Chapter 12
Combining Crowd Sensing and Social Data Mining with Agent-Based Simulation Using Mobile Agents Towards Augmented Virtuality

Stefan Bosse and Uwe H. Engel

Abstract Augmented reality is well known for extending the real world by adding computer-generated perceptual information and overlaid sensory information. In contrast, simulation worlds are commonly closed and rely on artificial social behaviour and synthetic sensory information generated by the simulator program or using data collected off-line by surveys. Agent-based modelling used for investigation and evaluation of social interaction and networking relies on parameterisable models. Finding accurate and representative parameter settings can be a challenge. In this work, a new simulation paradigm is introduced, providing augmented virtuality by coupling crowd sensing and social data mining with simulation worlds in real-time by using mobile agents in an unified way. A simple social network analysis case-study based on the Sakoda social interaction model and mobile crowd sensing demonstrates the capabilities of the new hybrid simulation method and the impact of collected real-world data on social simulation.

Keywords Agent-based modelling · Agent-based simulation · Agents · Mobile crowd sensing · Agent platforms

Introduction

The key concept of this work is the consideration of humans as sensors and the seamless integration of real-world sensors in social simulation. This concept is highly interdisciplinary and is a merit of social and computer science if the human sensor data will be coupled with social interaction and networking models. Social interaction has

S. Bosse (✉)
Department Mathematics and Computer Science, University of Bremen, Bremen, Germany
e-mail: sbosse@uni-bremen.de

U. H. Engel
Department Social Science, Social Science Methods Centre, University of Bremen, Bremen, Germany

P. Ahrweiler and M. Neumann (eds.), *Advances in Social Simulation*,
Springer Proceedings in Complexity,
https://doi.org/10.1007/978-3-030-61503-1_12

a high impact on the control and evolution of complex social systems, which should be addressed in this work.

Agent-based methods are established for modelling and studying of complex dynamic systems and for implementing distributed intelligent systems, e.g., in traffic and transportation control (see [1, 2]). Therefore, agent-based methods can be divided into the following main classes and paradigm [3]:

1. Agent-based Modelling (ABM)—Modelling of complex dynamic systems by using the agent behaviour and interaction model ⇒ **Physical agents**
2. Agent-based Computing (ABC)—Distributed and parallel computing using mobile agents related to mobile software processes ⇒ **Computational agents**
3. Agent-based Simulation (ABS)—Simulation of agents or using agents for simulation
4. Hybrid Agent-based Computation, Modelling, and Simulation (ABX) ⇒ **Combining physical and computational agents**.

The fourth paradigm is the novelty introduced in this work with application to social mobility and network simulation.

ABS is suitable for studying complex social systems with respect to interaction between individual entities, manipulation of the world, spatial movement, and emergence effects of groups of entities. The main advantage is the bottom-up modelling approach composing large-scale complex systems by simple entity models. The main disadvantage of ABM is the (over-) simplified entity behaviour and simplification of the world the entities acting in. Commonly, simulation bases on synthetic data or data retrieved by field studies. Many simulations and models lacking of diversity that are existing in real world. Commonly, sensor and model data (parameters) used in simulations (virtual world) is retrieved from experiments or field studies (real world), shown in Fig. 12.1. But there is neither a feedback from the virtual to the real world nor an interaction of the real world with the virtual world. ABC is mainly used to implement adaptive and self-organising distributed computing. In this work, a different simulation approach combining ABM, ABC, and ABS methodologies (ABX), is used deploying the widely used programming language JavaScript. JavaScript is an easy to learn programming language that was used originally in WEB development, but is increasingly used as a generic programming language and for ABS [4].

The *NetLogo* simulator is an established agent-based simulation tool used in social and natural sciences [5]. NetLogo bases on an observer-centered global bottom-up model, i.e., there is one script representing an observer that controls the entire simulation and the implements the agent behaviour. The *NetLogo* programming language (that is domain-specific) consists basically of sets and set iterator statements performing computations. In this work, a different simulation approach combining ABM, ABC, and ABS methodologies, is used deploying the widely used programming language JavaScript. JavaScript is an easy to learn programming language that was used originally in WEB development, but is increasingly used as a generic programming language and for ABS [4, 6].

