The Effect of Heterogeneous Agents in Socio-Cognitive Networks

Changkun Zhao, Jui-Te Tseng, Frank E. Ritter
The College of Information Sciences and Technology
The Pennsylvania State University

Keywords: Social Cognition, Heterogeneity, ACT-R, Agent-based Simulation

ABSTRACT: Many network simulations have used homogenous agents, that is, the population of agents is of one type. This study examined if heterogeneous agents created different socio-cognitive networks than homogeneous agents. Here, cognitive agents in a simple virtual world through their interactions formed networks. The agents were different in movement strategies and starting locations. We found that: (a) the difference between homogeneous agents and heterogeneous agents, and (b) the starting locations of agents, both affect network formation. The result of heterogeneous agents in network formation is different from the average results of homogeneous agents and indicates heterogeneity fundamentally changes network formation. Consequently, future network simulation should include heterogeneous agents.

1. Introduction


Nevertheless, all these existing researches of social agent models do not often consider individual differences. Human variability is an important parameter that should be considered more often. It is suggested by Ritter and Norling (2006), among many others, that variability plays an important role in influencing task completion within a group. Due to their different properties and behaviors, heterogeneous agents seem to be the best tool to use to model and account for this variability to improve the accuracy of future network simulations. With agent-based simulation, we can explore social network formation.

To explore the effect of heterogeneity of agents on networks, we designed two experiments with: (a) different starting locations, and (b) different types of agents. We found the results with heterogeneous agents were different from networks with homogeneous agents, and the values of heterogeneous agents were not the average of the values from two sets of homogeneous agents. Thus, running experiments with homogeneous agents will often be unpredictable of the real world, and including heterogeneous agents in simulations may reflect the real case should be considered more often in future work.

2. Background

This section will cover some previous studies and systems from four different perspectives that are used to create our agent networks.

2.1 Social agents modeling

A lot of work has been done in agent-based simulation. One of the first is Schelling (1971), who studied dynamic models of segregation. Other work includes Carley and Gasser (1999), who modeled organization theory; Carley and Newell (1994) discussed the Model Social Agent; Ritter et al. (2004) studied models related to behavior moderators. However, these studies do not have a large-scale simulation focused on networks.

Some further work expanded the simulation of heterogeneous agents to form a Multi-Agent System (MAS). A MAS is a complex systems with independent agents interacting with the environment and other agents. MAS's are required in some domains because the interaction between different subjects need to be handled (e.g., Stone and Veloso, 2000). MAS's can be used to solve complex and unpredictable problems (Sycara, 1998) because they split the problem into subproblems and solve them in parallel. The characteristics of MAS's are (a) each agent has incomplete information for solving the problem, (b) there is no global system control, (c) data are decentralized, and (d) computation is asynchronous (Sycara, 1998). The motivation to use MAS's includes: (a) the problem domain is too large to solve by a single agent, (b) the way to solve the problem can be done by interactions between different components of the system, (c) data need to be or is better to be distributed, and (d) some system performance can be improved, including computational efficiency, reliability, extensibility, robustness, maintainability, responsiveness, exibility, and reuse (Sycara, 1998).
Moreover, Stone and Veloso (2000) separated MAS's system into four categories based on agent variance and communication between agents: (a) homogeneous agents with no communication, (b) homogeneous agents with communication, (c) heterogeneous agents with no communication, and (d) heterogeneous agents with communication. In the second and fourth cases, if agents can communicate seamlessly, we can regard the whole system as a single complex agent. Our work mainly focused on the first and the third cases, they are homogeneous agents with no direct communication and heterogeneous agents with no direct communication (such as human agents).

Related work of MAS's includes: Davidsson (2001), who compared multi-agent based simulation (MABS) with parallel and distributed discrete event simulation, object oriented simulation, and dynamic micro simulation, and suggested some advantages of MABS, such as MABS is easy to simulate persons with different behavioral characteristics. Morgan et al. (2010) studied modeling of participation for small groups. As previous mentioned, this research differs because it focuses on a large-scale network formation simulation with heterogeneous agents.

