Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

ON THE EXTENSION OF SCHELLING'S SEGREGATION MODEL

Zeyu Liu
Xueping Li
Anahita Khojandi

Department of Industrial and Systems Engineering
The University of Tennessee, Knoxville
John D. Tickle Engineering Building
Knoxville, TN 37996, USA

Sanja Lazarova-Molnar

Mærsk Mc-Kinney Møller Institute
University of Southern Denmark
5230 Odense M
DENMARK

ABSTRACT

Schelling's social segregation model has been extensively studied over the years. A major implication of the model is that individual preferences of similarity lead to a collective segregation behavior. Schelling used Agent-Based Modeling (ABM) with uni-dimensional agents. In reality, people are multidimensional. This raises the question of whether multi-dimensionality can boost stability or reduce segregation in society. In this paper, we first adopt ABM to reconstruct Schelling's original model and discuss its convergence behaviors under different threshold levels. Then, we extend Schelling's model with multidimensional agents and investigate convergence behaviors of the model. Results suggest that if agents have high levels of demand for identical neighbors, the society might become less stable or even chaotic. Also, several experiments suggest that multidimensional agents are able to form a stable society that is not segregated, if agents prefer to stay adjacent to not only "identical" but also "similar" neighbors.

1 INTRODUCTION

Thomas Schelling developed a seminal agent-based model (ABM) in the 1970's (Schelling 1969; Schelling 1971) for which the result suggests that segregation in a society can form when individuals only adopt certain preferences of residing closer to those possessing similar attributes, but do not feel negatively towards those who are different. This model is enormously influential because the result provides with a simple but fundamental explanation of social segregation. Schelling modeled individuals as agents with only one attribute of two possible values (e.g. $\circ$ and $+$), distributing within some geometric shape that mimics the area of a city, in which some grids are occupied by agents and some are vacant so that agents can relocate. An agent is considered to be "satisfied" if the number of neighbors that are identical to the agent is above certain threshold. Unsatisfied agents are given the opportunity to relocate within the area of residence, to the nearest vacant positions that satisfy their needs. At each iteration, all unsatisfied agents can move once. Then, some previously satisfied agents may become unsatisfied; some previously unsatisfied agents will become satisfied. The model then goes to the next period, where currently unsatisfied agents are selected and given the chance to move again. The model runs in iterations until an equilibrium is reached, i.e, all agents are satisfied and no one is needed to relocate.

Many have studied Schelling's social segregation model and its variants. Clark (1991) evaluated Schelling's model using data collected through telephone interview conducted in several major cities in the U.S.. Pancs and Vriend (2007) analyzed Schelling's model using game theory approaches and came to the conclusion that social segregation is the result of best-responses of agents. Clark and Fossett (2008) studied Schelling's model under multicultural context, concluding a similar result as the original Schelling's model. However, their model adopted same uni-dimensional agent, where the only dimension is made possible to take multiple different values. Grauwin et al. (2012) utilized an analytical approach that proved the existence of a potential function of global configuration that is maximized under the stationary state. They also showed three analytic result: (i) that linear utility functions are the only ones that maximize collective social utilities; (ii) Schelling's original utility function boosts segregation at the expense of collective utility; (iii) the model converge to perfect segregation if agents have strict preference to mixed neighborhood and wish to stay in majority. Spaiser et al. (2018) applied Schelling's model to a complex dynamic system to the study of segregation in a school in Stockholm during 1990 and 2002. Thresholds regarding the school's ethnic composition, school quality and parents' income are established to prevent segregation. Paolillo and Lorenz (2018) considered the impact of shared tolerance towards ethnicity diversity, and extend Schelling's model by replacing two groups of agents with value-oriented and ethnicity-oriented. Results suggest that value-oriented agents help to reduce total segregation and strong ethnic homophily exacerbates overall segregation.