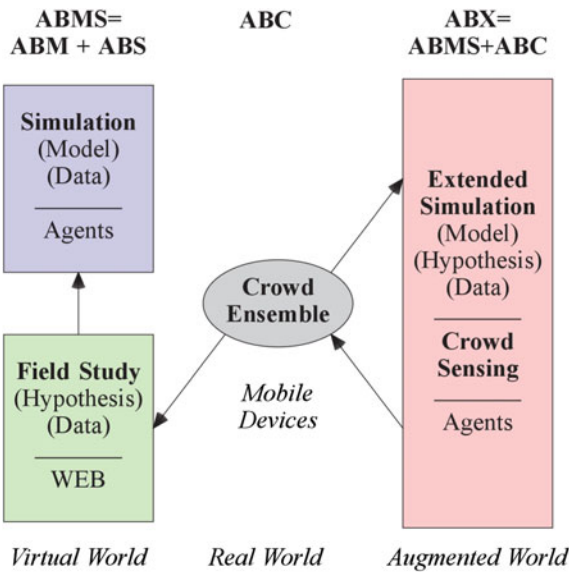

Fig. 12.1 (Left) Traditional ABMS with separate field studies (Right) New ABX combining modelling, simulation, and field studies via crowd sensing

Today administration and management of public services and infrastructure relies more and more on user data collected by many domestic and private devices including smart phones and Internet services. Simulation is increasingly used to study emergence effects of complex distributed systems. User data and user decision making has a large impact on public decision making processes, for example, plan-based traffic flow control. Furthermore, intelligent behaviour, i.e., cognitive, knowledge-based, adaptive, and self-organizing behaviour based on learning, emerges rapidly in today's machines and environments. Social science itself exerts influence on public opinion formation. We observe a rise of Artificial Intelligence (AI) in the shape of computational methods of analysis of data and meta data extracted, e.g., from the web. This statement alludes to the rise of "computational social science" (CSS), the accompanying shift from survey methods to computational methods and possible implications of this development: usage of different data, different methods, exposition to different threats to data quality (concerning sampling, selection effects, measurement effects, data analysis), and different rules of inference are likely to result in comparably different conclusions, public reports, and hence in different input to public opinion formation.

ABM and ABS relies on parameterisable behaviour models. The selection of appropriate and representative parameter sets are crucial for modelling real-world scenarios.

Mobile devices like smart phones are valuable sources for social data [7], either by participatory crowd sensing with explicit participation of users providing first class data (e.g., performing surveys or polls) or implicitly by opportunistic crowd sensing collecting secondary class data, i.e., traces of device sensor data delivering, e.g., actual position, ambient conditions, network connectivity, digital media interaction, and so on. Crowd sensing and Social Data Mining as a data source contribute more and more to investigations of digital traces in large-scale machine-human environments characterised by complex interactions and causalities between perception and action (decision making).

It is difficult to study such large-scale data collection, data mining, and their effect on societies and social interaction in field studies due to a lack of data. Agent-based modelling of socio-technical systems is well established [8], however commonly applied in an artificial world., i.e., a simulation is performed in virtual reality worlds only to derive and proof models under hard limitations. In this work, a new concept and framework for augmented virtual reality simulation is introduced, suitable, but not limited to, to investigate large-scale socio-technical systems. Mobile agents are used already successfully in-field crowd sensing [9]. In this work, mobile agents are used to combine in-field ubiquitous crowd sensing, e.g., performed by mobile devices, with simulation.

Crowd sensing (CWS) can be considered as part of field studies [9], either performed in a participatory or opportunistic way. A challenge in crowd sensing as well as classical field studies is user motivation in participation and incentive mechanisms. Opportunistic crowd sensing on mobile and ubiquitous devices requires self-organising and adaptation capabilities, which cannot be satisfied by classical and centralised WEB-based services. Mobile agents (ABC) as mobile software processes can overcome limitations occurring in centralised WEB infrastructures [9]. Mobile agents provide a seamless interface between real and virtual worlds (ABMS).