2.2 Network analysis measures
Network analysis studies status, dynamics, and predictions of complex networks. Many measures have been applied to represent or predict network information in biology, physics, economic, social sciences, etc. In this study, we will use these five measures to interpret our results.

Edges: An edge represents a connection between two nodes in the graph theory. In this work, an edge has a cognitive meaning; it is a bi-directional memory retention between agents.

Assortativity (coefficient): Assortativity is a measure of homophily between nodes in a network based on some label or value assigned to a node, usually a node's degree. Homophily implies that we tend to be similar to our friends (Easley & Kleinberg, 2010).

Global Cluster Coefficient: The cluster coefficient, or transitivity (Wasserman & Faust, 1994, p. 243), is a measure that indicates the nodes in a graph tend to cluster together. If the value is higher, then it means the nodes are grouped with higher density.

Diameter: In a graph the distance between two vertices is the number of edges that connected them together in the shortest path. It is also known as geodesic distance. Graph diameter is the maximum geodesic distance in a graph.

Reciprocity: Reciprocity measures the mutual connections between nodes in a directed network. Garlaschelli and Loffredo (2004) defined reciprocity as the correlation coefficient between the entries of the adjacency matrix of a directed graph.

2.3 ACT-R
ACT-R (Anderson, 2007) is a cognitive architecture as well as a unified theory of cognition. It tries to provide a fully functional system that produces all aspects of human behavior, particularly cognition. We chose ACT-R because its declarative memory mechanism enables us to fully implement the cognitive capacities that influence the emergence of networks as we discussed above.

2.4 VIPER
VIPER (Hiam et al., 2011; Zhao et al., 2012), is a text-based, multi-agent simulation environment that is built on top of NakedMud, a content-less MUD engine. VIPER’s server-client architecture support client agents running independently. The server and clients communicate through the Telnet protocol with text messages, a common MUD approach.

Zhao et al. (2011, 2012), and Kaulakis et al. (2012) have done experiments using VIPER to explore factors such as environment configuration, agent size, and knowledge. But none of these experiments have examined the effect of heterogeneous agents, which is what this study is concerned with.

In this study, we use VIPER to test the effect of heterogeneous agents in the network, and we decided to focus on the effect of starting locations, and the heterogeneity of navigation patterns.

3. Experiment settings
To test our assumption that the starting location and agent heterogeneity influence network formation, In the experiment, our ACT-R agents can “memorize” relations by adding a chunk in their declarative memory. When agents meet their “old friends”, the relation chunks will be recalled, and the activation of the relation chunks will be stimulated by reactivation.

To build a socio-cognitive network, VIPER can output all relation chunks of the agents. Consequently, we can build a network by merging of these relations chunks, and each link in the network will have an activation value to represent the relation weight between each pair of agents.

In the two experiments, we used 30 agents and the same map layout in VIPER. Figure 1 shows the layout is a 5 by 5 grid. The running time (300 s) was chosen based on our previous work (Zhao et al., 2012), where after 300 s, the number of links formed is relatively
stable. There was a delay of 16 s after every movement, and also a 4 s delay after other actions.

Figure 1. The layout of the 5 by 5 grid map

3.1 Experiment 1
Brantingham and Brantingham (1993) examined the relations between spatial environment and social network formations because spatial layout can sometimes determine social networks by alter meeting chances. Based on this theory, we examine the effect of different starting locations in network formation because it might alter the meeting chances and frequencies between agents. There are four types of starting locations: (1) center: all the agents start in the center of the map; (2) corner: all the agents start from a corner of the map; (3) random: the agents are assigned their starting location randomly with replacement; (4) equally distributed: the agents are assigned their starting location equally distributed across all the rooms. All four settings were run 10 times with randomly moving agents.