Several recent studies combined Schelling's social segregation model with Axelrod's culture dissemination model (Axelrod 1997), which is another influential agent-based model that studied the formation of communities in society. Differing from Schelling's model, Axelrod gave agents multiple features as dimensions of culture and each feature takes value from a set of traits. Agents in this model cannot move, but are able to influence their neighbors by their own culture traits. Gracia-Lázaro et al. (2009) combined Axelrod's cultural dissemination model with Schelling's social segregation model by introducing empty sites and allowing movements of agents to those sites if the *intolerance* of an agent exceed certain threshold. For small densities of sites, they showed that agents converge to a uni-cultural society quickly, but for larger densities of sites, communities of various cultures form because isolated settlements can no longer transmit cultural influences to each other. Neal and Neal (2014) focused more on interpersonal relationships and social networks based on the Schelling model with people of only two types. They explored cultural diversity and the sense of community of people in a society and defined the probability that two people became friends on *homophily* and *proximity*, i.e., similarity of two people and the distance between them. The result demonstrates a negative relationship between social diversity and the sense of community. Stivala et al. (2016) refuted such negative relationship by integrating Axelrod's model with Schelling's model. Specifically, they introduced multiple features for a single agent and allowed mutable features. They showed that under sufficiently large cultural diversity and mutable agents features, diversity and sense of community can co-exist.

In this paper, we partly combine Axelrod's model with Schelling's; that is, we replace uni-dimensional agents in the original Schelling model with multi-dimensional agents and define new parameters for the extended model so that other features of the model can be studied. We primarily focus on two aspects: social stability and social segregation. Here, we use the convergence property of a model to represent stability. If a model converges to some steady state, we say the model suggests a stable society; otherwise the society is chaotic. The significance of social segregation can be directly observed from the model output: whether agents with the same attributes are blocked together, or are they all mixed with other types of agents. In Section 2, we define Schelling's original model as well as our extended model using formal mathematical notations. In Section 3, we show simulation results of Schelling's model and our extended model. We also discuss convergence behaviors and segregation patterns of the two model. In Section 4, we summarize our work in this paper, draw conclusions and state future prospects of this work as well as limitations.

2 METHODS

2.1 Schelling's Model

We reconstruct Schelling's original model using formal mathematical notations, while generalizing several concepts so that the model can be easily extended. Consider a society of population M on an $L \times W$ area, where population density can be computed as $M/(L \times W)$. Each agent $i, i = 1, 2, \ldots, M$ is assumed to have a set of K attributes, denoted by $A_i := \{a_1, a_2, ..., a_K\}$, $K \in \mathbb{N}^+$, and each attribute a_k can take different values as $a_k \in \{1, 2, ..., q_k\}$, where $k = 1, 2, .., K$ and q_k denotes the maximal number of options for each attribute. We assume that q_k can take different values for different attributes a_k. In Schelling's model, agents are uni-dimensional, with only one attribute. In this case $K = 1$ and $A_i = \{a_1\}$. Schelling also assumed that two types of agents are present, which is interpreted as $a_1 \in \{1, 2\}$.

Initially, the population of M agents distribute among $L \times W$ area randomly, with even number of agents for attribute $a_1 = 1$ and $a_1 = 2$ (a_1 is the only attribute we have in this model). We use $N(i)$ to denote the set of neighbors of agent i. An agents decides where it it satisfied with current location of residence based on how many identical neighbors it has, which can be expressed as $1 - \delta_{A_iA_j}$, where $\delta_{A_iA_j}$ is the Hamming distance between A_i and A_j, i.e., the number of different items in two lists. The level of satisfaction can be modeled by an utility function

$$u(i) = \begin{cases} 1 & \text{if } \theta \geq \tau; \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where

$$\theta = \frac{\sum_{j \in N(i)} (1 - \delta_{A_iA_j})}{|N(i)|} = 1 - \frac{\sum_{j \in N(i)} \delta_{A_iA_j}}{|N(i)|} \tag{2}$$

is the fraction of neighbors identical to agent i, with respect to all the neighbors of i and τ is the threshold. In Schelling (1971), the threshold τ is set to be 50%. If the utility function $u(i) = 1$, agent i is satisfied with its current location; if $u(i) = 0$, under some probability

$$p_{move}(i) = 1 - u(i), \tag{3}$$

agent i move to the nearest location in the area that satisfies its demand. The same probability function applies when agent i is satisfied. In such case $p_{move} = 1 - u(i) = 1 - 1 = 0$.

2.2 Schelling's Model Extended

Schelling's original model is an oversimplified version of relocation behaviors in society, where two groups of people decides where to live solely on the neighborhood. However, in reality, different situations might lead to violations of these assumptions. For example, some people may have a different preference level than others regarding identical neighbors; some may not consider neighborhood to be the only deciding factor for relocation and most people in society have multi-dimensional attributes. These unaddressed issues of the original model offer opportunities to extensions that capture more realistic features of how communities are formed during relocation. In this section, we extend the original model so that it is capable of incorporating these new features.