Agent-Based Crowd Sensing, Chat Bots, and Digital Twins

Classical field studies are commonly serviced by a central WEB server and WEB-based content management. These field studies address usually participative crowd sensing. Opportunistic crowd sensing can extend user classes and user numbers significantly. One promising approach supporting the human-as-sensor concept is the deployment of autonomous or semi-autonomous chat bot agents that can interact with humans on a wide range of devices, media, and social platforms. Mobile agents feature self-organising and self-adaptation. Commonly, simulation is performed with artificial agent models derived from theoretical considerations or experimental data. Augmented virtuality enables dynamic simulations with agents representing real humans (or crowds). By using crowd sensing it is possible to create digital twins of real humans based on a parameterisable behaviour and interaction model. The parameters of artificial humans in the simulation represented by agents are collected by sensor data, i.e., surveys optionally fusioned with physical sensors like GPS.

It is assumed that there is a parameterisable behaviour model M of individual entities, i.e., a model of social interaction, which is used to model digital twins representing individual real humans (P: Parameter set, S: Sensor set, A: Action set):

$$M(S, P) : S \times P \rightarrow A \tag{12.1}$$

The individual behaviour model will result in a specific system behaviour showing emergence effects (e.g., segregation and group formation).

Mobile chat bot agents fulfill two tasks in crowd sensing: (I) Sensing of data from the user and user devices and (II) Negotiation between the crowd sourcer and the user.

Extended Simulation Architecture

The software framework couples virtual and real worlds by using computational agents (chat bots) operating in virtual and real worlds and physical agents in simulation. The chat bot agents are the interface between humans and simulation.

The entire crowd sensing and simulation architecture consists of the following components:

1. Crowd sensing software (Mobile App and WEB Browser)
2. Unified agent processing platform based on JavaScript (JavaScript Agent Machine, *JAM*)
3. Agent-based simulation with Internet connectivity providing two different agent types:
   - **Physical agents** representing individual artificial humans or any other physical entity in the simulation world;
   - **Computational agents** representing mobile software, i.e., used for distributed data processing and digital communication, and implementing chat bots, available in the simulation and the Internet;
4. Chat dialogues, Chat bots and Mobile agents (computational);
5. Knowledge-based Question–Answer Systems.

The principle system architecture is shown in Fig. 12.2. Details can be found in [4]. *JAM* platform nodes exist in the simulator, in Internet relays used for the interconnectivity of CWS and simulation, and application programs for crowd sensing operating, e.g., on mobile devices. The simulation world consists of a set of virtual *JAM* platforms that are controlled by a physical *JAM* node. A physical *JAM* node can be connected to the Internet, virtual nodes can be connected with each other (simulating communication).

In the simulation model, there is no significant difference between physical and computational agents. The main difference is mobility (in the simulation world). Physical agents are bound to a virtual *JAM* node (*vJAM*), and the agent is mobile by its platform, whereas computational agents can migrate between platforms and between virtual and real worlds. Physical agents can access an extended programming interface (API) similar to the well-known *NetLogo* simulation model providing agent control like agent creation, movement or visual changes. An expressive *ask* operator is provided, too.

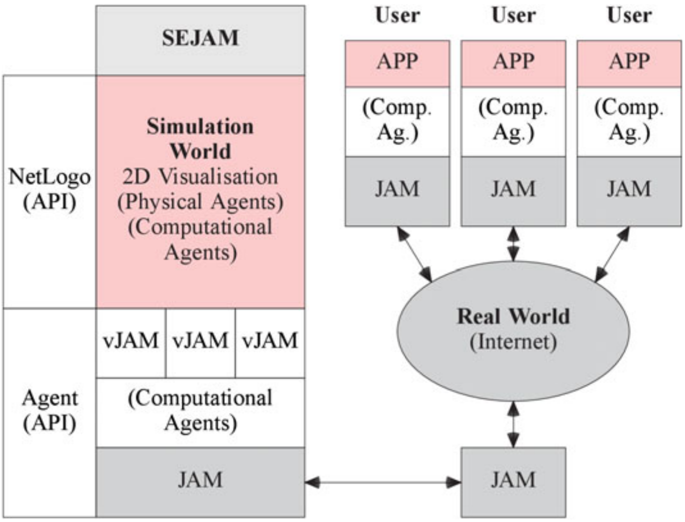

Fig. 12.2 Principle concept of closed-loop simulation for augmented virtuality: (Left) Simulation framework based on the JAM platform (Right) Mobile and non-mobile devices executing the JAM platform connected with the virtual simulation world (via the Internet)

Social Interaction and Segregation with the Sakoda Model

To demonstrate the augmented virtuality approach combining agent-based simulation with agent-based crowd sensing the *Sakoda* model [10] was chosen as a simple social interaction and behaviour model between groups of individual humans. It poses self-organising behaviour (emergence) and structures of social groups by segregation.