3.2 Experiment 2
In the second experiment, we wanted to examine the effect of heterogeneous agents in the network formation. We used three sets of agents: (1) All agents are random-walk agents. (2) All agents are round-walk agents (moving along a 3x3 square route). (3) Half of the agents are random-walk agents and half of the agents are round-walk agents. There are also four types of starting locations as Experiment 1: (1) center: all the agents start from the center of the map; (2) corner: all the agents start from the lower left corner of the map; (3) random: the agents are assigned their starting location randomly; (4) equally distributed: the agents are assigned their starting location equally distributed across all the rooms. Based on these factors, we have 12 conditions, and we ran each one 10 times.

4. Result and analyses
In the following sections, we use egocentric data from the agents for our network analysis. If an agent A has a memory record of another agent B, there is a directed link from A to B with a memory activation value as a link weight. Zhao et al. (2013) found that having a threshold of the activation values changes the total number of links in a network. With perfect memory, the total number of edges formed in a network is larger than the condition with a memory activation threshold. Figure 2 shows a sample egocentric network that we created based on the experiment output.

Figure 2. A sample egocentric network

4.1 Result of Experiment 1
The network measures are shown in Figure 3. The associated statistical tests are shown in Table 1.

**Edges.** Edges represent the relationships between agents. The cumulative data shows an overview of the network formation. The maximum count of edges is 30 * 29 = 870 bi-directional links (a relation can be recalled on both sides). Figure 3(a) shows that the starting locations influenced the number of edges formed during simulation. The center starting location and the corner starting location are two starting locations where agents all started in the same room. As a result, these two types of starting locations encouraged agents to form more relationships. On the other hand, randomly distributed starting locations and equally distributed starting locations have higher assortativity and there is no significant difference between these two data sets (with t(10) = 1.255). This again explained that the starting location is a factor that influenced network formation, and randomly distributed starting locations and equally distributed starting locations are similar.

**Assortativity.** Assortativity implies that nodes tend to be similar to their connected nodes. Figure 3(b) shows that the data sets with the center starting location and the corner starting location have disassortativity as the t-value between them is significant (t=6.05, N=10, see Table 1). On the other hand, the data sets with randomly distributed starting locations and equally distributed starting locations have higher assortativity and there is no significant difference between these two data sets (with t(10) = 1.255). This again explained that the starting location is a factor that influenced network formation, and randomly distributed starting locations and equally distributed starting locations are similar.
Figure 3. Network measures for different starting locations.

**Global cluster coefficient.** In graph theory, a clustering coefficient indicates how much the nodes in a graph cluster together. The global cluster coefficient gives the global indication of such properties. Figure 3(c) shows that with the center starting location or the corner starting location, the network is highly clustered together because the global cluster coefficient is around 1.0. On the other hand, the global cluster coefficient with the random distributed starting location and equally distributed starting location are around 0.7, which is less clustered than the center starting location and the corner starting location. The t-value between the center and corner starting location is 2.498, and the t-value between randomly distributed and equally distributed starting locations is 0.786. Both t-values show a strong relation between two data sets.

**Diameter.** Figure 3(d) shows that the diameter with the center and the corner starting location share similar numbers, which are around 5, and the t-value of this pair is 0.526. On the other hand, the random and the equally distributed starting location have similar diameter values, which are about 12, and the t-value of this pair is 0.058. This measurement also supports our other measurements that the agents tend to cluster together when the starting location is the center or the corner.

**Reciprocity.** This measure is the probability of the mutual connections between nodes in a directed network. Figure 3(e) shows that this attribute is different from other attributes with only the corner starting location has difference than other starting locations, and other starting locations do not have a reliable difference.
Table 1. The t-test values for experiment 1 between starting location pairs (N=10).
Values in bold are reliable, df=19, p < .05, two-tailed.

