2007). The spread of influenza and other respiratory viruses: complexities and conjectures. *Clinical Infectious Diseases*, 45(3), pp. 353–359 (cit. on p. 7).

HAMILL, L. and N. GILBERT (2010). Simulating large social networks in agent-based models: A social circle model. *Emergence: Complexity and Organization*, 12(4), pp. 78–94 (cit. on pp. 71, 88, 127).

HARRISON, M. (2013). *Disease and the Modern World: 1500 to the Present Day*. John Wiley & Sons (cit. on p. 36).

HAWE, G. I., G. COATES, D. T. WILSON and R. S. CROUCH (Nov. 2012). Agent-based simulation for large-scale emergency response. *ACM Computing Surveys*, 45(1), pp. 1–51 (cit. on pp. 6, 8, 17, 31, 40, 43).

HAYS, J. N. (2009). *The Burdens of Disease: Epidemics and Human Response in Western History*. Rutgers University Press (cit. on p. 36).

HELBING, D. (2012). Agent-based modeling. *Social self-organization*. Springer, pp. 25–70 (cit. on pp. 16, 20).

HELBING, D. and S. BALIETTI (2013). How to Do Agent-Based Simulations in the Future : From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design Why Develop and Use Agent-Based Models? *Social Self-Organization*. Santa Fe Institute, pp. 25–70 (cit. on pp. 5, 8).

HESAN, R., A. GHORBANI and V. DIGNUM (2015). Modelling environments in ABMS: A system dynamics approach. *Multi-Agent-Based Simulation XV*. Springer, pp. 41–52 (cit. on pp. 31, 41, 67).

HEVNER, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, 19(2), pp. 1–6 (cit. on p. 11).

HILL, R. and R. DUNBAR (2003). Social network size in humans. *Human Nature*, 14(1), pp. 53–72 (cit. on p. 88).

HINDRIKS, K. V., F. S. DE BOER, W. VAN DER HOEK and J.-J. C. MEYER (1999). Agent Programming in 3APL. *Autonomous Agents and Multi-Agent Systems*, 2(4), pp. 357–401 (cit. on p. 4).

HOLZHAUER, S. (2010). *Developing a Social Network Analysis and Visualization Module for Repast Models*. Vol. 4. Kassel University Press GmbH (cit. on p. 19).

HORIZONKEY (2007). *Chinese Social Pattern: A Small Circle*. Tech. rep. HorizonKey, pp. 1–2 (cit. on p. 129).

HOUK, V. N., D. C. KENT, J. H. BAKER, K. SORENSEN and G. D. HANZEL (1968). The Byrd study: in-depth analysis of a micro-outbreak of tuberculosis in a closed environment. *Archives of Environmental Health: An International Journal*, 16(1), pp. 4–6 (cit. on p. 89).

HU, X, A MUZY and L NTAIMO (2005). A hybrid agent-cellular space modeling approach for fire spread and suppression simulation. *Proceedings of the 2005 Winter Simulation Conference*. IEEE, pp. 248–255 (cit. on pp. 4, 16).

HUANG, C.-Y., T.-H. WEN and Y.-S. TSAI (2013). FLUed: A Novel Four-Layer Model for Simulating Epidemic Dynamics and Assessing Intervention Policies. *Journal of Applied Mathematics*, 2013, pp. 1–20 (cit. on p. 80).

HUANG, Y. (2010). Modeling the severe acute respiratory syndrome (SARS) outbreak in Beijing: an agent-based approach. *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application*. New York, NY, USA: ACM (cit. on pp. 6, 39, 78).

HUBER, M. J. (1999). JAM : A BDI-theoretic mobile agent architecture. *Proceedings of the Third Annual Conference on Autonomous Agents*. Ed. by O. Etzioni, J. P. Muller and J. M. Bradshaw. New York, USA: ACM New York, pp. 236–243 (cit. on p. 23).

HÜBNER, J. F., O. BOISSIER and R. H. BORDINI (2011). A normative programming language for multi-agent organisations. *Annals of Mathematics and Artificial Intelligence*, 62(1-2), pp. 27–53 (cit. on p. 32).

HUNTBACH, M. and G. RINGWOOD (2003). *Agent-Oriented Programming: From Prolog to Guarded Definite Clauses*. Lecture Notes in Computer Science. Springer Berlin Heidelberg (cit. on p. 4).

IANNELLI, M, F. MILNER and A PUGLIESE (1992). Analytical and numerical results for the age-structured SIS epidemic model with mixed inter-intracohort transmission. *SIAM Journal on Mathematical Analysis*, 23(3), pp. 662–688 (cit. on p. 37).

INGRAND, F. F., M. P. GEORGE and A. S. RAO (1992). An architecture for real-time reasoning and system control. *IEEE Expert*, 7(6), pp. 33–44 (cit. on p. 23).

ISTRATE, G., A. HANSSON, C. D. TALLMAN, L. CUELLAR-HENGARTNER and N. HENGARTNER (2006). Towards generative activity-based models for large-scale socio-technical simulations. *Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects*. Chicago, IL, pp. 215–230 (cit. on p. 48).

IZQUIERDO, L. R., N. M. GOTTS and J. G. POLHILL (2003). *FEARLUS-W: an agent-based model of river basin land use and water management*. Tech. rep. Utrecht University, The Netherlands (cit. on p. 3).

JACOBS, P. H., N. A. LANG and A. VERBRAECK (2002). D-SOL: A Distributed JAVA based Discrete Event Simulation Architecture. *Proceedings of the 2002 Winter Simulation Conference*. San Diego, CA, USA: IEEE, pp. 793–800 (cit. on p. 90).

JENNINGS, N. and M. WOOLDRIDGE (2012). *Agent Technology: Foundations, Applications, and Markets*. Springer Berlin Heidelberg (cit. on p. 21).

JENNINGS, N. R. (2000). On agent-based software engineering. *Artificial Intelligence*, 117(2), pp. 277–296 (cit. on p. 3).

JIANG, Y., S. MEMBER and J. C. JIANG (2014). Understanding social networks from a multiagent perspective. *IEEE Transactions on Parallel and Distributed Systems*, 25(10), pp. 2743–2759 (cit. on p. 29).

KAKAS, A, L AMGOUD and G KERN-ISBERNER (2012). ABA: Argumentation based agents. *Eigth International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2011)*. Ed. by P. McBurney, S. Parsons and I. Rahwan. Springer Berlin Heidelberg, pp. 9–27 (cit. on p. 30).

KERMACK, W. O. and A. G. MCKENDRICK (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 115. 772. The Royal Society, pp. 700–721 (cit. on p. 37).

KITE, M. E. (1996). Age, gender, and occupational label: A test of social role theory. *Psychology of Women Quarterly*, 20(3), pp. 361–374 (cit. on p. 51).

KOLLINGBAUM, M. J. and T. J. NORMAN (2003). NoA-a normative agent architecture. *IJCAI*, pp. 1465–1466 (cit. on p. 24).

KOLP, M., P. GIORGINI and J. MYLOPOULOS (2006). Multi-agent architectures as organizational structures. *Autonomous Agents and Multi-Agent Systems*, 13(1), pp. 3–25 (cit. on p. 28).

KOPMAN, S., M. I. AKBAS and D. TURGUT (2012). EpidemicSim: Epidemic simulation system with realistic mobility. *37th Annual IEEE Conference on Local Computer Networks*. Clearwater, FL: IEEE, pp. 659–665 (cit. on p. 64).

KOSSINETS, G. and D. J. WATTS (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115, pp. 405–450 (cit. on p. 29).

KROGSTIE, J., A. OPDAHL and S. BRINKKEMPER (2007). *Conceptual Modelling in Information Systems Engineering*. Springer (cit. on p. 48).

LÄMMEL, G., D. GRETHER and K. NAGEL (2010). The representation and implementation of time-dependent inundation in large-scale microscopic evacuation simulations. *Transportation Research Part C: Emerging Technologies*, 18(1), pp. 84 –98 (cit. on p. 17).

LANGE, D. B. and O. MITSURU (1998). *Programming and Deploying Java Mobile Agents Aglets*. Addison-Wesley Longman Publishing Co., Inc. (cit. on p. 22).

LAST, J. M. (2001). *A Dictionary of Epidemiology*. Vol. 141. Oxford Univ Press (cit. on p. 35).

LAVER, M. and E. SERGENTI (2011). *Party Competition: An Agent-Based Model: An Agent-Based Model*. Princeton University Press (cit. on p. 16).

LEES, M., B. LOGAN and R. MINSON (2005). Modelling environments for distributed simulation. *Environments for Multi-Agent Systems*. Ed. by D. Weyns, H. V. D. Parunak and F. Michel. Vol. 3374. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 150–167 (cit. on p. 4).

LI, G. and J. ZHEN (2005). Global stability of an SEI epidemic model with general contact rate. *Chaos, Solitons & Fractals*, 23(3), pp. 997–1004 (cit. on p. 38).

LI, X., W. MAO, D. ZENG and F. WANG (2008). Agent-based social simulation and modeling in social computing. *Proceedings of the IEEE ISI 2008 PAISI, PACCF, and SOCO International Workshops on Intelligence and Security Informatics*. Vol. 5075. Springer-Verlag, pp. 401–412 (cit. on pp. 1, 16).

LI, X., J. CAO, Y. HE and Y. CHEN (2006). MADESE: A simulation environment for mobile agent. *The Sixth IEEE International Conference on Computer and Information Technology (CIT'06)*. Seoul: IEEE, pp. 86–86 (cit. on p. 22).

LIN, T. and D. WANG (2014). Social networks and joint/solo activity-travel behavior. *Transportation Research Part A: Policy and Practice*, 68, pp. 18–31 (cit. on p. 30).

LINARD, C., N. PONÇON, D. FONTENILLE and E. F. LAMBIN (2009). A multi-agent simulation to assess the risk of malaria re-emergence in southern France. *Ecological Modelling*, 220(2), pp. 160–174 (cit. on pp. 6, 39).

LINGNAU, A. (1999). Simulating mobile agent systems with Swarm. *Proceedings of First International Symposium on Agent Systems and Applications and Third International Symposium on Mobile Agents*. Palm Springs, CA: IEEE, pp. 272–273 (cit. on pp. 6, 18).

LIU, J., D. M. NICOL and K. TAN (2001). Lock-free scheduling of logical processes in parallel simulation. *Proceedings of the Fifteenth Workshop on Parallel and Distributed Simulation*. IEEE Computer Society. Washington, DC, USA, pp. 22–31 (cit. on p. 130).

LIU, X. and Y. TAKEUCHI (2006). Spread of disease with transport-related infection and entry screening. *Journal of Theoretical Biology*, 242, pp. 517–528 (cit. on p. 42).

LLOYD-SMITH, J. O., D. MOLLISON, C. METCALF, P. KLEPAC and J. HEESTERBEEK (2015). Challenges in modelling infectious disease dynamics: preface. *Epidemics*, 10, pp. iii–iv (cit. on p. 36).

LUCK, M. and M. D'INVERNO (2001). A conceptual framework for agent definition and development. *The Computer Journal*, 44(1), pp. 1–20 (cit. on p. 21).

LUKE, S., C. CIOFFI-REVILLA, L. PANAIT and K. SULLIVAN (2004). Mason: A new multi-agent simulation toolkit. *Proceedings of the 2004 Swarmfest Workshop*. Vol. 8. University of Michigan (cit. on p. 5).

LUNA, F. and B. STEFANSSON (2012). *Economic Simulations in Swarm: Agent-Based Modelling and Object Oriented Programming*. Advances in Computational Economics. Springer US (cit. on p. 4).

LYTINEN, S. L. and S. F. RAILSBACK (2012). The evolution of agent-based simulation platforms: A review of netlogo 5.0 and relogo. *Proceedings of the Fourth International Symposium on Agent-Based Modeling and Simulation* (cit. on p. 19).

MACAL, C. M. and M. J. NORTH (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3), pp. 151–162 (cit. on pp. 15, 16, 19–21, 24, 27).

MACAL, C., T. HOWE et al. (2006). *Who's your neighbor? Neighbor identification for agent-based modeling*. Tech. rep. Argonne National Laboratory (ANL) (cit. on p. 27).

MACAL, C. M., M. J. NORTH, N. COLLIER, V. M. DUKIC, D. S. LAUDERDALE, M. Z. DAVID, R. S. DAUM, P. SHUMM, J. A. EVANS, J. R. WILDER and D. T. WEGENER (2012). Modeling the spread of community-associated MRSA. *Proceedings of the 2012Winter Simulation Conference*. Berlin: IEEE, pp. 1–12 (cit. on pp. 39, 40).

MAO, L. (2014). Modeling triple-diffusions of infectious diseases, information, and preventive behaviors through a metropolitan social network - an agent-based simulation. *Applied Geography*, 50, pp. 31–39 (cit. on p. 78).

MARSELLA, S., D. PYNADATH and S. READ (2004). PsychSim: Agent-based modeling of social interactions and influence. *Proceedings of the International Conference on Cognitive Modeling - ICCM 2004*. Pittsburgh, Pennsylvania, USA, pp. 243–248 (cit. on p. 72).