People in society are rarely uni-dimensional. When interacting with each other, people evaluate different aspects to decide whether they have something in common, or how "similar" they are. Under such scenario, the model assumes that for each agent i, there is a set of attributes $A_i = \{a_1, a_2, ...a_K\}$, where $K \geq 1$ and for each attribute a_k, there's more than two value options, $a_k \in \{1, 2, ..., q_k\}$, where $q_k \geq 2$. If this is the case, we will have in total $\prod_{k=1}^{K} q_k$ types of agents, each with a unique attribute vector. With the expansion of agent types, agents will find it more difficult to be surrounded by others that are "identical" to themselves. This is analogous to what is observed in real life: people rarely find neighbors or friends that have everything in common with them. Rather, they are satisfied when surrounding with "similar" people, i.e., people with whom they share "something". To incorporate this idea into our model, we use Hamming distance (Hamming 1950) of two attribute vector to model the "similarity" between agents. For two agents i and j, the Hamming distance $\delta_{A_iA_j}$ shows how many attributes that they have are different. We use this to define similarity of agents i and j to be

$$\sigma(i, j) = 1 - \frac{\delta_{A_iA_j}}{K}. \tag{4}$$

Here, different levels of threshold can be applied to reflect how people value similarity when interacting with others. We denote such threshold by τ_s and call it the *similarity threshold*. For an agent i, the set of similar neighbors is $S_i := \{j : j \in N(i), \sigma(i, j) \geq \tau_s\}$. We call all neighbors j in the set S_i *effective neighbors* of agent i, i.e., neighbors satisfy agent i's requirements. We assume that the utility function used in Schelling's original model still applies, which gives

$$u(i) = \begin{cases} 1 & \text{if } \theta_i \geq \tau_u; \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where $\theta_i = \frac{|S_i|}{|N(i)|}$ and τ_u is the utility threshold. When $\theta_i \geq \tau_u$, agent i is satisfied with its current neighborhood and does not plan to move. When $\theta_i < \tau_u$, agent i is not satisfied and may move to another location with probability $p_{move}(i) = 1 - u(i)$, which is the same function as Equation 3.

3 ANALYSIS

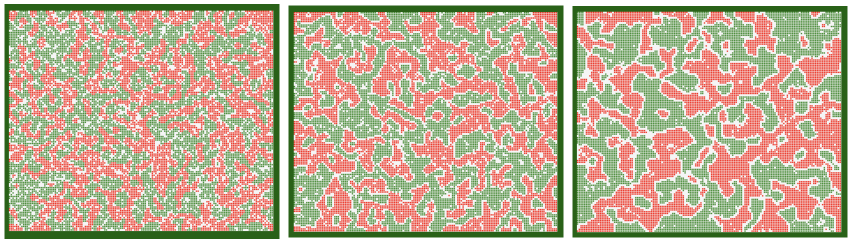
The agent-based models were built on AnyLogic (Personal Learning Edition, version 8.3.2), in reference to AnyLogic's *Schelling Segregation* model. In our model, the $L \times W$ residential area is built by the "Canvas" entity in AnyLogic. For connectivity, we used Moore Neighborhood, i.e., each agent has eight neighbors. Length and width of this canvas is set to $L = 150$ and $W = 125$, which offers 18,750 empty "houses". Agents are built based on the "Agent" entity in AnyLogic and are distributed randomly over the canvas in the initialization of each simulation. Attributes and colors are assigned to agents as their "parameters". Values for each attribute is assumed to distribute uniformly among all agents. Before presenting the result, *convergence* and *steady state* should be defined. A *steady state* of the model is a state where the probability of moving equals zero for every agent, i.e., none of the agents is relocating and all agents are satisfied with their current neighborhood. We say that the model *converges* if the model can reach a steady state in a number of iterations. Note that if a model reaches the steady state, it will there for all following iterations unless parameters are changed. However, true convergence of a model is hard to prove without theoretical analysis and it is impractical to run the model indefinitely. Thus, we say the model does not *converge in time limit*, if after a period of time, the model cannot reach the aforementioned steady state.

In each of the experiments, we use a simple but informative indicator to measure how segregated a society can be. Specifically, we add up the number of identical neighbors of all agents. Suppose for an agent i, the number of identical neighbor is N_i', then the total segregation level of a society is measured by $\sum_{i=1}^{M} N_i'$. In what follows, this indicator is applied to Schelling's model as well as extended models, providing a quantitative measurement of social segregation.