<table>
<thead>
<tr>
<th>Starting location pairs</th>
<th>Edges</th>
<th>Assortativity</th>
<th>Global Cluster Coefficient</th>
<th>Diameter</th>
<th>Reciprocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center/Corner</td>
<td>7.15</td>
<td>6.05</td>
<td>2.50</td>
<td>0.53</td>
<td>12.33</td>
</tr>
<tr>
<td>Center/Random</td>
<td>34.99</td>
<td>13.51</td>
<td>20.59</td>
<td>5.61</td>
<td>0.15</td>
</tr>
<tr>
<td>Center/Equally Distributed</td>
<td>47.32</td>
<td>19.84</td>
<td>49.84</td>
<td>15.24</td>
<td>0.50</td>
</tr>
<tr>
<td>Corner/Random</td>
<td>39.01</td>
<td>10.86</td>
<td>21.21</td>
<td>5.68</td>
<td>8.17</td>
</tr>
<tr>
<td>Corner/Equally Distributed</td>
<td>49.57</td>
<td>17.15</td>
<td>53.99</td>
<td>15.05</td>
<td>13.94</td>
</tr>
<tr>
<td>Random/Equally Distributed</td>
<td>1.30</td>
<td>1.26</td>
<td>0.79</td>
<td>0.06</td>
<td>0.48</td>
</tr>
</tbody>
</table>

4.3 Result of Experiment 2
The result of experiment 2 are more complex because we examined a larger set of parameters and conducted 60 t-test analyses. Figure 4 shows the five network measures for the parameter sets. Table 2 provides he t-test results, where significant results are bolded.

Edges. Figure 4(a) shows that starting locations had a greater influence than the types of agents. When the starting location is randomly distributed or equally distributed, even the difference between round-walk agents and round-walk agents is not significant, so heterogeneity does not have much effect. However, when the starting location is center or corner, there is the difference between random-walk agents and round-walk agents, and heterogeneity agents also form different number of edges. The only insignificant difference is between round-walk agents and heterogeneous agents with the center starting locations.

Assortativity. The center starting location and the corner starting location conditions have disassortativity and the data sets with random starting locations and equally distributed starting locations have higher assortativity. There is an exception, when round-walk agents start in the corner; the data have variance that leads to high values of assortativity. This is because the round-walk agents sometimes become stuck at the edges of the grid.

Figure 4. Network measures for different starting locations and agent types.
Considering heterogeneity, we found that with the center starting location, the corner starting location, and randomly distributed locations, heterogeneous agent set has the mean value between the two homogeneous sets. But when the starting location is randomly distributed, heterogeneous agents have less assortativity than the other two types of homogeneous agent sets and the value of assortativity is near 0. Table 2 includes t-values of the assortativity between agent set pairs. It shows that heterogeneity influences assortativity in most of the cases.

**Global cluster coefficient.** Experiment 1 shows that when the starting location is at the center or corner, the agents are highly clustered together because the global cluster coefficient is around 1.0. In Figure 4(c), it also reproduces the result and the difference between homogeneous agents and heterogeneous agents is less than 0.25.

Also, though the values of global cluster coefficient with random-walk agents and round-walk agents are both around 1, heterogeneous agents have lower values. This shows heterogeneous agent sets change network properties.

On the other hand, when the starting location is randomly distributed, heterogeneous agents have lower values; when the starting location is equally distributed: heterogeneous agents have a value between the two homogeneous agents sets. The value shows that heterogeneous agents have different networks based on this global cluster coefficient. In Table 4, the fourth column presents the values of the global cluster coefficient between agent set pairs.

**Diameter.** Experiment 1 shows that with the center or the corner starting location, the diameter of the network formation graphs is smaller. The result in Figure 4(d) follows this analysis, and the heterogeneous agents also moderate the network formation. With the center starting location and the equally distributed starting locations, heterogeneous agents have a value between homogeneous agents. But with the corner starting location, heterogeneous agents have a higher diameter; with the randomly distributed starting location, heterogeneous agents have a lower diameter.

**Reciprocity.** Figure 4(e) shows that this attribute is different from other attributes, where only the corner starting location has a difference than other starting locations. The reciprocity with round-walk agents and the corner starting location is near 1.0, but when using heterogeneous agents, the value is closer to the results with only random-walk agents. Table 4 includes the t-values of the reciprocity between agent set pairs.