MCKENZIE, B. and M. RAPINO (2011). *Commuting in the United States: 2009*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau (cit. on p. 78).

MCNEILL, W. (2010). *Plagues and Peoples*. Anchor (cit. on p. 36).

MEI, S., P. SLOOT, R. QUAX, Y. ZHU and W. WANG (2010). Complex agent networks explaining the HIV epidemic among homosexual men in Amsterdam. *Mathematics and Computers in Simulation*, 80(5), pp. 1018–1030 (cit. on p. 39).

MEI, S., B. CHEN, Y. ZHU, M. LEES, A. BOUKHANOVSKY and P. M. A. SLOOT (2015). Simulating city-level airborne infectious diseases. *Computers, Environment and Urban Systems*, 51, pp. 97–105 (cit. on p. 64).

MILI, R. Z., R STEINER and E OLADIMEJI (2006). DIVAs: illustrating an abstract architecture for agent-environment simulation systems. *Multiagent and Grid Systems - An International Journal*, 2, pp. 505–525 (cit. on p. 31).

MILLER, E. J. and M. J. ROORDA (2003). A prototype model of household activity/travel scheduling. *Transportation Research Record*, 1831(416), pp. 1–20 (cit. on p. 30).

MINAR, N. (1996). *The Swarm simulation system : A yoolkit for building multi-agent simulations*. Tech. rep. Santa Fe, NM: Santa Fe Institute, pp. 1–11 (cit. on pp. 6, 18).

MKANSI, M. and E. A. ACHEAMPONG (2012). Research philosophy debates and classifications: students' dilemma. *Electronic Journal of Business Research Methods*, 10(2), pp. 132–140 (cit. on p. 10).

MNISZEWSKI, S. M., S. Y. DEL VALLE, P. D. STROUD, J. M. RIESE and S. J. SYDORIAK (2008). Pandemic simulation of antivirals + school closures: buying time until strain-specific vaccine is available. *Computational and Mathematical Organization Theory*, 14(3), pp. 209–221 (cit. on pp. 114, 115).

MONTES, G (2012). Using artificial societies to understand the impact of teacher student match on academic performance: The case of same race effects. *Journal of Artificial Societies and Social Simulation*, 15(4), p. 8 (cit. on p. 18).

MONTGOMERY, M. and J. CASTERLINE (1996). Social learning, social influence, and new models of fertility. *Population and Development Review*, 22(1996), pp. 151–175 (cit. on p. 72).

MORSE, S. S. (1995). Factors in the emergence of infectious diseases. *Emerging infectious diseases*, 1(1), pp. 1–7 (cit. on p. 7).

MORTON, S. and P. REDMOND (2015). The odd one out: Revisiting and investigating the gender imbalance in ICT study choices. *Journal of Applied Computing and Information Technology*, 19(1), pp. 1–5 (cit. on pp. 12, 13).

MOSSONG, J., N. HENS, M. JIT, P. BEUTELS, K. AURANEN, R. MIKOLAJCZYK, M. MASSARI, S. SALMASO, G. S. TOMBA, J. WALLINGA, J. HEIJNE, M. SADKOWSKA-TODYS, M. ROSINSKA and W. J. EDMUNDS (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3), pp. 381–391 (cit. on pp. 37, 38, 42, 43, 48, 51).

MÜLLER, J. and M. PISCHEL (1993). *The Agent Architecture InteRRaP: Concept and Application*. DFKI (cit. on p. 22).

MYERS, M. (2013). *Qualitative Research in Business and Management*. SAGE Publications (cit. on p. 11).

NAGEL, K. and M. RICKERT (2001). Parallel implementation of the TRANSIMS microsimulation. *Parallel Computing*, 27, pp. 1611–1639 (cit. on pp. 6, 16, 18, 64).

NANCE, R. E. (1996). *A History of Discrete Event Simulation Programming Languages*. ACM (cit. on p. 130).

NAVARRO, L., F. FLACHER and V. CORRUBLE (2011). Dynamic level of detail for large scale agent-base urban simulations. *10th International Conference on Autonomous Agents and Multiagent Systems 2011*. Taipei, Taiwan: International Foundation for Autonomous Agents and Multiagent Systems, pp. 701–708 (cit. on pp. 16, 71, 127).

NEWELL, A., P. ROSENBLOOM and J. LAIRD (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33(1), pp. 1–64 (cit. on p. 22).

NORTH, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press (cit. on p. 32).

NORTH, M. J. and C. M. MACAL (2011). Product design patterns for agent-based modeling. *Proceedings of the 2011 Winter Simulation Conference (WSC)*. 1996. IEEE, pp. 3082–3093 (cit. on p. 9).

NWANA, H. S. (1996). Software agents: An overview. *The Knowledge Engineering Review*, 11(03), pp. 205–244 (cit. on p. 22).

OKUYAMA, F. Y., R. H. BORDINI and A. C. DA ROCHA COSTA (2009). A distributed normative infrastructure for situated multi-agent organisations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 5397. Springer-Verlag Berlin Heidelberg, pp. 29–46 (cit. on p. 67).

OMICINI, A., A. RICCI and M. VIROLI (2008). Artifacts in the A & A meta-model for multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 17(May), pp. 432–456 (cit. on p. 31).

ORLIKOWSKI, W. J. and J. J. BAROUDI (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1), pp. 1–28 (cit. on p. 10).

PAGE, C. L., F. BOUSQUET and I. BAKAM (2000). CORMAS: A multiagent simulation toolkit to model natural and social dynamics at multiple scales. *Proceedings of Workshop on "The ecology of scales"*. Wageningen, The Netherlands, pp. 1–20 (cit. on p. 5).

PALESHI, A and G. EVANS (2011). Simulation of mitigation strategies for a pandemic influenza. *Proceedings of the 2011 Winter Simulation Conference*. 2009. IEEE, pp. 1345–1353 (cit. on p. 115).

PAN, Z., A. D. CHEOK, H. YANG, J. ZHU and J. SHI (2006). Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics*, 30(1), pp. 20–28 (cit. on p. 16).

PARKER, J. and J. M. EPSTEIN (2011). A distributed platform for global-scale agent-based models of disease transmission. *ACM Transactions on Modeling and Computer Simulation*, 22(1), pp. 1–25 (cit. on pp. 7, 39, 42–44, 109).

PARKER, M. T. (2001). What is Ascape and why should you care. *Journal of Artificial Societies and Social Simulation*, 4(1), pp. 1–5 (cit. on p. 5).

PAS, E. I. (1984). The effect of selected sociodemographic characteristics on daily travel-activity behavior. *Environment and Planning A*, 16(5), pp. 571–581 (cit. on p. 48).

PASTOR-SATORRAS, R., C. CASTELLANO, P. VAN MIEGHEM and A. VESPIGNANI (2014). Epidemic processes in complex networks. *Review of Modern Physics*, 87(3), pp. 925–979 (cit. on pp. 8, 39).

PASTOR, O., V. PELECHANO, E. INSFRÁN and J. GÓMEZ (1998). From object oriented conceptual modeling to automated programming in java. *Conceptual Modeling-ER'98*. Springer, pp. 183–196 (cit. on p. 47).

PAUL, D. (2002). Agent based social simulation: A computer science view. *Journal of Artificial Societies and Social Simulation*, 5(1), pp. 1–7 (cit. on pp. 2, 3).

PEFFERS, K., T. TUUNANEN, M. A. ROTHENBERGER and S. CHATTERJEE (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), pp. 45–77 (cit. on p. 12).

PELECHANO, N, J. M. ALLBECK and N. I. BADLER (2007). Controlling individual agents in high-density crowd simulation. *Proceedings of the 2007 ACM SIGGRAPH Eurographics Symposium on Computer Animation*. Ed. by D. Metaxas and J. Popovic. Vol. 1. Aire-la-Ville, Switzerland, Switzerland: The Eurographics Association, pp. 1–10 (cit. on p. 18).

PEREZ, L. and S. DRAGICEVIC (Jan. 2009). An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8(50), pp. 1–17 (cit. on pp. 37, 39, 40, 42, 109).

POSLAD, S. (2007). Specifying protocols for multi-agent systems interaction. *ACM Transactions on Autonomous and Adaptive Systems*, 2(4), pp. 1–24 (cit. on pp. 30, 56).

PRATT, S. C., D. J. SUMPTER, E. B. MALLON and N. R. FRANKS (2005). An agent-based model of collective nest choice by the ant temnothorax albipennis. *Animal Behaviour*, 70(5), pp. 1023–1036 (cit. on p. 16).

PUJOL, J. and A FLACHE (2005). How can social networks ever become complex? Modelling the emergence of complex networks from local social exchanges. *Journal of Artificial Societies and Social Simulation*, 8(4), pp. 1–18 (cit. on pp. 28, 70).

RAILSBACK, S. F., S. L. LYTINEN and S. K. JACKSON (2006). Agent-based Simulation Platforms: Review and Development Recommendations. *Simulation*, 82(9), pp. 609–623 (cit. on pp. 4, 16, 19, 20).

RAKOWSKI, F., M. GRUZIEL, M. KRYCH and J. P. RADOMSKI (2010). Large Scale Daily Contacts and Mobility Model - an Individual-Based Countrywide Simulation Study for Poland. *Journal of Artificial Societies and Social Simulation*, 13(1), pp. 1–13 (cit. on pp. 7, 39, 42–44, 78, 109).

RANA, L. (2008). *Geographical Thought*. Concept Pub. Co. (cit. on p. 11).

RANEY, B. and K. NAGEL (2003). An agent-based simulation model of Swiss travel: first results. *Networks and Spatial Economics*, 3(1), pp. 23–42 (cit. on p. 64).

RANEY, B., A. VOELLMY and N. CETIN (2002). Towards a microscopic traffic simulation of all of Switzerland. *ICCS '02 Proceedings of the International Conference on Computational Science*. Springer-Verlag London, pp. 371–380 (cit. on p. 18).

RAO, A. S. (2009). AgentSpeak(L): BDI Agents speak out in a logical computable language. *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*. Secaucus, NJ, USA: Springer-Verlag New York, pp. 42–55 (cit. on p. 4).

RAZAVI, S. N., N. GAUD, N. MOZAYANI and A. KOUKAM (2010). Multi-agent based simulations using fast multipole method: application to large scale simulations of flocking dynamical systems. *Artificial Intelligence Review*, 35(1), pp. 53–72 (cit. on p. 16).

RITZER, G. (2001). *Explorations in Social Theory: From Metatheorizing to Rationalization*. SAGE Publications (cit. on p. 2).

ROBINSON, S. (2008). Conceptual modelling for simulation Part I: definition and requirements. *Journal of the Operational Research Society*, 59(3), pp. 278–290 (cit. on p. 9).

ROBINSON, S. (2010). Conceptual modelling: Who needs it. *SCS M&S Magazine*, 2, pp. 1–7 (cit. on p. 9).

ROMAN, S. and S. KATERINA (2012). The usability of agent-based simulation in decision support system of E-commerce architecture. *International Journal of Information Engineering and Electronic Business*, 4(1), pp. 10–17 (cit. on p. 16).

RONALD, N., T. ARENTZE and H. TIMMERMANS (2012a). Modeling social interactions between individuals for joint activity scheduling. *Transportation Research Part B: Methodological*, 46(2), pp. 276–290 (cit. on pp. 28, 30, 112).

RONALD, N., V. DIGNUM, C. JONKER, T. ARENTZE and H. TIMMERMANS (2012b). On the engineering of agent-based simulations of social activities with social networks. *Information and Software Technology*, 54(6), pp. 625–638 (cit. on p. 43).

RUCKER, R. (2006). *The Lifebox, the Seashell, and the Soul: What Gnarly Computation Taught Me About Ultimate Reality, the Meaning of Life, and How to Be Happy*. Basic Books (cit. on p. 1).

RUEF, M., H. E. ALDRICH and N. M. CARTER (2003). The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review*, 68(2), pp. 195–222 (cit. on p. 29).

RUSSELL, S. and P. NORVIG (2003). *Artificial Intelligence: A Modern Approach (2nd ed.)* Prentice Hall (cit. on p. 22).

SAID, L. B., T. BOURON and A. DROGOUL (2002). Agent-based interaction analysis of consumer behavior. *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: part 1*. New York, NY, USA: ACM, pp. 184–190 (cit. on p. 16).

SALAMON, L. M. (2002). *The Tools of Government: A Guide to the New Governance*. Oxford University Press (cit. on p. 67).

SARJOUGHIAN, H., B. ZEIGLER and S. HALL (2001). A layered modeling and simulation architecture for agent-based system development. *Proceedings of the IEEE*, 89(2), pp. 201–213 (cit. on p. 9).

SAUNDERS, M., P. LEWIS and A. THORNHILL (2012). *Research Methods for Business Students, 6/e*. Pearson Education India (cit. on pp. 9, 10).

SAVARIMUTHU, B. T. R., S. CRANEFIELD, M. PURVIS and M. PURVIS (2008). *Role Model Based Mechanism for Norm Emergence in Artificial Agent Societies*. Springer (cit. on pp. 8, 31).