For the sake of simplicity, there is a two-dimensional grid world that consists of places at discrete locations (x, y). An artificial agent occupies one place of the grid. Maximal one agent can occupy a place. The agents can move on the grid and can change their living position.

It is assumed that there are two groups related to the classes a and b of individuals. The social interaction is characterised by different attitudes [10] of an individual between different and among same groups given by four parameters:

$$S_{ab} = (s_{aa}, s_{ab}, s_{ba}, s_{bb}), S_{abcd} = \begin{pmatrix} s_{aa} & s_{ab} & s_{ac} & s_{ad} \\ s_{ba} & s_{bb} & s_{bc} & s_{bd} \\ s_{ca} & s_{cb} & s_{cc} & s_{cd} \\ s_{da} & s_{db} & s_{dc} & s_{dd} \end{pmatrix} \quad (12.2)$$

The model is not limited to two groups of individuals. The S vector can be extended to four or more groups (or generalised) by a n-dimensional matrix. Here, the S vector or matrix is the model parameter set $M(P)$ that can be used to create diverse digital twins based on surveys and crowd sensing.

The world model consists of N places x_i. Each place can be occupied by none or one agent either of group α or β, expressed by the variable $x_i = \{0, -1, 1\}$, or generalised $x_i = \{0, 1, 2, 3, 4,...,n\}$ with n groups. The social expectation of an individual i at place x_i is given by:

$$f_i(x_i) = \sum_{k=1}^{N} J_{ik}\delta_s(x_i, x_k) \tag{12.3}$$

The parameter J_{ik} is a measure of the social distance (equal one for Moore neighbourhood with distance one), decreasing for longer distances. The parameter δ expresses the attitude to a neighbour place, given by (for the general case of n different groups):

$$\delta_s(x_i, x_k) = \begin{cases} s_{\alpha\beta}\ , & \text{if } x_i \neq 0 \text{ and } x_k \neq 0 \text{ with } \alpha = x_i,\ \beta = x_k \\ 0\ , & \text{otherwise} \end{cases} \tag{12.4}$$

An individual agent ag_i of group α or β is able to change its position by migrating from an actual place x_i to another place x_q if this place is not occupied ($x_q = 0$) and if $f_i(x_q) > f_i(x_i)$. The computation of the neighbouring social expectation f values is opportunistic, i.e., if f is computed for a neighbouring place, it is assumed that the agent occupies this neighbour place if the place is free, and the current original place x_i is omitted for this computation. Any other already occupied places are kept unchanged for the computation of a particular f value. From the set of neighbouring places and their particular social expectations for the specific agent the best place is chosen for migration (if there is a better place than the current with the above condition). In this work, spatial social distances in the range 1–30 place units are considered.

Originally, the entire world consists of individual agents interacting in the world based on one specific set of attitude parameters S. In this work, the model is generalised by assigning individual entities its own set S retrieved from real humans by crowd sensing, or at least different configurations of the S vector classifying social behaviour among the groups. Furthermore, the set of entities can be extended by humans and bots (intelligent machines) belonging to a group class, too.

Model Parameters and Crowd Sensing

Creation of virtual digital twins is the aim of the crowd sensing. The crowd sensing is performed with chat bot agents. One stationary agent is operating on a user device, e.g., a smart phone, and another mobile agent is responsible to perform a survey (either participatory with a former negotiation or opportunistic ad-hoc).

The results of the survey, a set of questions, are used to derive the following simulation model parameters:

```
parameters = {
  group : string "a"|"b",
  social-distance: number [1-100],
  social-attitudes : [saa,sab,sba,sbb] ,
  mobility : number [0-1],
  position : {x:number,y:number}
}
```

The *group* parameter sorts the user in one of two classes *a*/*b*, the *social-distance* parameter is an estimation of the social interaction distance, the *social-attitudes* parameter is the S vector, but limited to a sub-set of all possible S vector combinations (discussed below), and the *mobility* parameter is a probability to migrate from one place to another. The position (in cartesian coordinates) is derived from the living centre of the user (global position data, GPS) and mapped on the simulation world (x, y). The S vector parameter determines the spatial social organisation structure. Typical examples of the S vectors with relation to social behaviour are [10]:

- (1, −1, −1, 1): Typical segregation with strong and isolated group clusters
- (0, −1, −1, 0): Mutual suspicion
- (1, −1, 1, −1): Social climbers
- (1, 1, −1, −1): Social workers
- (1, 1, 1, 1): Inclusion.