<table>
<thead>
<tr>
<th>Starting location</th>
<th>Agent set pairs</th>
<th>Edges</th>
<th>Assortativity</th>
<th>Global Cluster Coefficient</th>
<th>Diameter</th>
<th>Reciprocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center</td>
<td>Random/round</td>
<td>9.66</td>
<td>10.69</td>
<td>2.70</td>
<td>3.70</td>
<td>18.37</td>
</tr>
<tr>
<td>Center</td>
<td>Random/Hetero</td>
<td>10.31</td>
<td>3.46</td>
<td>6.87</td>
<td>4.11</td>
<td>16.10</td>
</tr>
<tr>
<td>Center</td>
<td>Round/hetero</td>
<td>0.99</td>
<td>6.33</td>
<td>2.55</td>
<td>1.52</td>
<td>1.08</td>
</tr>
<tr>
<td>Corner</td>
<td>Random/round</td>
<td>14.18</td>
<td>2.69</td>
<td>1.00</td>
<td>3.43</td>
<td>22.64</td>
</tr>
<tr>
<td>Corner</td>
<td>Random/Hetero</td>
<td>7.21</td>
<td>3.35</td>
<td>3.87</td>
<td>2.09</td>
<td>4.96</td>
</tr>
<tr>
<td>Corner</td>
<td>Round/hetero</td>
<td>19.57</td>
<td>1.99</td>
<td>3.87</td>
<td>4.45</td>
<td>52.04</td>
</tr>
<tr>
<td>Randomly</td>
<td>Random/round</td>
<td>0.69</td>
<td>4.15</td>
<td>6.48</td>
<td>1.73</td>
<td>2.93</td>
</tr>
<tr>
<td>Randomly</td>
<td>Random/Hetero</td>
<td>0.23</td>
<td>5.01</td>
<td>8.14</td>
<td>2.45</td>
<td>2.72</td>
</tr>
<tr>
<td>Randomly</td>
<td>Round/hetero</td>
<td>0.51</td>
<td>0.85</td>
<td>0.08</td>
<td>1.26</td>
<td>0.95</td>
</tr>
<tr>
<td>Equally</td>
<td>Random/round</td>
<td>1.04</td>
<td>6.38</td>
<td>16.15</td>
<td>6.22</td>
<td>10.11</td>
</tr>
<tr>
<td>Equally</td>
<td>Random/Hetero</td>
<td>1.78</td>
<td>2.42</td>
<td>3.37</td>
<td>2.67</td>
<td>3.15</td>
</tr>
<tr>
<td>Equally</td>
<td>Round/hetero</td>
<td>0.72</td>
<td>2.95</td>
<td>9.77</td>
<td>3.75</td>
<td>6.35</td>
</tr>
</tbody>
</table>

5. **Discussion and Conclusion**

With two different models, which use different, simple navigation strategies, we ran experiments with both homogeneous agent sets and heterogeneous agent sets in different starting locations in a cognitive agent-based simulation.

5.1 **Experiment 1**

The result shows that the starting locations of agents influence the network formation strongly. Testing four kinds of starting locations, we found that the result of center location and corner location is very similar, as well as the randomly distributed location and the equally distributed location.

For the edge count, the mean value of the randomly distributed starting locations and the equally distributed starting locations are around 200 edges, where the mean value of the center starting location and the corner starting location are around 600 edges. This implies that center and corner locations provide a better chance for agents to meet.

For assortativity, the results of the randomly distributed
starting locations and the equally distributed starting locations show assortativity, but the results of the center starting location and the corner starting location do not show assortativity as the measure between them is not significant.

For the global cluster coefficient, the mean value of the randomly and the equally distributed starting locations are around 0.7, where the mean value of the center starting location and the corner starting location are around 1. Comparing the two sets of global cluster coefficient results, it is clear to the randomly and the equally distributed starting locations could makes the network tighter than center and corner starting locations. The measures of diameter also support this conclusion as the randomly and equally distributed starting location (diameter around 5) has much smaller diameter measures than the center and corner starting locations (diameter around 12).