SAVARIMUTHU, B. T. R., R. ARULANANDAM and M. PURVIS (2011). Aspects of active norm learning and the effect of lying on norm emergence in agent societies. *Agents in Principle, Agents in Practice*. Ed. by D. Kinny, J. Y. jen Hsu, G. Governatori and A. K. Ghose. Vol. 7047. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 36–50 (cit. on p. 32).

SAWYER, R. K. (Feb. 2003). Artificial societies: Multiagent systems and the micro-macro link in sociological theory. *Sociological Methods & Research*, 31(3), pp. 325–363 (cit. on p. 17).

SCHELLING, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, pp. 143–186 (cit. on p. 2).

SCHREINEMACHERS, P. and T. BERGER (2011). An agent-based simulation model of human–environment interactions in agricultural systems. *Environmental Modelling & Software*, 26(7), pp. 845–859 (cit. on p. 16).

SCHUTT, R. (2014). *Investigating the Social World: The Process and Practice of Research*. SAGE Publications (cit. on p. 1).

SCHWARZ, R. (1995). *Speedup Limits for Tightly-coupled Parallel Computations*. Springer (cit. on p. 130).

SCHWERDTFEGER, T. (1984). DYNEMO: A model for the simulation of traffic flow in motorway networks. *Proceedings of the 9th international Symposium on Transportation and Traffic Theory*. Utrecht, The Netherlands: VNU Science Press, pp. 65–87 (cit. on p. 130).

SHANNON, R. (1998). Introduction to the art and science of simulation. *Proceedings of the 1998 Winter Simulation Conference*, pp. 7–14 (cit. on pp. 13, 15).

SHOHAM, Y. (1993). *Agent-oriented programming* (cit. on pp. 4, 21).

SHOHAM, Y. (1997). An overview of agent-oriented programming. *Software Agents*, 4, pp. 271–290 (cit. on p. 4).

SHULGIN, B., L. STONE and Z. AGUR (1998). Pulse vaccination strategy in the SIR epidemic model. *Bulletin of Mathematical Biology*, 60(6), pp. 1123–1148 (cit. on p. 37).

SIBERTIN-BLANC, C., F AMBLARD and M MAILLIARD (2006). A coordination framework based on the Sociology of Organized Action. *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems*. Vol. 3913. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 3–17 (cit. on p. 29).

SIEBERS, P. O., C. M. MACAL, J GARNETT, D BUXTON and M PIDD (Sept. 2010). Discrete-event simulation is dead, long live agent-based simulation! *Journal of Simulation*, 4(3), pp. 204–210 (cit. on pp. 15, 18).

SIEBERS, P.-O., U. AICKELIN, H. CELIA and C. CLEGG (2007). A multi-agent simulation of retail management practices. *Proceedings of the 2007 Summer Computer Simulation Conference*. Society for Computer Simulation International, pp. 959–966 (cit. on p. 15).

SIERHUIS, M. (2001). Brahms: A multi-agent modeling and simulation language for work system analysis and design. PhD thesis. University of Amsterdam (cit. on pp. 5, 7).

SIERRA, C., N. R. JENNINGS, P. NORIEGA and S. PARSONS (1998). A framework for argumentation-based negotiation. *Proceedings of the 4th International Workshop on Intelligent Agents IV, Agent Theories, Architectures, and Languages.* Vol. 1365. Springer-Verlag London, pp. 177–192 (cit. on p. 30).

ŠIŠLÁK, D., P. VOLF, M. JAKOB and M. PĚCHOUČEK (2009). Distributed platform for large-scale agent-based simulations. *Trends in Techniques, Concepts and Design.* Vol. 5920. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 16–32 (cit. on p. 120).

SLOMAN, A., B. LOGAN and C. RICH (1999). Building cognitively rich agents. *Communications of the ACM,* 42(3), pp. 71–77 (cit. on p. 22).

SMITH, E. R. and F. R. CONREY (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review,* 11(1), pp. 87–104 (cit. on p. 6).

SQUAZZONI, F., W. JAGER and B. EDMONDS (Dec. 2013). Social simulation in the social sciences: A brief overview. *Social Science Computer Review,* 32(3), pp. 279–294 (cit. on pp. 1–3, 27, 29).

STEEL, T., D. KUIPER and R. Z. WENKSTERN (2010). Context-aware virtual agents in open environments. *2010 Sixth International Conference on Autonomic and Autonomous Systems.* Cancun: IEEE, pp. 90–96 (cit. on p. 67).

STEFANSSON, B. (1997). Swarm: An object oriented simulation platform applied to markets and organizations. *Evolutionary Programming VI.* Springer, pp. 59–71 (cit. on p. 18).

STOCKER, R, D. GREEN and D NEWTH (2001). Consensus and cohesion in simulated social networks. *Journal of Artificial Societies and Social Simulation,* 4(4), pp. 1–14 (cit. on p. 95).

STRAIF-BOURGEOIS, S., R. RATARD and M. KRETZSCHMAR (2014). Infectious disease epidemiology. *Handbook of Epidemiology.* Springer, pp. 2041–2119 (cit. on p. 36).

STROUD, P and S. D. VALLE (2007). Spatial dynamics of pandemic influenza in a massive artificial society. *Journal of Artificial Societies and Social Simulation,* 10(4), pp. 1–18 (cit. on pp. 6, 7, 39, 43, 44, 48, 78, 88–90, 98–101).

SUN, R. (2009). Cognitive architectures and multi-agent social simulation. *Multi-Agent Systems for Society.* Heidelberg: Springer-Verlag Berlin, pp. 7–21 (cit. on p. 24).

SUTCLIFFE, A., D. WANG and R. DUNBAR (2012). Social relationships and the emergence of social networks. *Journal of Artificial Societies and Social Simulation,* 15(4) (cit. on p. 70).

SUZUMURA, T, C HOUNGKAEW and H KANEZASHI (2014). Towards billion-scale social simulations. *Proceedings of the 2014 Winter Simulation Conference.* Piscataway, NJ, USA: IEEE, pp. 781–792 (cit. on p. 7).

SUZUMURA, T. and H. KANEZASHI (2012). Highly scalable X10-based agent simulation platform and its application to large-scale traffic simulation. *2012 IEEE/ACM 16th International Symposium on Distributed Simulation and Real Time Applications.* Dublin: IEEE, pp. 243–250 (cit. on p. 18).

SWANSON, J. (2010). *God, Science and the Universe: The Integration of Religion and Science.* Strategic Book Publishing (cit. on p. 11).

SWARUP, S., S. EUBANK and M. MARATHE (2014). Computational epidemiology as a challenge domain for multiagent systems. *13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014).* Ed. by A. Lomuscio, P. Scerri, A. Bazzan and M. Huhns. Paris, France: IFAAMAS, pp. 1173–1176 (cit. on p. 41).

TA, N., Y. CHAI and M. GUAN (2015). Suburbanization, daily lifestyle and space-behavior interaction in Beijing (In Chinese). *Acta Geographica Sinica*, 70(8), pp. 1271–1280 (cit. on pp. 82, 83).

TAKEUCHI, Y., X. LIU and J. CUI (2007). Global dynamics of SIS models with transport-related infection. *Journal of Mathematical Analysis and Applications*, 329, pp. 1460–1471 (cit. on p. 42).

TAMBE, M. (1997). Agent architectures for flexible, practical teamwork. *National Conference on Artificial Intelligence (AAAI)*. AAAI Press, pp. 22–28 (cit. on p. 23).

TATARA, E., M. NORTH, T. HOWE, N. COLLIER and J. VOS (2006). An introduction to Repast modeling by using a simple predator-prey example. *Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects*. Argonne, IL USA, pp. 83–94 (cit. on pp. 5, 7, 18).

TAYLOR, S. (2014). *Agent-based Modeling and Simulation*. Palgrave Macmillan (cit. on p. 5).

TESFATSION, L. (2001). Introduction to the special issue on agent-based computational economics. *Journal of Economic Dynamics and Control*, 25(3), pp. 281–293 (cit. on p. 16).

THEODOROPOULOS, G. and Y. ZHANG (2006). Large scale distributed simulation on the grid. *Proceedings of the Sixth IEEE International Symposium on Cluster Computing and the Grid Workshops (CCGRIDW'06)*. Singapore: IEEE, pp. 1–8 (cit. on p. 7).

TOBIAS, R. and C. HOFMANN (2004). Evaluation of free Java-libraries for social-scientific agent based simulation. *Journal of Artificial Societies and Social Simulation*, 7(1) (cit. on p. 19).

TOIVONEN, R., L. KOVANEN, M. KIVELÄ, J.-P. ONNELA, J. SARAMÄKI and K. KASKI (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks*, 31(4), pp. 240–254 (cit. on p. 29).

TOROCZKAI, Z. and H. GUCLU (2007). Proximity networks and epidemics. *Physica A: Statistical Mechanics and its Applications*, 378(1), pp. 68–75 (cit. on p. 90).

TRIANDIS, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, 96(3) (cit. on p. 78).

TYSZER, J. (2012). *Object-oriented Computer Simulation of Discrete-event Systems*. Vol. 10. Springer Science & Business Media (cit. on p. 4).

UHRMACHER, A. M. and K. GUGLER (2000). Distributed, parallel simulation of multiple, deliberative agents. *Proceedings of the Fourteenth Workshop on Parallel and Distributed Simulation*. Bologne: IEEE, pp. 101–108 (cit. on p. 4).

UHRMACHER, A. M., P TYSCHLER and D TYSCHLER (1997). Concepts of object-and agent-oriented simulation. *Transactions of the Society for Computer Simulation*, 14(2), pp. 59–67 (cit. on p. 8).

VAISHNAVI, V. K., W. KUECHLER and W. KUECHLER JR (2007). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. CRC Press (cit. on p. 12).

VAN DE COEVERING, P. and T. SCHWANEN (2006). Re-evaluating the impact of urban form on travel patterns in Europe and North-America. *Transport Policy*, 13(3), pp. 229–239 (cit. on p. 78).

VASCONCELOS, W. W., A. GARCÍA-CAMINO, D. GAERTNER, J. A. RODRÍGUEZ-AGUILAR and P. NORIEGA (2012). Distributed norm management for multi-agent systems. *Expert Systems with Applications*, 39(5), pp. 5990–5999 (cit. on p. 32).

Viganò, F., N. Fornara and M. Colombetti (2006). An event driven approach to norms in artificial institutions. *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems*. Springer, pp. 142–154 (cit. on p. 32).

Vikhorev, K., N. Alechina and B. Logan (2009). The ARTS real-time agent architecture. *CEUR Workshop Proceedings*. Vol. 494. Springer Berlin Heidelberg, pp. 286–292 (cit. on p. 24).

WHO (2007). *The World Health Report 2007: a safer future: global public health security in the 21st century*. Tech. rep. WHO (cit. on p. 6).

Waddell, P. (2012). Introduction to Urban Simulation: Design and Development of Operational Models. *Handbook of Transport Geography and Spatial Systems*. Ed. by D. A. Hensher, K. J. Button, K. E. Haynes and P. R. Stopher. Elsevier, pp. 1–35 (cit. on p. 16).

Walker, L. (2009). Implementing a large scale social simulation using the New Zealand BeSTGRID computer network : A case study. *Proceeding of the 18th World IMACS / MODSIM Congress*. Cairns, Australia, pp. 1073–1079 (cit. on p. 16).

Wan, H. and J. Cui (2007). An SEIS epidemic model with transport-related infection. *Journal of Theoretical Biology*, 247, pp. 507–524 (cit. on p. 42).

Wang, D., Y. Chai and F. Li (2011). Built environment diversities and activity–travel behaviour variations in Beijing, China. *Journal of Transport Geography*, 19(6), pp. 1173–1186 (cit. on pp. 92, 94).

Wang, K.-Y. (2014). How change of public transportation usage reveals fear of the SARS virus in a city. *PloS one*, 9(3) (cit. on p. 41).

Wang, X. and A. Collins (2014). Popularity or proclivity? Revisiting agent heterogeneity in network formation. *Proceedings of the 2014 Winter Simulation Conference*. Savanah, GA: IEEE, pp. 3084–3095 (cit. on p. 70).

Wang, Y, M Lees and W Cai (2012). Grid-based partitioning for large-scale distributed agent-based crowd simulation. *2012 Winter Simulation Conference*. Berlin: IEEE, pp. 1–12 (cit. on pp. 4, 7).

Ward, R. L. and W. V. Huang (1992). Simulation with object oriented programming. *Computers & Industrial Engineering*, 23(1), pp. 219–222 (cit. on p. 4).

Weber, R. (2004). The rhetoric of positivism versus interpretivism: A personal view. *MIS Quarterly*, 28(1), pp. 3–7 (cit. on p. 11).

Wei, C. and K. V. Hindriks (2013). An agent-based cognitive robot architecture. *Programming Multi-Agent Systems*. Ed. by M. Dastani, J. F. Hubner and B. Logan. Vol. 7837. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 54–71 (cit. on p. 5).