3.1 The Original Model

The major result of Schelling's model is well-known: although no one has an explicit negative feeling towards others that are different, a segregated society still forms when everyone is finally satisfied. In this section, we focus on the convergence analysis of Schelling's model. By using the same setting as Schelling's model, that is, 50% of the similarity threshold, we observed the same pattern (Figure 1a), in an area of population density 80% (15000 agents). Model time limit is set to be 1000 iterations, i.e., we stop the model if it cannot reach steady states within 1000 iterations. This limit is decided by running the model multiple times using different random seeds. This limit is chosen among a set of predetermined limits of 100, 300, 500, 1000, 5000, 10000. Under all limits larger than 1000 iteration, convergence of the model has similar behavior, so that 1000 is chosen as the limit.

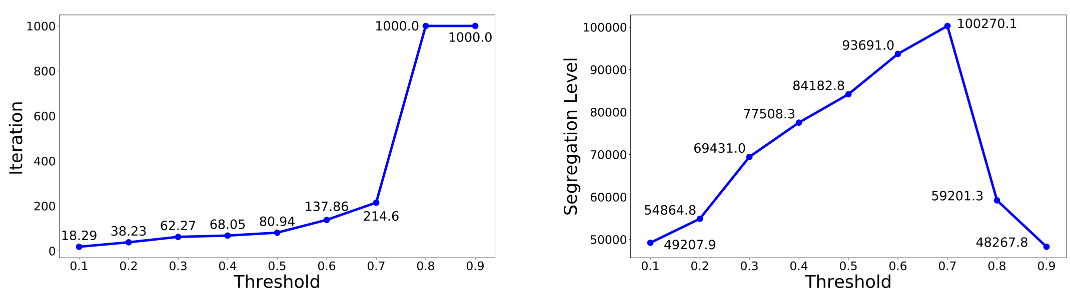
Boundaries of segregated communities change when model parameters change. Figure 1 shows the result of three different thresholds, form 50% to 70%. Boundaries between two classes of agents become clearer, suggesting an escalation of segregation behaviors. Under 50% threshold, agents desire four identical neighbors, which is easily satisfied. With the threshold increasing, say, to 70%, agents are no longer



(a) $\tau = 50\%$ (b) $\tau = 60\%$ (c) $\tau = 70\%$

Figure 1: Illustration results of the original Schelling model, using a fixed seed. (a) $\tau = 50\%$, the model converge into a steady state, with segregated patterns; boundaries are blurry. (b) $\tau = 60\%$, the model converge into a steady state, with segregated patterns; boundaries become clear. (c) $\tau = 70\%$, the model converge into a steady state, with segregated patterns; boundaries are most clear.

indifferent about “no neighbor” and “different neighbor”. For example, if some agent has 4 identical neighbor and 4 adjacent areas are left blank, the agent is satisfied. Now, if two different neighbors move into two of the blank areas, the percentage θ drops from 100% to 66.7%, which is below the threshold 70% and the agent becomes unsatisfied.



(a) (b)

Figure 2: Results of 100 runs using different random seeds. Mean value of 100 runs are calculated and plotted in the figure. (a) shows the changing of convergence iterations. When the threshold is 80% or 90%, the model does not converge within 1000 iterations. (b) shows the changing of social segregation level when the model is converged or reaches maximum iteration limit (1000 iterations). When the model does not converge within 1000 iterations, segregation levels are still recorded. Although the segregation level is low, many agents are unsatisfied and the society is highly unstable.

To evaluate the model and account for its stochastic nature, we ran 100 experiment per instance to collect data regarding model convergence and segregation degree. Figure 2a shows how many iterations it takes for models with thresholds ranging from 10% to 90% to converge. In the graph, lines represents the changing in mean values of 100 instances. Confidence intervals are too narrow to see if plotted on the figure, thus omitted. A clear increasing pattern can be observed. When agents have a higher similarity threshold, their preference towards identical neighbors alters their indifference in such a way that they prefer no neighbors rather than neighbors with different attributes. When the threshold is high enough, say 80%, no one can be satisfied with its current location, because everyone prefers identical neighbors over no neighbor. Figure 2b shows how social segregation levels change with thresholds. Social segregation level increases as threshold increases. When the threshold is high, the model does not converge within the time limit and agents keep moving to other locations, which renders a lower segregation level.