Experiment and Evaluation

The initial simulation was carried out with a unified $S = (1, -1, -1, 1)$ setting for all agents of both classes leading to classical segregation structures (strong isolated clusters), shown in Fig. 12.3. The social interaction distance was fixed $r = 3$; In a second run, digital twins retrieved form crowd sensing surveys were added to the simulation dynamically. Some results are shown in Fig. 12.4. Now, the S vector and the social distance r depend on the answers given by the (real) humans, which can differ from the initial S setting and the social interaction distance r. These agents (if their S differs from the basic model) create a disturbance in the segregation patterns. Agents with $s_{\alpha\beta} = 1$ and $s_{\alpha\alpha} = 1$ can be integrated in both groups and are able to bind different groups close together (see the development and movement of the blue and red clusters inside the red circle in Fig. 12.4).

The crowd sensing extends the simulation with the following dynamics and changes:

1. Disturbance of the synthetic simulation with digital twins not conforming to an initial parameterisation of the artificial individuals (affecting S vector, spatial social interaction distance, mobility)
2. Adaptation and change of the fraction of different groups (commonly a equally distributed fraction from each group is assumed)
3. Convergence and divergence of the emergent behaviour of group formation (spatial organisation structures).

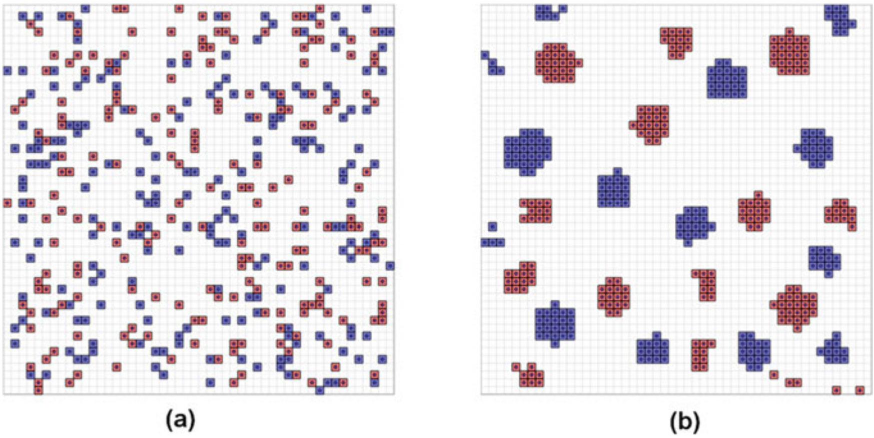

Fig. 12.3 **a** Simulation world consisting of 200/200 artificial agents of class a/b (blue/red squares), randomly distributed **b** Simulation world after social organisation based on mobility forming strong isolated homogeneous clusters with S = (1, −1, −1, 1) and r = 3 interaction radius after 200 simulation steps

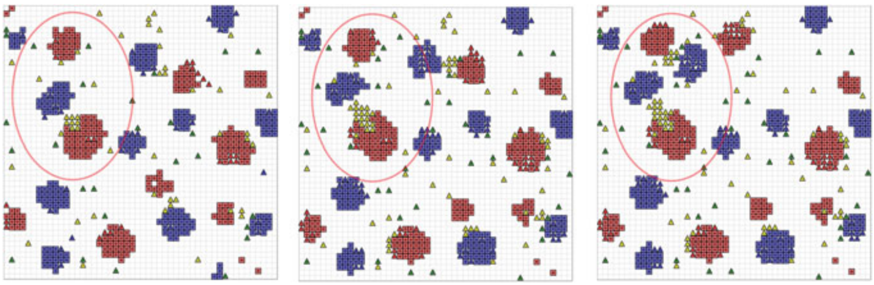

Fig. 12.4 Simulation world at different simulation times (500/1000/1500 steps) consisting of 200/200 a/b class agents (blue/red squares) all with S = (1, −1, −1, 1) and r = 3 parameter settings and additionally up to 200 digital twins (triangles with colour based on individual S/r parameters)

The comparison of the simulation with the single and the multi-parameter social behaviour derived by CWS is shown in Table 12.1. The mean distance between group *a* and *b* clusters (in mesh grid units measured between the center points of group clusters) show no significant difference, but the mean distance between clusters of same group increase by about 20% as well as an increased variance in the distance (increased disturbance).