Weidlich, A. and D. Veit (2008). A critical survey of agent-based wholesale electricity market models. *Energy Economics*, 30(4), pp. 1728–1759 (cit. on p. 16).

Weiss, G. (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, p. 619 (cit. on p. 28).

Wilenski, U. (1999). *Netlogo*. Tech. rep. Center for Connected Learning and Computer-Based Modeling (cit. on pp. 6, 8, 18).

Wooldridge, M. and N Jennings (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2), pp. 115–152 (cit. on pp. 4, 22).

Xu, F., C. Connell McCluskey and R. Cressman (2013). Spatial spread of an epidemic through public transportation systems with a hub. *Mathematical Biosciences*, 246(1), pp. 164–175 (cit. on p. 42).

Xu, R. and Z. Ma (2009). Stability of a delayed SIRS epidemic model with a nonlinear incidence rate. *Chaos, Solitons & Fractals*, 41(5), pp. 2319–2325 (cit. on p. 38).

YUE, S., C. YANWEI et al. (2013). Daily activity space of suburban mega-community residents in Beijing based on GPS data (In Chinese). *Acta Geographica Sinica*, 68(4), pp. 506–516 (cit. on p. 83).

ZAWADOWSKI, A., R KARADI and J KERTESZ (2002). Price drops, fluctuations, and correlation in a multi-agent model of stock markets. *Physica A: Statistical Mechanics and its Applications*, 316(1), pp. 403 –412 (cit. on p. 16).

ZEIGLER, B. and R. RADA (1984). Abstraction in methodology: A framework for computer support. *Information Processing & Management*, 20(1-2), pp. 63–79 (cit. on p. 4).

ZEIGLER, B. P., Y. MOON, D. KIM and G. BALL (1997). The DEVS environment for high-performance modeling and simulation. *Computing in Science & Engineering*, (3), pp. 61–71 (cit. on p. 8).

ZEIGLER, B. P., H. PRAEHOFER and T. G. KIM (2000). *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. Academic Press (cit. on pp. 4, 130).

ZHANG, B, W. K. V. CHAN and S. V. UKKUSURI (2014). On the modelling of transportation evacuation: an agent-based discrete-event hybrid-space approach. *Journal of Simulation*, 8(4), pp. 1–12 (cit. on p. 47).

ZHANG, L., G. CHANG, S. ZHU and C. XIONG (2012). Integrating an agent-based travel behavior model with large-scale microscopic traffic simulation for corridor-level and subarea transportation operations and planning applications. *Journal of Urban Planning and Development*, 139(2), pp. 94–103 (cit. on p. 18).

ZHANG, L., W. YANG, J. WANG and Q. RAO (2013). Large-scale agent-based transport simulation in Shanghai, China. *Transportation Research Record: Journal of the Transportation Research Board*, 2399, pp. 34–43 (cit. on pp. 18, 64).

ZHANG, M., A. VERBRAECK, R. MENG and X. QIU (2015). Agent-based Modeling of Large-scale Complex Social Interactions. *Proceedings of the 3rd ACM Conference on Principles of Advanced Discrete Simulation*. ACM. New York, NY, USA, pp. 197–198 (cit. on p. 30).

ZHAO, P., B. LÜ and G. DE ROO (2011). Impact of the jobs-housing balance on urban commuting in Beijing in the transformation era. *Journal of Transport Geography*, 19(1), pp. 59–69 (cit. on p. 96).

ZHAO, Y. and A. W. SADEK (2012). Large-scale agent-based traffic micro-simulation: Experiences with model refinement, calibration, validation and application. *Procedia Computer Science*, 10(2), pp. 815–820 (cit. on p. 18).

ZHENG, H., Y.-J. SON, Y.-C. CHIU, L. HEAD, Y. FENG, H. XI, S. KIM and M. HICKMAN (2013). *A Primer for Agent-Based Simulation and Modeling in Transportation Applications*. Tech. rep. (cit. on p. 15).

ZHOU, J., N. N. CHUNG, L. Y. CHEW and C. H. LAI (2012). Epidemic spreading induced by diversity of agents' mobility. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 86, pp. 1–7 (cit. on pp. 40, 42).

ZIA, K, K FARRAHI, a RIENER and a FERSCHA (2013). An agent-based parallel geo-simulation of urban mobility during city-scale evacuation. *Simulation*, 89(10), pp. 1184–1214 (cit. on pp. 7, 18).

ZIA, K., A. RIENER, K. FARRAHI and A. FERSCHA (2012). A new opportunity to urban evacuation analysis: Very large scale simulations of social agent systems in Repast HPC. *2012 ACM/IEEE/SCS 26th Workshop on Principles of Advanced and Distributed Simulation*. Zhangjiajie, China: IEEE, pp. 233–242 (cit. on pp. 7, 17).

ZUCKERMAN, I., S. KRAUS and J. S. ROSENSCHEIN (2012). The adversarial activity model for bounded rational agents. *Autonomous Agents and Multi-Agent Systems*, 24(3), pp. 374–409 (cit. on p. 23).

ZUK, T., F. RAKOWSKI and J. P. RADOMSKI (2009). Probabilistic model of influenza virus transmissibility at various temperature and humidity conditions. *Computational Biology and Chemistry*, 33(4), pp. 339–43 (cit. on p. 69).

List of Tables

List of Figures

Summary

Large-scale agent-based social simulation is gradually proving to be a versatile methodological approach for studying human societies, which could make contributions from policy making in social science, to distributed artificial intelligence and agent technology in computer science, and to theory and modeling practice in computer simulation systems. Simultaneously, the application areas of large-scale agent-based social simulation vary a lot as well, from daily transportation on a city/country level, to large-scale emergency response, prediction of social change, and analysis of social structure.

However, current large-scale agent-based social simulation practice is facing difficulties in balancing model complexity and simulation performance. The wide adoption of distributed/parallel mechanism in current large-scale agent-based social simulation has proven to be an efficient solution to achieve system performance and scalability. On the other hand, the trade-off is usually the simplification of the model precision including agent behavior, agent environment and the social networks and interactions, which are proven to be important to understand social phenomena in complex social systems.

Based on the existing challenges, this thesis introduces a novel conceptual model for large-scale agent-based social simulation development, gives out the reference implementation of the proposed model components, and presents a simulation study of a case of epidemic prediction and control in the city of Beijing. This conceptual model can be considered as a hybrid model mixing a general agent-based conceptual model and the discrete event simulation paradigm.

For the concept of agent in the proposed conceptual model, this thesis presents a new way for implementation. A reference implementation of an agent is constituted by three main parts: (1) agent object, (2) activity pattern, and (3) multi-level decision-making module. With this design, the implemented agents can carry out a lot of complex activities and show diverse behaviors, such as traveling around and joining non-predefined social activities, while staying "simple" and "small" enough for scalability consideration.

For the concept of a social network, this thesis presents a new method to generate social networks dynamically for simulating interactions among a group of agents on a large scale during a simulation run. This thesis borrows from the concept of 'social reach' in a social circle model, and proposes the concept of 'social

similarity' to generate the special type of social networks, friendship. Using the generated entire social network, agents in this model are able to communicate for scheduling joint social activities. When executing joint social activities, a functional entity called 'activity group' is generated to organize and manage the participants, and a social contact network emerges from the execution.

Compared to the concept of agent environments in general ABM conceptual models, the introduced conceptual model separates the concept of an agent environment into physical container, social regulation and functional entity, which overcomes the limitations on environmental completeness in other ABM models and provides flexibility in simulating different system scenarios.

The concept of a physical container is introduced to represent the physical environment where agents stay. Typical physical containers are school, classroom, office, bedroom, train, etc. Physical containers are organized hierarchically. Moreover, this concept makes it much easier to include a transportation component in a social simulation model, which is achieved by considering vehicles as movable physical containers in the model.

The concept of social regulation, borrowed from the multi-agent system community, is used to model artifacts that can guide and influence agent behavior globally (rules/norms/institutions). With this concept, agents can respond to different situations during a simulation run. For example, regulating agents' behavior during a disease outbreak is an indispensable part at a large-scale agent-based epidemic simulation. How agents would respond to interventions during a disease outbreak would have a big impact on the model outcomes.

The concept of functional entity, borrowed from the object-oriented paradigm, is used to model the extra objects in the system that can influence or directly change attributes of either agents, physical containers or social regulations. For example, a disease is modeled as a functional entity to change agents' healthy status. Temperature can be modeled as a functional entity to change the transmission probability of a disease in a specified location (physical container).

Using reference implementations of these concepts, a model of a large-scale artificial city of Beijing is constructed as a case study to test policies for controlling the spread of disease among the full population (19.6 million). This case study can be considered as a proof of concept which exemplifies how large-scale social systems with complex human behavior and social interactions can be modeled with the help of the proposed conceptual model, but still gains reasonable performance. It also indicates potential use in other social science areas, such as microscopic transportation systems and city level evacuation planning.

Samenvatting (in Dutch)

Grootschalige agentgebaseerde sociale simulatie is uitgegroeid tot een veelzijdige methodologische aanpak voor het bestuderen van samenlevingen, die onder andere kan bijdragen aan het opstellen van beleid (vanuit de sociale wetenschappen), aan gedistribueerde kunstmatige intelligentie en agenttechnologie in de computerwetenschappen en aan de theorie en modelleringspraktijk van computergebaseerde simulatiesystemen. Er zijn vele verschillende toepassingsgebieden van grootschalige agentgebaseerde sociale simulatie, zoals het dagelijks transport op het niveau van een stad of een land, grootschalige rampenbestrijding, voorspelling van sociale veranderingen en analyse van sociale structuren.

Eén van de moeilijkheden bij de toepassing van grootschalige agentgebaseerde sociale simulatie is de balans tussen de complexiteit van het model en de doorlooptijd van de simulatie. Er bestaan goede gedistribueerde en parallelle oplossingen die een efficiënte oplossing bieden voor het bereiken van zowel systeemprestaties als schaalbaarheid. Echter, het gebruik van deze mechanismen leidt meestal tot het sterk vereenvoudigen van de modellen waarbij ook het gedrag van de agent, de omgeving van de agent en de sociale netwerken en interacties vereenvoudigd worden. Bij het bestuderen en begrijpen van verschijnselen in complexe systemen zijn dit juist de belangrijke eigenschappen.

Om dit probleem op te lossen introduceert dit proefschrift een nieuw conceptueel model voor de ontwikkeling van grootschalige agentgebaseerde sociale simulatie en geeft het een referentie-implementatie van de voorgestelde componenten voor het model. Als test wordt het model toegepast in een simulatiestudie waarin de verspreiding van epidemieën in de stad Beijing en beleid om de verspreiding tegen te gaan bestudeerd worden. Het conceptuele model is een hybride model waarin een algemeen agentgebaseerd conceptueel model met een discreet-simulatieparadigma gecombineerd wordt.

Het conceptuele model bevat een nieuwe implementatie van het concept agent. De referentie-implementatie van een agent heeft drie hoofdonderdelen: (1) het agent-object, (2) het patroon van activiteiten en (3) de multi-level besluitvormingsmodule. In dit ontwerp kunnen de geïmplementeerde agenten veel verschillende complexe activiteiten uitvoeren en verschillend gedrag vertonen, zoals rondreizen en aan niet-voorgedefinieerde sociale activiteiten deelnemen. Tegelijkertijd zijn ze "simpel" en "klein" genoeg om schaalbaarheid te garanderen.

Voor het concept sociaal netwerk wordt een nieuwe methode aangereikt waarmee sociale netwerken voor interacties tussen agenten op grote schaal dynamisch gegenereerd kunnen worden tijdens een simulatierun. Het proefschrift bouwt voort op het concept 'sociale reikwijdte' uit de sociale wetenschappen, en stelt het concept 'sociale overeenkomst' voor om vriendschapsnetwerken als een speciaal type sociaal netwerk te genereren. Door deze gegeneerde sociale netwerken hebben agenten de mogelijkheid om gezamenlijke activiteiten te organiseren. Wanneer deze gezamenlijke activiteiten plaatsvinden, wordt een functionele entiteit, die 'activiteitengroep' genoemd wordt, gegenereerd die de deelnemers bijeen probeert te roepen en aanstuurt. Met deze uitvoering ontstaan netwerken van sociale contacten in het model.

Anders dan in andere agentgebaseerde conceptuele modellen, wordt de agentomgeving in het voorgestelde conceptuele model opgedeeld in een fysieke container, sociale regels en een functionele entiteit. Dit vermindert de beperkingen van compleetheid van de omgeving binnen andere agentgebaseerde modellen en biedt flexibiliteit bij het simuleren van verschillende scenario's.

Het concept fysieke container representeert de fysieke omgeving waar de agenten zich bevinden. Typische fysieke containers zijn bijvoorbeeld school, klaslokaal, slaapkamer enzovoort. Fysieke containers zijn hiërarchisch gestructureerd, waar klaslokalen onderdeel uitmaken van een school. Dit concept maakt het ook makkelijker om een transportcomponent aan een sociaal simulatiemodel toe te voegen. Dit wordt geïmplementeerd door voertuigen als bewegende fysieke containers te beschouwen.