Reciprocity is the only value that the results of the corner starting location is different from other three, where the mean value is 0.65, and the other three are between 0.45 and 0.5. These results show that when doing a network formation simulation with spatial settings, one should consider and report the starting locations of agents and how they are distributed. Dunbar (1998) claims that on average we have 150 friends memorized. If we take the number into consideration and want a normalized simulation, then the results of putting agents distributed on a map is closer to it. On the other hand, starting agents together simulates other situations.

5.2 Experiment 2
This experiment shows that heterogeneous agents in comparison with homogeneous agents change the network formation, though the effect on edges count, assortativity, global cluster coefficient, diameter, and reciprocity is not as strong as changing starting locations. When we compare the property values between agent model sets, we found that putting two types of agents together in equal percentages, the properties of the network graph are not simply the average of networks with two types of homogeneous agents. Moreover, the values from heterogeneous agents do not always fall between the two sets of homogeneous agents that make them up.

For the edge count, the mean value of the randomly distributed starting locations and the equally distributed starting locations are around 200. With the center starting location and the corner starting location, the values have variances between different agent sets. It implies that heterogeneous agents formed fewer edges than the other two types of agents alone.

For assortativity, most of the results of the randomly distributed starting locations and the equally distributed starting locations have assortativity, but the results of the center starting location and the corner starting location have disassortativity. Round-walk agents with the corner starting location is an exception, where the value possible to be associative and disassociative. But homogeneous agents make networks less associative for both associativity and disassociativity.

For the global cluster coefficient, the mean value of the randomly distributed starting locations and the equally distributed starting locations are between 0.5 and 0.7, where the mean value of the center starting location and the corner starting location are around 1. The heterogeneous agents lower the value of global cluster coefficient except when with the equally distributed starting locations.

For diameter, with the center starting location, heterogeneous agents have value between homogeneous agents; with the corner starting location, heterogeneous agents have value greater than homogeneous agents; with randomly distributed starting locations, heterogeneous agents have value smaller than homogeneous agents; with equally distributed starting location, heterogeneous agents have values between homogeneous agents but closer to the random-walk agents.

For reciprocity, with the center starting location, heterogeneous agents have a value between homogeneous agents but closer to the round-walk agents; with the corner starting location, heterogeneous agents have a smaller value than homogeneous agents; with randomly distributed starting locations, heterogeneous agents have a smaller value than homogeneous agents; With equally distributed starting location, heterogeneous agents have a value between homogeneous agents but closer to the random-walk agents.

Together, these results suggest heterogeneous sets of agents change network formation. Using homogeneous agents appears to not represent diversity in agents when doing simulation, and those variances of agents needs to be considered. This also suggests that if we run the simulation with the agent model set that do not reflect the real condition, the simulation may fail to represent the real cases.

Consequently, this work shows that heterogeneity of agents is important; it influences every measures we examined of a network. Heterogeneous agents should be used to model humans if the humans are diverse.

5.3 Future work
This work opens up several new areas. Firstly, we will explore the effect of cognitive heterogeneity. More
specifically, we will modify cognitive parameters such as forgetting speed and mental noise in ACT-R to implement heterogeneous agents with different cognitive capabilities. Secondly, we will extend VIPER to support more complex behavior such as verbal communications and collaborations to achieve more realistic model compositions.

6. References

Author Biographies
CHANGKUN ZHAO is a PhD candidate in the College of Information Sciences and Technology. He is interested in spatial memory modeling and information processing theory.

JUI-TE TSENG recently received his master degree in computer science and engineering at Penn State.

FRANK E. RITTER is on the faculty of the College of IST, an interdisciplinary academic unit at Penn State to study how people process information using technology. He edits the *Oxford Series on Cognitive Models and Architectures* and is an editorial board member of *Cognitive Systems Research, Human Factors*, and *IEEE SMC: Human-Machine Systems.*