Remarkable is the creation of heterogeneous clusters (group *a* and *b*) by digital twin agents with a variation of social behaviour. The mean cluster size of these meta clusters are 4 times larger than the mean size of single clusters. The mean distance gap of *a* to *b* groups in such meta clusters is about 5 grid units and much lower than the mean distance in other areas.

Table 12.1 Comparison of the mean distance of blue (group a) and red (group b) clusters (ab) and the mean distance of same group distances (aa, bb) without[1] and with[2] social behaviour variations by CWS

Group	Mean	Variance	Min	Max	#Clust	Mean(cs)
ab[1]	10.6	3.7	5.6	20.8	–	–
aa[1]	14.5	2.9	10.4	18.5	11	16
bb[1]	12.4	2.2	8.5	16.9	13	15
ab[2]	11.7	3.7	6.9	20.3	4 (11)	64
aa[2]	17.0	5.5	9.9	28.5	11	18
bb[2]	16.9	3.9	10.8	22.5	10	21

Finally, the injection of the digital twins with behaviour variations leads to a broader and scattered area coverage, weakening segregation.

Conclusion

The MAS simulation framework, introduced in this work, is suitable to combine social and computational simulations with real-world interaction at run-time by integrating crowd sensing and by using mobile agents. A simple use-case demonstrated social interaction and social structure formation based on a parametrisable Sakoda model extended with digital twins retrieved by agent-based crowd sensing integrated in and extending agent-based simulation. The crowd sensing surveys performed by mobile agents is used to create digital twins of real humans in the simulation world (with respect to the social interaction model and mobility) based on individual surveys via a chat bot dialogue. The injected twins posing behaviour variance introduce disturbance in the simulation and the emergence outcome.

References

1. F.-Y. Wang, Agent-based control for networked traffic management systems. Intell. Transp. Syst. (2005)
2. H. Hamidi, A. Kamankesh, An approach to intelligent traffic management system using a multi-agent system. Int. J. ITS Res. **16** 112–124 (2018)
3. O. Baqueiro, Y.J. Wang, P. McBurney, F. Coenen, *Integrating Data Mining and Agent Based Modeling and Simulation*, in *ICDM 2009, LNAI 5633*, (2008)
4. S. Bosse, U. Engel, *Augmented Virtual Reality: Combining Crowd Sensing and Social Data Mining with Large-Scale Simulation Using Mobile Agents for Future Smart Cities*, Proceedings, Volume 4, ECSA-5 5th International Electronic Conference on Sensors and Applications 15–30 November, (2018). doi: https://doi.org/10.3390/ecsa-5-05762
5. U. Wilensky, W. Rand, *An Introduction to Agent-Based Modeling Modeling Natural, Social, and Engineered Complex Systems with NetLogo*, MIT Press, (2015)

6. T. Calenda, M.D. Benedetti, F. Messina, G. Pappalardo, C. Santoro, *AgentSimJS: A Web-based Multi-agent simulator with 3D capabilities*, (2016)
7. R.K. Ganti, F. Ye, H. Lei, Mobile crowd sensing: current state and future challenges. IEEE Commun. Mag. 49(11) (2011)
8. K.H. van Dam, I. Lukszo, N. Zofia. (eds.), *Agent-Based Modelling of Socio-Technical Systems*, (Springer Berlin, 2013)
9. T. Leppänen, J.Á. Lacasia, Y. Tobe, K. Sezaki, J. Riekki, Mobile crowdsensing with mobile agents. Auton. Agent. Multi-Agent Syst. **31**(1), 1–35 (2017)
10. P. Medina, E. Goles, R. Zarama, S. Rica, Self-Organized societies: on the sakoda model of social interactions. Complexity (2017)