Het concept sociale regels is gebruikt om artefacten te modellen die het gedrag van actoren kunnen leiden en beïnvloeden (regels / normen / instituties). Met dit concept kunnen agenten tijdens een simulatierun reageren op verschillende situaties. Het reguleren van het gedrag van agenten tijdens een ziekte-uitbraak is bijvoorbeeld onmisbaar in een grootschalige agentgebaseerde epidemische simulatie. De manier waarop agenten reageren op verspreiding van de ziekte en afgekondigde maatregelen tijdens een uitbraak heeft veel invloed op de modeluitkomsten.

Het concept functionele entiteit, geleend uit het object georiënteerde paradigma, is gebruikt om extra objecten in het systeem te modelleren die attributen van agenten, fysieke containers en sociale regulaties kunnen beïnvloeden of veranderen. Zo wordt bijvoorbeeld een ziekte gemodelleerd als een functionele entiteit die de gezondheidstoestand van een agent verandert. En temperatuur kan gemodelleerd worden als een functionele entiteit die de kans op oplopen van een ziekte op een specifieke locatie (in een fysieke container) verandert.

Met de referentie-implementatie van deze concepten is een grootschalig model van de stad Beijing opgesteld waarmee verschillende beleidsopties om de verspreiding van een ziekte onder controle te krijgen getest kunnen worden. Deze case study kan beschouwd worden als een proof of concept die laat zien hoe een grootschalig sociaal systeem met complex menselijk gedrag en sociale interacties gemodelleerd kan worden. Het voorgestelde conceptuele model leidt tot een simulatie met goede uitkomsten en een redelijke snelheid. De resultaten wijzen ook op mogelijkheden om het conceptueel model te gebruiken in andere gebieden uit de sociale wetenschap, zoals bestuderen van transportsystemen en evacuatieplanning.

SIKS Dissertation Series

1998

1 Johan van den Akker (CWI) *DEGAS: An Active, Temporal Database of Autonomous Objects*
2 Floris Wiesman (UM) *Information Retrieval by Graphically Browsing Meta-Information*
3 Ans Steuten (TUD) *A Contribution to the Linguistic Analysis of Business Conversations*
4 Dennis Breuker (UM) *Memory versus Search in Games*
5 E. W. Oskamp (RUL) *Computerondersteuning bij Straftoemeting*

1999

1 Mark Sloof (VUA) *Physiology of Quality Change Modelling: Automated modelling of*
2 Rob Potharst (EUR) *Classification using decision trees and neural nets*
3 Don Beal (UM) *The Nature of Minimax Search*
4 Jacques Penders (UM) *The practical Art of Moving Physical Objects*
5 Aldo de Moor (KUB) *Empowering Communities: A Method for the Legitimate User-Driven*
6 Niek J. E. Wijngaards (VUA) *Re-design of compositional systems*
7 David Spelt (UT) *Verification support for object database design*
8 Jacques H. J. Lenting (UM) *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism*

2000

1 Frank Niessink (VUA) *Perspectives on Improving Software Maintenance*
2 Koen Holtman (TUe) *Prototyping of CMS Storage Management*
3 Carolien M. T. Metselaar (UvA) *Sociaal-organisatorische gevolgen van kennistechnologie*
4 Geert de Haan (VUA) *ETAG, A Formal Model of Competence Knowledge for User Interface*
5 Ruud van der Pol (UM) *Knowledge-based Query Formulation in Information Retrieval*
6 Rogier van Eijk (UU) *Programming Languages for Agent Communication*
7 Niels Peek (UU) *Decision-theoretic Planning of Clinical Patient Management*
8 Veerle Coupé (EUR) *Sensitivity Analyis of Decision-Theoretic Networks*
9 Florian Waas (CWI) *Principles of Probabilistic Query Optimization*
10 Niels Nes (CWI) *Image Database Management System Design Considerations, Algorithms and Architecture*
11 Jonas Karlsson (CWI) *Scalable Distributed Data Structures for Database Management*

2001

1 Silja Renooij (UU) *Qualitative Approaches to Quantifying Probabilistic Networks*
2 Koen Hindriks (UU) *Agent Programming Languages: Programming with Mental Models*
3 Maarten van Someren (UvA) *Learning as problem solving*
4 Evgueni Smirnov (UM) *Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets*
5 Jacco van Ossenbruggen (VUA) *Processing Structured Hypermedia: A Matter of Style*
6 Martijn van Welie (VUA) *Task-based User Interface Design*
7 Bastiaan Schonhage (VUA) *Diva: Architectural Perspectives on Information Visualization*
8 Pascal van Eck (VUA) *A Compositional Semantic Structure for Multi-Agent Systems Dynamics*
9 Pieter Jan 't Hoen (RUL) *Towards Distributed Development of Large Object-Oriented Models*

10 Maarten Sierhuis (UvA) *Modeling and Simulating Work Practice*
11 Tom M. van Engers (VUA) *Knowledge Management*

2002

1 Nico Lassing (VUA) *Architecture-Level Modifiability Analysis*
2 Roelof van Zwol (UT) *Modelling and searching web-based document collections*
3 Henk Ernst Blok (UT) *Database Optimization Aspects for Information Retrieval*
4 Juan Roberto Castelo Valdueza (UU) *The Discrete Acyclic Digraph Markov Model in Data Mining*
5 Radu Serban (VUA) *The Private Cyberspace Modeling Electronic*
6 Laurens Mommers (UL) *Applied legal epistemology: Building a knowledge-based ontology of*
7 Peter Boncz (CWI) *Monet: A Next-Generation DBMS Kernel For Query-Intensive*
8 Jaap Gordijn (VUA) *Value Based Requirements Engineering: Exploring Innovative*
9 Willem-Jan van den Heuvel (KUB) *Integrating Modern Business Applications with Objectified Legacy*
10 Brian Sheppard (UM) *Towards Perfect Play of Scrabble*
11 Wouter C. A. Wijngaards (VUA) *Agent Based Modelling of Dynamics: Biological and Organisational Applications*
12 Albrecht Schmidt (UvA) *Processing XML in Database Systems*
13 Hongjing Wu (TUe) *A Reference Architecture for Adaptive Hypermedia Applications*
14 Wieke de Vries (UU) *Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems*
15 Rik Eshuis (UT) *Semantics and Verification of UML Activity Diagrams for Workflow Modelling*
16 Pieter van Langen (VUA) *The Anatomy of Design: Foundations, Models and Applications*
17 Stefan Manegold (UvA) *Understanding, Modeling, and Improving Main-Memory Database Performance*

2003

1 Heiner Stuckenschmidt (VUA) *Ontology-Based Information Sharing in Weakly Structured Environments*
2 Jan Broersen (VUA) *Modal Action Logics for Reasoning About Reactive Systems*
3 Martijn Schuemie (TUD) *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*
4 Milan Petkovic (UT) *Content-Based Video Retrieval Supported by Database Technology*
5 Jos Lehmann (UvA) *Causation in Artificial Intelligence and Law: A modelling approach*
6 Boris van Schooten (UT) *Development and specification of virtual environments*
7 Machiel Jansen (UvA) *Formal Explorations of Knowledge Intensive Tasks*
8 Yongping Ran (UM) *Repair Based Scheduling*
9 Rens Kortmann (UM) *The resolution of visually guided behaviour*
10 Andreas Lincke (UvT) *Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture*
11 Simon Keizer (UT) *Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks*
12 Roeland Ordelman (UT) *Dutch speech recognition in multimedia information retrieval*
13 Jeroen Donkers (UM) *Nosce Hostem: Searching with Opponent Models*
14 Stijn Hoppenbrouwers (KUN) *Freezing Language: Conceptualisation Processes across ICT-Supported Organisations*
15 Mathijs de Weerdt (TUD) *Plan Merging in Multi-Agent Systems*
16 Menzo Windhouwer (CWI) *Feature Grammar Systems: Incremental Maintenance of Indexes to Digital Media Warehouses*
17 David Jansen (UT) *Extensions of Statecharts with Probability, Time, and Stochastic Timing*
18 Levente Kocsis (UM) *Learning Search Decisions*

2004

1 Virginia Dignum (UU) *A Model for Organizational Interaction: Based on Agents, Founded in Logic*
2 Lai Xu (UvT) *Monitoring Multi-party Contracts for E-business*
3 Perry Groot (VUA) *A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving*
4 Chris van Aart (UvA) *Organizational Principles for Multi-Agent Architectures*
5 Viara Popova (EUR) *Knowledge discovery and monotonicity*

6 Bart-Jan Hommes (TUD) *The Evaluation of Business Process Modeling Techniques*
7 Elise Boltjes (UM) *Voorbeeldig onderwijs: voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes*
8 Joop Verbeek (UM) *Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële gegevensuitwisseling en digitale expertise*
9 Martin Caminada (VUA) *For the Sake of the Argument: explorations into argument-based reasoning*
10 Suzanne Kabel (UvA) *Knowledge-rich indexing of learning-objects*
11 Michel Klein (VUA) *Change Management for Distributed Ontologies*
12 The Duy Bui (UT) *Creating emotions and facial expressions for embodied agents*
13 Wojciech Jamroga (UT) *Using Multiple Models of Reality: On Agents who Know how to Play*
14 Paul Harrenstein (UU) *Logic in Conflict. Logical Explorations in Strategic Equilibrium*
15 Arno Knobbe (UU) *Multi-Relational Data Mining*
16 Federico Divina (VUA) *Hybrid Genetic Relational Search for Inductive Learning*
17 Mark Winands (UM) *Informed Search in Complex Games*
18 Vania Bessa Machado (UvA) *Supporting the Construction of Qualitative Knowledge Models*
19 Thijs Westerveld (UT) *Using generative probabilistic models for multimedia retrieval*
20 Madelon Evers (Nyenrode) *Learning from Design: facilitating multidisciplinary design teams*

2005

1 Floor Verdenius (UvA) *Methodological Aspects of Designing Induction-Based Applications*
2 Erik van der Werf (UM) *AI techniques for the game of Go*
3 Franc Grootjen (RUN) *A Pragmatic Approach to the Conceptualisation of Language*
4 Nirvana Meratnia (UT) *Towards Database Support for Moving Object data*
5 Gabriel Infante-Lopez (UvA) *Two-Level Probabilistic Grammars for Natural Language Parsing*
6 Pieter Spronck (UM) *Adaptive Game AI*
7 Flavius Frasincar (TUe) *Hypermedia Presentation Generation for Semantic Web Information Systems*
8 Richard Vdovjak (TUe) *A Model-driven Approach for Building Distributed Ontology-based Web Applications*
9 Jeen Broekstra (VUA) *Storage, Querying and Inferencing for Semantic Web Languages*
10 Anders Bouwer (UvA) *Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments*
11 Elth Ogston (VUA) *Agent Based Matchmaking and Clustering: A Decentralized Approach to Search*
12 Csaba Boer (EUR) *Distributed Simulation in Industry*
13 Fred Hamburg (UL) *Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen*
14 Borys Omelayenko (VUA) *Web-Service configuration on the Semantic Web: Exploring how semantics meets pragmatics*
15 Tibor Bosse (VUA) *Analysis of the Dynamics of Cognitive Processes*
16 Joris Graaumans (UU) *Usability of XML Query Languages*
17 Boris Shishkov (TUD) *Software Specification Based on Re-usable Business Components*
18 Danielle Sent (UU) *Test-selection strategies for probabilistic networks*
19 Michel van Dartel (UM) *Situated Representation*
20 Cristina Coteanu (UL) *Cyber Consumer Law, State of the Art and Perspectives*
21 Wijnand Derks (UT) *Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics*

2006

1 Samuil Angelov (TUe) *Foundations of B2B Electronic Contracting*
2 Cristina Chisalita (VUA) *Contextual issues in the design and use of information technology in organizations*
3 Noor Christoph (UvA) *The role of metacognitive skills in learning to solve problems*
4 Marta Sabou (VUA) *Building Web Service Ontologies*
5 Cees Pierik (UU) *Validation Techniques for Object-Oriented Proof Outlines*
6 Ziv Baida (VUA) *Software-aided Service Bundling: Intelligent Methods & Tools for Graphical Service Modeling*
7 Marko Smiljanic (UT) *XML schema matching: balancing efficiency and effectiveness by means of clustering*

8 Eelco Herder (UT) *Forward, Back and Home Again: Analyzing User Behavior on the Web*
9 Mohamed Wahdan (UM) *Automatic Formulation of the Auditor's Opinion*
10 Ronny Siebes (VUA) *Semantic Routing in Peer-to-Peer Systems*
11 Joeri van Ruth (UT) *Flattening Queries over Nested Data Types*
12 Bert Bongers (VUA) *Interactivation: Towards an e-cology of people, our technological environment, and the arts*
13 Henk-Jan Lebbink (UU) *Dialogue and Decision Games for Information Exchanging Agents*
14 Johan Hoorn (VUA) *Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change*
15 Rainer Malik (UU) *CONAN: Text Mining in the Biomedical Domain*
16 Carsten Riggelsen (UU) *Approximation Methods for Efficient Learning of Bayesian Networks*
17 Stacey Nagata (UU) *User Assistance for Multitasking with Interruptions on a Mobile Device*
18 Valentin Zhizhkun (UvA) *Graph transformation for Natural Language Processing*
19 Birna van Riemsdijk (UU) *Cognitive Agent Programming: A Semantic Approach*
20 Marina Velikova (UvT) *Monotone models for prediction in data mining*
21 Bas van Gils (RUN) *Aptness on the Web*
22 Paul de Vrieze (RUN) *Fundaments of Adaptive Personalisation*
23 Ion Juvina (UU) *Development of Cognitive Model for Navigating on the Web*
24 Laura Hollink (VUA) *Semantic Annotation for Retrieval of Visual Resources*
25 Madalina Drugan (UU) *Conditional log-likelihood MDL and Evolutionary MCMC*
26 Vojkan Mihajlovic (UT) *Score Region Algebra: A Flexible Framework for Structured Information Retrieval*
27 Stefano Bocconi (CWI) *Vox Populi: generating video documentaries from semantically annotated media repositories*
28 Borkur Sigurbjornsson (UvA) *Focused Information Access using XML Element Retrieval*

2007

1 Kees Leune (UvT) *Access Control and Service-Oriented Architectures*
2 Wouter Teepe (RUG) *Reconciling Information Exchange and Confidentiality: A Formal Approach*
3 Peter Mika (VUA) *Social Networks and the Semantic Web*
4 Jurriaan van Diggelen (UU) *Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach*
5 Bart Schermer (UL) *Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance*
6 Gilad Mishne (UvA) *Applied Text Analytics for Blogs*
7 Natasa Jovanovic' (UT) *To Whom It May Concern: Addressee Identification in Face-to-Face Meetings*
8 Mark Hoogendoorn (VUA) *Modeling of Change in Multi-Agent Organizations*
9 David Mobach (VUA) *Agent-Based Mediated Service Negotiation*
10 Huib Aldewereld (UU) *Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols*
11 Natalia Stash (TUe) *Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System*
12 Marcel van Gerven (RUN) *Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty*
13 Rutger Rienks (UT) *Meetings in Smart Environments: Implications of Progressing Technology*
14 Niek Bergboer (UM) *Context-Based Image Analysis*
15 Joyca Lacroix (UM) *NIM: a Situated Computational Memory Model*
16 Davide Grossi (UU) *Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems*
17 Theodore Charitos (UU) *Reasoning with Dynamic Networks in Practice*
18 Bart Orriens (UvT) *On the development an management of adaptive business collaborations*
19 David Levy (UM) *Intimate relationships with artificial partners*
20 Slinger Jansen (UU) *Customer Configuration Updating in a Software Supply Network*
21 Karianne Vermaas (UU) *Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005*
22 Zlatko Zlatev (UT) *Goal-oriented design of value and process models from patterns*
23 Peter Barna (TUe) *Specification of Application Logic in Web Information Systems*

24 Georgina Ramírez Camps (CWI) *Structural Features in XML Retrieval*

25 Joost Schalken (VUA) *Empirical Investigations in Software Process Improvement*

2008

1 Katalin Boer-Sorbán (EUR) *Agent-Based Simulation of Financial Markets: A modular, continuous-time approach*

2 Alexei Sharpanskykh (VUA) *On Computer-Aided Methods for Modeling and Analysis of Organizations*

3 Vera Hollink (UvA) *Optimizing hierarchical menus: a usage-based approach*

4 Ander de Keijzer (UT) *Management of Uncertain Data: towards unattended integration*

5 Bela Mutschler (UT) *Modeling and simulating causal dependencies on process-aware information systems from a cost perspective*

6 Arjen Hommersom (RUN) *On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective*

7 Peter van Rosmalen (OU) *Supporting the tutor in the design and support of adaptive e-learning*

8 Janneke Bolt (UU) *Bayesian Networks: Aspects of Approximate Inference*

9 Christof van Nimwegen (UU) *The paradox of the guided user: assistance can be counter-effective*

10 Wauter Bosma (UT) *Discourse oriented summarization*

11 Vera Kartseva (VUA) *Designing Controls for Network Organizations: A Value-Based Approach*

12 Jozsef Farkas (RUN) *A Semiotically Oriented Cognitive Model of Knowledge Representation*

13 Caterina Carraciolo (UvA) *Topic Driven Access to Scientific Handbooks*

14 Arthur van Bunningen (UT) *Context-Aware Querying: Better Answers with Less Effort*

15 Martijn van Otterlo (UT) *The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains*

16 Henriette van Vugt (VUA) *Embodied agents from a user's perspective*

17 Martin Op 't Land (TUD) *Applying Architecture and Ontology to the Splitting and Allying of Enterprises*

18 Guido de Croon (UM) *Adaptive Active Vision*

19 Henning Rode (UT) *From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search*

20 Rex Arendsen (UvA) *Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven*

21 Krisztian Balog (UvA) *People Search in the Enterprise*

22 Henk Koning (UU) *Communication of IT-Architecture*

23 Stefan Visscher (UU) *Bayesian network models for the management of ventilator-associated pneumonia*

24 Zharko Aleksovski (VUA) *Using background knowledge in ontology matching*

25 Geert Jonker (UU) *Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency*

26 Marijn Huijbregts (UT) *Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled*

27 Hubert Vogten (OU) *Design and Implementation Strategies for IMS Learning Design*

28 Ildiko Flesch (RUN) *On the Use of Independence Relations in Bayesian Networks*

29 Dennis Reidsma (UT) *Annotations and Subjective Machines: Of Annotators, Embodied Agents, Users, and Other Humans*

30 Wouter van Atteveldt (VUA) *Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content*

31 Loes Braun (UM) *Pro-Active Medical Information Retrieval*

32 Trung H. Bui (UT) *Toward Affective Dialogue Management using Partially Observable Markov Decision Processes*

33 Frank Terpstra (UvA) *Scientific Workflow Design: theoretical and practical issues*

34 Jeroen de Knijf (UU) *Studies in Frequent Tree Mining*

35 Ben Torben Nielsen (UvT) *Dendritic morphologies: function shapes structure*

2009

1 Rasa Jurgelenaite (RUN) *Symmetric Causal Independence Models*

2 Willem Robert van Hage (VUA) *Evaluating Ontology-Alignment Techniques*

3 Hans Stol (UvT) *A Framework for Evidence-based Policy Making Using IT*

4 Josephine Nabukenya (RUN) *Improving the Quality of Organisational Policy Making using Collaboration Engineering*
5 Sietse Overbeek (RUN) *Bridging Supply and Demand for Knowledge Intensive Tasks: Based on Knowledge, Cognition, and Quality*
6 Muhammad Subianto (UU) *Understanding Classification*
7 Ronald Poppe (UT) *Discriminative Vision-Based Recovery and Recognition of Human Motion*
8 Volker Nannen (VUA) *Evolutionary Agent-Based Policy Analysis in Dynamic Environments*
9 Benjamin Kanagwa (RUN) *Design, Discovery and Construction of Service-oriented Systems*
10 Jan Wielemaker (UvA) *Logic programming for knowledge-intensive interactive applications*
11 Alexander Boer (UvA) *Legal Theory, Sources of Law & the Semantic Web*
12 Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin) *Operating Guidelines for Services*
13 Steven de Jong (UM) *Fairness in Multi-Agent Systems*
14 Maksym Korotkiy (VUA) *From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)*
15 Rinke Hoekstra (UvA) *Ontology Representation: Design Patterns and Ontologies that Make Sense*
16 Fritz Reul (UvT) *New Architectures in Computer Chess*
17 Laurens van der Maaten (UvT) *Feature Extraction from Visual Data*
18 Fabian Groffen (CWI) *Armada, An Evolving Database System*
19 Valentin Robu (CWI) *Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets*
20 Bob van der Vecht (UU) *Adjustable Autonomy: Controling Influences on Decision Making*
21 Stijn Vanderlooy (UM) *Ranking and Reliable Classification*
22 Pavel Serdyukov (UT) *Search For Expertise: Going beyond direct evidence*
23 Peter Hofgesang (VUA) *Modelling Web Usage in a Changing Environment*
24 Annerieke Heuvelink (VUA) *Cognitive Models for Training Simulations*
25 Alex van Ballegooij (CWI) *RAM: Array Database Management through Relational Mapping*
26 Fernando Koch (UU) *An Agent-Based Model for the Development of Intelligent Mobile Services*
27 Christian Glahn (OU) *Contextual Support of social Engagement and Reflection on the Web*
28 Sander Evers (UT) *Sensor Data Management with Probabilistic Models*
29 Stanislav Pokraev (UT) *Model-Driven Semantic Integration of Service-Oriented Applications*
30 Marcin Zukowski (CWI) *Balancing vectorized query execution with bandwidth-optimized storage*
31 Sofiya Katrenko (UvA) *A Closer Look at Learning Relations from Text*
32 Rik Farenhorst (VUA) *Architectural Knowledge Management: Supporting Architects and Auditors*
33 Khiet Truong (UT) *How Does Real Affect Affect Affect Recognition In Speech?*
34 Inge van de Weerd (UU) *Advancing in Software Product Management: An Incremental Method Engineering Approach*
35 Wouter Koelewijn (UL) *Privacy en Politiegegevens: Over geautomatiseerde normatieve informatie-uitwisseling*
36 Marco Kalz (OUN) *Placement Support for Learners in Learning Networks*
37 Hendrik Drachsler (OUN) *Navigation Support for Learners in Informal Learning Networks*
38 Riina Vuorikari (OU) *Tags and self-organisation: a metadata ecology for learning resources in a multilingual context*
39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin) *Service Substitution: A Behavioral Approach Based on Petri Nets*
40 Stephan Raaijmakers (UvT) *Multinomial Language Learning: Investigations into the Geometry of Language*
41 Igor Berezhnyy (UvT) *Digital Analysis of Paintings*
42 Toine Bogers (UvT) *Recommender Systems for Social Bookmarking*
43 Virginia Nunes Leal Franqueira (UT) *Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients*
44 Roberto Santana Tapia (UT) *Assessing Business-IT Alignment in Networked Organizations*
45 Jilles Vreeken (UU) *Making Pattern Mining Useful*
46 Loredana Afanasiev (UvA) *Querying XML: Benchmarks and Recursion*

2010

1 Matthijs van Leeuwen (UU) *Patterns that Matter*
2 Ingo Wassink (UT) *Work flows in Life Science*
3 Joost Geurts (CWI) *A Document Engineering Model and Processing Framework for Multimedia documents*
4 Olga Kulyk (UT) *Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments*
5 Claudia Hauff (UT) *Predicting the Effectiveness of Queries and Retrieval Systems*
6 Sander Bakkes (UvT) *Rapid Adaptation of Video Game AI*
7 Wim Fikkert (UT) *Gesture interaction at a Distance*
8 Krzysztof Siewicz (UL) *Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments*
9 Hugo Kielman (UL) *A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging*
10 Rebecca Ong (UL) *Mobile Communication and Protection of Children*
11 Adriaan Ter Mors (TUD) *The world according to MARP: Multi-Agent Route Planning*
12 Susan van den Braak (UU) *Sensemaking software for crime analysis*
13 Gianluigi Folino (RUN) *High Performance Data Mining using Bio-inspired techniques*
14 Sander van Splunter (VUA) *Automated Web Service Reconfiguration*
15 Lianne Bodenstaff (UT) *Managing Dependency Relations in Inter-Organizational Models*
16 Sicco Verwer (TUD) *Efficient Identification of Timed Automata, theory and practice*
17 Spyros Kotoulas (VUA) *Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications*
18 Charlotte Gerritsen (VUA) *Caught in the Act: Investigating Crime by Agent-Based Simulation*
19 Henriette Cramer (UvA) *People's Responses to Autonomous and Adaptive Systems*
20 Ivo Swartjes (UT) *Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative*
21 Harold van Heerde (UT) *Privacy-aware data management by means of data degradation*
22 Michiel Hildebrand (CWI) *End-user Support for Access to Heterogeneous Linked Data*
23 Bas Steunebrink (UU) *The Logical Structure of Emotions*
24 Zulfiqar Ali Memon (VUA) *Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective*
25 Ying Zhang (CWI) *XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines*
26 Marten Voulon (UL) *Automatisch contracteren*
27 Arne Koopman (UU) *Characteristic Relational Patterns*
28 Stratos Idreos (CWI) *Database Cracking: Towards Auto-tuning Database Kernels*
29 Marieke van Erp (UvT) *Accessing Natural History: Discoveries in data cleaning, structuring, and retrieval*
30 Victor de Boer (UvA) *Ontology Enrichment from Heterogeneous Sources on the Web*
31 Marcel Hiel (UvT) *An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems*
32 Robin Aly (UT) *Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval*
33 Teduh Dirgahayu (UT) *Interaction Design in Service Compositions*
34 Dolf Trieschnigg (UT) *Proof of Concept: Concept-based Biomedical Information Retrieval*
35 Jose Janssen (OU) *Paving the Way for Lifelong Learning: Facilitating competence development through a learning path specification*
36 Niels Lohmann (TUe) *Correctness of services and their composition*
37 Dirk Fahland (TUe) *From Scenarios to components*
38 Ghazanfar Farooq Siddiqui (VUA) *Integrative modeling of emotions in virtual agents*
39 Mark van Assem (VUA) *Converting and Integrating Vocabularies for the Semantic Web*
40 Guillaume Chaslot (UM) *Monte-Carlo Tree Search*
41 Sybren de Kinderen (VUA) *Needs-driven service bundling in a multi-supplier setting: the computational e3-service approach*
42 Peter van Kranenburg (UU) *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*
43 Pieter Bellekens (TUe) *An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain*
44 Vasilios Andrikopoulos (UvT) *A theory and model for the evolution of software services*

45 Vincent Pijpers (VUA) *e3alignment: Exploring Inter-Organizational Business-ICT Alignment*
46 Chen Li (UT) *Mining Process Model Variants: Challenges, Techniques, Examples*
47 Jahn-Takeshi Saito (UM) *Solving difficult game positions*
48 Bouke Huurnink (UvA) *Search in Audiovisual Broadcast Archives*
49 Alia Khairia Amin (CWI) *Understanding and supporting information seeking tasks in multiple sources*
50 Peter-Paul van Maanen (VUA) *Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention*
51 Edgar Meij (UvA) *Combining Concepts and Language Models for Information Access*

2011

1 Botond Cseke (RUN) *Variational Algorithms for Bayesian Inference in Latent Gaussian Models*
2 Nick Tinnemeier (UU) *Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language*
3 Jan Martijn van der Werf (TUe) *Compositional Design and Verification of Component-Based Information Systems*
4 Hado van Hasselt (UU) *Insights in Reinforcement Learning: Formal analysis and empirical evaluation of temporal-difference*
5 Base van der Raadt (VUA) *Enterprise Architecture Coming of Age: Increasing the Performance of an Emerging Discipline*
6 Yiwen Wang (TUe) *Semantically-Enhanced Recommendations in Cultural Heritage*
7 Yujia Cao (UT) *Multimodal Information Presentation for High Load Human Computer Interaction*
8 Nieske Vergunst (UU) *BDI-based Generation of Robust Task-Oriented Dialogues*
9 Tim de Jong (OU) *Contextualised Mobile Media for Learning*
10 Bart Bogaert (UvT) *Cloud Content Contention*
11 Dhaval Vyas (UT) *Designing for Awareness: An Experience-focused HCI Perspective*
12 Carmen Bratosin (TUe) *Grid Architecture for Distributed Process Mining*
13 Xiaoyu Mao (UvT) *Airport under Control. Multiagent Scheduling for Airport Ground Handling*
14 Milan Lovric (EUR) *Behavioral Finance and Agent-Based Artificial Markets*
15 Marijn Koolen (UvA) *The Meaning of Structure: the Value of Link Evidence for Information Retrieval*
16 Maarten Schadd (UM) *Selective Search in Games of Different Complexity*
17 Jiyin He (UvA) *Exploring Topic Structure: Coherence, Diversity and Relatedness*
18 Mark Ponsen (UM) *Strategic Decision-Making in complex games*
19 Ellen Rusman (OU) *The Mind ' s Eye on Personal Profiles*
20 Qing Gu (VUA) *Guiding service-oriented software engineering: A view-based approach*
21 Linda Terlouw (TUD) *Modularization and Specification of Service-Oriented Systems*
22 Junte Zhang (UvA) *System Evaluation of Archival Description and Access*
23 Wouter Weerkamp (UvA) *Finding People and their Utterances in Social Media*
24 Herwin van Welbergen (UT) *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*
25 Syed Waqar ul Qounain Jaffry (VUA) *Analysis and Validation of Models for Trust Dynamics*
26 Matthijs Aart Pontier (VUA) *Virtual Agents for Human Communication: Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*
27 Aniel Bhulai (VUA) *Dynamic website optimization through autonomous management of design patterns*
28 Rianne Kaptein (UvA) *Effective Focused Retrieval by Exploiting Query Context and Document Structure*
29 Faisal Kamiran (TUe) *Discrimination-aware Classification*
30 Egon van den Broek (UT) *Affective Signal Processing (ASP): Unraveling the mystery of emotions*
31 Ludo Waltman (EUR) *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*
32 Nees-Jan van Eck (EUR) *Methodological Advances in Bibliometric Mapping of Science*
33 Tom van der Weide (UU) *Arguing to Motivate Decisions*
34 Paolo Turrini (UU) *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*
35 Maaike Harbers (UU) *Explaining Agent Behavior in Virtual Training*

36 Erik van der Spek (UU) *Experiments in serious game design: a cognitive approach*

37 Adriana Burlutiu (RUN) *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference*

38 Nyree Lemmens (UM) *Bee-inspired Distributed Optimization*

39 Joost Westra (UU) *Organizing Adaptation using Agents in Serious Games*

40 Viktor Clerc (VUA) *Architectural Knowledge Management in Global Software Development*

41 Luan Ibraimi (UT) *Cryptographically Enforced Distributed Data Access Control*

42 Michal Sindlar (UU) *Explaining Behavior through Mental State Attribution*

43 Henk van der Schuur (UU) *Process Improvement through Software Operation Knowledge*

44 Boris Reuderink (UT) *Robust Brain-Computer Interfaces*

45 Herman Stehouwer (UvT) *Statistical Language Models for Alternative Sequence Selection*

46 Beibei Hu (TUD) *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work*

47 Azizi Bin Ab Aziz (VUA) *Exploring Computational Models for Intelligent Support of Persons with Depression*

48 Mark Ter Maat (UT) *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*

49 Andreea Niculescu (UT) *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality*

2012

1 Terry Kakeeto (UvT) *Relationship Marketing for SMEs in Uganda*

2 Muhammad Umair (VUA) *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models*

3 Adam Vanya (VUA) *Supporting Architecture Evolution by Mining Software Repositories*

4 Jurriaan Souer (UU) *Development of Content Management System-based Web Applications*

5 Marijn Plomp (UU) *Maturing Interorganisational Information Systems*

6 Wolfgang Reinhardt (OU) *Awareness Support for Knowledge Workers in Research Networks*

7 Rianne van Lambalgen (VUA) *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions*

8 Gerben de Vries (UvA) *Kernel Methods for Vessel Trajectories*

9 Ricardo Neisse (UT) *Trust and Privacy Management Support for Context-Aware Service Platforms*

10 David Smits (TUe) *Towards a Generic Distributed Adaptive Hypermedia Environment*

11 J. C. B. Rantham Prabhakara (TUe) *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*

12 Kees van der Sluijs (TUe) *Model Driven Design and Data Integration in Semantic Web Information Systems*

13 Suleman Shahid (UvT) *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions*

14 Evgeny Knutov (TUe) *Generic Adaptation Framework for Unifying Adaptive Web-based Systems*

15 Natalie van der Wal (VUA) *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes*

16 Fiemke Both (VUA) *Helping people by understanding them: Ambient Agents supporting task execution and depression treatment*

17 Amal Elgammal (UvT) *Towards a Comprehensive Framework for Business Process Compliance*

18 Eltjo Poort (VUA) *Improving Solution Architecting Practices*

19 Helen Schonenberg (TUe) *What's Next? Operational Support for Business Process Execution*

20 Ali Bahramisharif (RUN) *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing*

21 Roberto Cornacchia (TUD) *Querying Sparse Matrices for Information Retrieval*

22 Thijs Vis (UvT) *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?*

23 Christian Muehl (UT) *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction*

24 Laurens van der Werff (UT) *Evaluation of Noisy Transcripts for Spoken Document Retrieval*

25 Silja Eckartz (UT) *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application*

26 Emile de Maat (UvA) *Making Sense of Legal Text*

27 Hayrettin Gurkok (UT) *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games*

28 Nancy Pascall (UvT) *Engendering Technology Empowering Women*
29 Almer Tigelaar (UT) *Peer-to-Peer Information Retrieval*
30 Alina Pommeranz (TUD) *Designing Human-Centered Systems for Reflective Decision Making*
31 Emily Bagarukayo (RUN) *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure*
32 Wietske Visser (TUD) *Qualitative multi-criteria preference representation and reasoning*
33 Rory Sie (OUN) *Coalitions in Cooperation Networks (COCOON)*
34 Pavol Jancura (RUN) *Evolutionary analysis in PPI networks and applications*
35 Evert Haasdijk (VUA) *Never Too Old To Learn: On-line Evolution of Controllers in Swarm- and Modular Robotics*
36 Denis Ssebugwawo (RUN) *Analysis and Evaluation of Collaborative Modeling Processes*
37 Agnes Nakakawa (RUN) *A Collaboration Process for Enterprise Architecture Creation*
38 Selmar Smit (VUA) *Parameter Tuning and Scientific Testing in Evolutionary Algorithms*
39 Hassan Fatemi (UT) *Risk-aware design of value and coordination networks*
40 Agus Gunawan (UvT) *Information Access for SMEs in Indonesia*
41 Sebastian Kelle (OU) *Game Design Patterns for Learning*
42 Dominique Verpoorten (OU) *Reflection Amplifiers in self-regulated Learning*
43 Anna Tordai (VUA) *On Combining Alignment Techniques*
44 Benedikt Kratz (UvT) *A Model and Language for Business-aware Transactions*
45 Simon Carter (UvA) *Exploration and Exploitation of Multilingual Data for Statistical Machine Translation*
46 Manos Tsagkias (UvA) *Mining Social Media: Tracking Content and Predicting Behavior*
47 Jorn Bakker (TUe) *Handling Abrupt Changes in Evolving Time-series Data*
48 Michael Kaisers (UM) *Learning against Learning: Evolutionary dynamics of reinforcement learning algorithms in strategic interactions*
49 Steven van Kervel (TUD) *Ontologogy driven Enterprise Information Systems Engineering*
50 Jeroen de Jong (TUD) *Heuristics in Dynamic Sceduling: a practical framework with a case study in elevator dispatching*

2013

1 Viorel Milea (EUR) *News Analytics for Financial Decision Support*
2 Erietta Liarou (CWI) *MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing*
3 Szymon Klarman (VUA) *Reasoning with Contexts in Description Logics*
4 Chetan Yadati (TUD) *Coordinating autonomous planning and scheduling*
5 Dulce Pumareja (UT) *Groupware Requirements Evolutions Patterns*
6 Romulo Goncalves (CWI) *The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience*
7 Giel van Lankveld (UvT) *Quantifying Individual Player Differences*
8 Robbert-Jan Merk (VUA) *Making enemies: cognitive modeling for opponent agents in fighter pilot simulators*
9 Fabio Gori (RUN) *Metagenomic Data Analysis: Computational Methods and Applications*
10 Jeewanie Jayasinghe Arachchige (UvT) *A Unified Modeling Framework for Service Design*
11 Evangelos Pournaras (TUD) *Multi-level Reconfigurable Self-organization in Overlay Services*
12 Marian Razavian (VUA) *Knowledge-driven Migration to Services*
13 Mohammad Safiri (UT) *Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly*
14 Jafar Tanha (UvA) *Ensemble Approaches to Semi-Supervised Learning Learning*
15 Daniel Hennes (UM) *Multiagent Learning: Dynamic Games and Applications*
16 Eric Kok (UU) *Exploring the practical benefits of argumentation in multi-agent deliberation*
17 Koen Kok (VUA) *The PowerMatcher: Smart Coordination for the Smart Electricity Grid*
18 Jeroen Janssens (UvT) *Outlier Selection and One-Class Classification*
19 Renze Steenhuizen (TUD) *Coordinated Multi-Agent Planning and Scheduling*
20 Katja Hofmann (UvA) *Fast and Reliable Online Learning to Rank for Information Retrieval*
21 Sander Wubben (UvT) *Text-to-text generation by monolingual machine translation*
22 Tom Claassen (RUN) *Causal Discovery and Logic*

23 Patricio de Alencar Silva (UvT) *Value Activity Monitoring*

24 Haitham Bou Ammar (UM) *Automated Transfer in Reinforcement Learning*

25 Agnieszka Anna Latoszek-Berendsen (UM) *Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System*

26 Alireza Zarghami (UT) *Architectural Support for Dynamic Homecare Service Provisioning*

27 Mohammad Huq (UT) *Inference-based Framework Managing Data Provenance*

28 Frans van der Sluis (UT) *When Complexity becomes Interesting: An Inquiry into the Information eXperience*

29 Iwan de Kok (UT) *Listening Heads*

30 Joyce Nakatumba (TUe) *Resource-Aware Business Process Management: Analysis and Support*

31 Dinh Khoa Nguyen (UvT) *Blueprint Model and Language for Engineering Cloud Applications*

32 Kamakshi Rajagopal (OUN) *Networking For Learning: The role of Networking in a Lifelong Learner's Professional Development*

33 Qi Gao (TUD) *User Modeling and Personalization in the Microblogging Sphere*

34 Kien Tjin-Kam-Jet (UT) *Distributed Deep Web Search*

35 Abdallah El Ali (UvA) *Minimal Mobile Human Computer Interaction*

36 Than Lam Hoang (TUe) *Pattern Mining in Data Streams*

37 Dirk Börner (OUN) *Ambient Learning Displays*

38 Eelco den Heijer (VUA) *Autonomous Evolutionary Art*

39 Joop de Jong (TUD) *A Method for Enterprise Ontology based Design of Enterprise Information Systems*

40 Pim Nijssen (UM) *Monte-Carlo Tree Search for Multi-Player Games*

41 Jochem Liem (UvA) *Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning*

42 Léon Planken (TUD) *Algorithms for Simple Temporal Reasoning*

43 Marc Bron (UvA) *Exploration and Contextualization through Interaction and Concepts*

2014

1 Nicola Barile (UU) *Studies in Learning Monotone Models from Data*

2 Fiona Tuliyano (RUN) *Combining System Dynamics with a Domain Modeling Method*

3 Sergio Raul Duarte Torres (UT) *Information Retrieval for Children: Search Behavior and Solutions*

4 Hanna Jochmann-Mannak (UT) *Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation*

5 Jurriaan van Reijsen (UU) *Knowledge Perspectives on Advancing Dynamic Capability*

6 Damian Tamburri (VUA) *Supporting Networked Software Development*

7 Arya Adriansyah (TUe) *Aligning Observed and Modeled Behavior*

8 Samur Araujo (TUD) *Data Integration over Distributed and Heterogeneous Data Endpoints*

9 Philip Jackson (UvT) *Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language*

10 Ivan Salvador Razo Zapata (VUA) *Service Value Networks*

11 Janneke van der Zwaan (TUD) *An Empathic Virtual Buddy for Social Support*

12 Willem van Willigen (VUA) *Look Ma, No Hands: Aspects of Autonomous Vehicle Control*

13 Arlette van Wissen (VUA) *Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains*

14 Yangyang Shi (TUD) *Language Models With Meta-information*

15 Natalya Mogles (VUA) *Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare*

16 Krystyna Milian (VUA) *Supporting trial recruitment and design by automatically interpreting eligibility criteria*

17 Kathrin Dentler (VUA) *Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability*

18 Mattijs Ghijsen (UvA) *Methods and Models for the Design and Study of Dynamic Agent Organizations*

19 Vinicius Ramos (TUe) *Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support*

20 Mena Habib (UT) *Named Entity Extraction and Disambiguation for Informal Text: The Missing Link*

21 Kassidy Clark (TUD) *Negotiation and Monitoring in Open Environments*

22 Marieke Peeters (UU) *Personalized Educational Games: Developing agent-supported scenario-based training*

23 Eleftherios Sidirourgos (UvA/CWI) *Space Efficient Indexes for the Big Data Era*

24 Davide Ceolin (VUA) *Trusting Semi-structured Web Data*

25 Martijn Lappenschaar (RUN) *New network models for the analysis of disease interaction*

26 Tim Baarslag (TUD) *What to Bid and When to Stop*

27 Rui Jorge Almeida (EUR) *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty*

28 Anna Chmielowiec (VUA) *Decentralized k-Clique Matching*

29 Jaap Kabbedijk (UU) *Variability in Multi-Tenant Enterprise Software*

30 Peter de Cock (UvT) *Anticipating Criminal Behaviour*

31 Leo van Moergestel (UU) *Agent Technology in Agile Multiparallel Manufacturing and Product Support*

32 Naser Ayat (UvA) *On Entity Resolution in Probabilistic Data*

33 Tesfa Tegegne (RUN) *Service Discovery in eHealth*

34 Christina Manteli (VUA) *The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems*

35 Joost van Ooijen (UU) *Cognitive Agents in Virtual Worlds: A Middleware Design Approach*

36 Joos Buijs (TUe) *Flexible Evolutionary Algorithms for Mining Structured Process Models*

37 Maral Dadvar (UT) *Experts and Machines United Against Cyberbullying*

38 Danny Plass-Oude Bos (UT) *Making brain-computer interfaces better: improving usability through post-processing*

39 Jasmina Maric (UvT) *Web Communities, Immigration, and Social Capital*

40 Walter Omona (RUN) *A Framework for Knowledge Management Using ICT in Higher Education*

41 Frederic Hogenboom (EUR) *Automated Detection of Financial Events in News Text*

42 Carsten Eijckhof (CWI/TUD) *Contextual Multidimensional Relevance Models*

43 Kevin Vlaanderen (UU) *Supporting Process Improvement using Method Increments*

44 Paulien Meesters (UvT) *Intelligent Blauw: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden*

45 Birgit Schmitz (OUN) *Mobile Games for Learning: A Pattern-Based Approach*

46 Ke Tao (TUD) *Social Web Data Analytics: Relevance, Redundancy, Diversity*

47 Shangsong Liang (UvA) *Fusion and Diversification in Information Retrieval*

2015

1 Niels Netten (UvA) *Machine Learning for Relevance of Information in Crisis Response*

2 Faiza Bukhsh (UvT) *Smart auditing: Innovative Compliance Checking in Customs Controls*

3 Twan van Laarhoven (RUN) *Machine learning for network data*

4 Howard Spoelstra (OUN) *Collaborations in Open Learning Environments*

5 Christoph Bösch (UT) *Cryptographically Enforced Search Pattern Hiding*

6 Farideh Heidari (TUD) *Business Process Quality Computation: Computing Non-Functional Requirements to Improve Business Processes*

7 Maria-Hendrike Peetz (UvA) *Time-Aware Online Reputation Analysis*

8 Jie Jiang (TUD) *Organizational Compliance: An agent-based model for designing and evaluating organizational interactions*

9 Randy Klaassen (UT) *HCI Perspectives on Behavior Change Support Systems*

10 Henry Hermans (OUN) *OpenU: design of an integrated system to support lifelong learning*

11 Yongming Luo (TUe) *Designing algorithms for big graph datasets: A study of computing bisimulation and joins*

12 Julie M. Birkholz (VUA) *Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks*

13 Giuseppe Procaccianti (VUA) *Energy-Efficient Software*

14 Bart van Straalen (UT) *A cognitive approach to modeling bad news conversations*

15 Klaas Andries de Graaf (VUA) *Ontology-based Software Architecture Documentation*

16 Changyun Wei (UT) *Cognitive Coordination for Cooperative Multi-Robot Teamwork*

17 André van Cleeff (UT) *Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs*

18 Holger Pirk (CWI) *Waste Not, Want Not!: Managing Relational Data in Asymmetric Memories*

19 Bernardo Tabuenca (OUN) *Ubiquitous Technology for Lifelong Learners*
20 Loïs Vanhée (UU) *Using Culture and Values to Support Flexible Coordination*
21 Sibren Fetter (OUN) *Using Peer-Support to Expand and Stabilize Online Learning*
22 Zhemin Zhu (UT) *Co-occurrence Rate Networks*
23 Luit Gazendam (VUA) *Cataloguer Support in Cultural Heritage*
24 Richard Berendsen (UvA) *Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation*
25 Steven Woudenberg (UU) *Bayesian Tools for Early Disease Detection*
26 Alexander Hogenboom (EUR) *Sentiment Analysis of Text Guided by Semantics and Structure*
27 Sándor Héman (CWI) *Updating compressed column-stores*
28 Janet Bagorogoza (TiU) *Knowledge Management and High Performance: The Uganda Financial Institutions Model for HPO*
29 Hendrik Baier (UM) *Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains*
30 Kiavash Bahreini (OUN) *Real-time Multimodal Emotion Recognition in E-Learning*
31 Yakup Koç (TUD) *On Robustness of Power Grids*
32 Jerome Gard (UL) *Corporate Venture Management in SMEs*
33 Frederik Schadd (UM) *Ontology Mapping with Auxiliary Resources*
34 Victor de Graaff (UT) *Geosocial Recommender Systems*
35 Junchao Xu (TUD) *Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction*

2016

1 Syed Saiden Abbas (RUN) *Recognition of Shapes by Humans and Machines*
2 Michiel Christiaan Meulendijk (UU) *Optimizing medication reviews through decision support: prescribing a better pill to swallow*
3 Maya Sappelli (RUN) *Knowledge Work in Context: User Centered Knowledge Worker Support*
4 Laurens Rietveld (VUA) *Publishing and Consuming Linked Data*
5 Evgeny Sherkhonov (UvA) *Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers*
6 Michel Wilson (TUD) *Robust scheduling in an uncertain environment*
7 Jeroen de Man (VUA) *Measuring and modeling negative emotions for virtual training*
8 Matje van de Camp (TiU) *A Link to the Past: Constructing Historical Social Networks from Unstructured Data*
9 Archana Nottamkandath (VUA) *Trusting Crowdsourced Information on Cultural Artefacts*
10 George Karafotias (VUA) *Parameter Control for Evolutionary Algorithms*
11 Anne Schuth (UvA) *Search Engines that Learn from Their Users*
12 Max Knobbout (UU) *Logics for Modelling and Verifying Normative Multi-Agent Systems*
13 Nana Baah Gyan (VU) *The Web, Speech Technologies and Rural Development in West Africa: An ICT4D Approach*
14 Ravi Khadka (UU) *Revisiting Legacy Software System Modernization*
15 Steffen Michels (RUN) *Hybrid Probabilistic Logics: Theoretical Aspects, Algorithms and Experiments*
16 Guangliang Li (UvA) *Socially Intelligent Autonomous Agents that Learn from Human Reward*
17 Berend Weel (VUA) *Towards Embodied Evolution of Robot Organisms*
18 Albert Meroño Peñuela (VUA) *Refining Statistical Data on the Web*
19 Julia Efremova (TUe) *Mining Social Structures from Genealogical Data*
20 Daan Odijk (UvA) *Context & Semantics in News & Web Search*
21 Alejandro Moreno Célleri (UT) *From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground*
22 Grace Lewis (VU) *Software Architecture Strategies for Cyber-Foraging Systems*
23 Fei Cai (UVA) *Query Auto Completion in Information Retrieval*
24 Brend Wanders (UT) *Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach*
25 Julia Kiseleva (TU/e) *Using Contextual Information to Understand Searching and Browsing Behavior*
26 Dilhan Thilakarathne (VU) *In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains*
27 Wen Li (TUD) *Understanding Geo-spatial Information on Social Media*

Curriculum Vitae

Mingxin Zhang was born in NanTong, China on 13th July, 1986. Mingxin started his undergraduate study in the school of Mechanical Engineering and Automation at National University of Defense Technology (NUDT) in China from 2004. In 2008, he started his master study in System Simulation with a priority recommendation (no examination required), under the supervision of Prof. Kedi Huang. After two years of study and research experience, he received a master degree in engineering and started his PhD in the same area since February 2011 at NUDT. In September 2011, Mingxin was awarded a PhD scholarship from the China Scholarship Council (CSC) and started his PhD study in large-scale agent-based social simulation at Delft University of Technology in The Netherlands, under the supervision of Prof. Alexander Verbraeck. During his PhD study, he worked closely with his former colleagues at NUDT on a case study of epidemic prediction and control in the city of Beijing. From September 2015, Mingxin participated in a research project in collaboration with Nissan Research Center (NRC) in USA.

Mingxin's main research interests are in agent-based modeling, social networks and environmental modeling with a focus on large-scale social systems, such as epidemic prediction and transportation analysis.

Publications

1. Mingxin Zhang, Alexander Verbraeck, Rongqing Meng, Bin Chen, Xiaogang Qiu. Modeling Spatial Contacts for Epidemic Prediction in a Large-scale Artificial City. Journal of Artificial Societies and Social Simulation (First Revision)
2. Mingxin Zhang, Rongqing Meng, Alexander Verbraeck. Including Public Transportation into a Large-scale Agent-based Model for Epidemic Predication and Control. 2015 Summer Simulation Multi-Conference (SummerSim'15). 2015.07
3. Mingxin Zhang, Alexander Verbraeck, Rongqing Meng, Xiaogang Qiu. Agent-based Modeling of Large-scale Complex Social Interactions. 2015 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation (PADS'15). 2015.06
4. Mingxin Zhang, Alexander Verbraeck. Large-scale Agent-based Modeling and Simulation. 2014 Winter Simulation Conference (PhD colloquium). 2014.12
5. Mingxin Zhang, Alexander Verbraeck. A Composable PRS-based Agent-based Meta-Model for Multi-agent Simulation Using the DEVS Framework. 2014 Spring Simulation Multi-Conference (SpringSim'14). 2014.04
6. Mingxin Zhang, Mamadou Seck, Alexander Verbraeck. A DEVS-based M&S Method for Large-scale Multi-agent Systems. 2013 Summer Simulation Multi-Conference (SummerSim'13). 2013.07
7. Mingxin Zhang. Constructing a cognitive agent model using DEVS framework for Multi-agent simulation. 15th European Agent Systems Summer School (Student Session). 2013.07