3.2 The Extended Model

Table 1: Agent types and corresponding colors for 2-attribute model.

Agent attributes	Color	Agent attributes	Color
$(a_1 = 1, a_2 = 1)$	Green	$(a_1 = 1, a_2 = 2)$	Yellow
$(a_1 = 1, a_2 = 1)$	Red	$(a_1 = 1, a_2 = 2)$	Blue

Table 2: Agent types and corresponding colors for 3-attribute model.

Agent attributes	Color	Agent attributes	Color
$(a_1 = 1, a_2 = 1, a_3 = 1)$	Green	$(a_1 = 1, a_2 = 1, a_3 = 2)$	Yellow
$(a_1 = 1, a_2 = 2, a_3 = 1)$	Red	$(a_1 = 1, a_2 = 2, a_3 = 2)$	Blue
$(a_1 = 2, a_2 = 1, a_3 = 1)$	Orange	$(a_1 = 2, a_2 = 1, a_3 = 2)$	Purple
$(a_1 = 2, a_2 = 2, a_3 = 1)$	Grey	$(a_1 = 2, a_2 = 2, a_3 = 2)$	Cyan

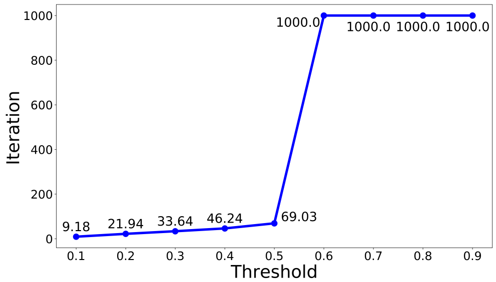
In order to analyze properties of the extended model but retain the convenience of visual observation, we build two basic models with 2 and 3 attributes each, where in both models the attributes draw values from set $\{1,2\}$. This gives 4 types of agents in the 2-attributes model and 8 types of agents in the 3-attributes model. Note that the model can be easily extended to have more attributes and agent types. In this work, 2 and 3 attributes are adequate for several important results and smaller scale models are more suitable for direct observation. Table 1 and 2 shows the colors and corresponding attributes for each type for 2-attribute and 3-attribute model, respectively. In the following analysis, illustrative figures are plotted using fixed random seeds so that they are repeatable. Indicators such as convergence iterations and segregation levels are evaluated on 100 runs of each model with fixed parameters but different random seeds.

3.2.1 Convergence

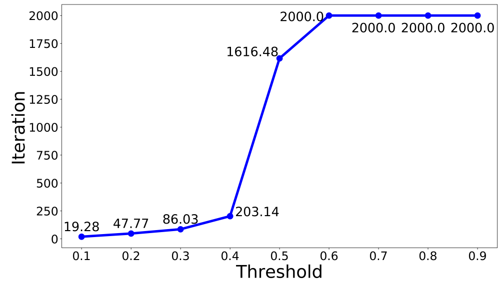
Firstly, we investigate the convergence behavior of the extended model. In this section, different time limits are imposed to different models. Time limits are chosen in a way similar to that of Schelling's original model. For 2-attribute model, we use 1000 iterations as the time limit, whereas for the 3-attribute model, we generally use 2000 iterations, unless otherwise specified.

To find what impact the utility threshold τ_u could have on convergence, we set τ_s to be 100%, i.e., a neighbor of an agent is effective if all their attributes are the same, for both 2-attributes and 3-attributes model, and simulate on different τ_u. Figure 3 shows the changing of iterations of different τ_u for 2-attribute and 3-attribute model. To account for the stochastic nature of the model, we also ran each instance 100 times. Mean of the 100 runs are calculated and plotted on the figure. As before, the scale of confidence intervals are too small compared to the scale of maximal iteration and are omitted in figure. When τ_u is low, both 2-attribute and 3-attribute models converge quickly. As τ_u increases, models takes longer to converge. When τ_u becomes high enough, say 60%, the models do not converge in time limit.

However, for several cases, not converging in time limit does not necessarily mean that model does not converge at all. For example, Figure 7c shows converged states of the model when τ_u is 60%. To obtain this result, utility threshold is raised twice when the model reaches a steady state for smaller thresholds. Firstly, the model is run under $\tau_u = 50\%$, until it converges. Figure 7b demonstrate the converged state of this instance using fixed random seed. Then, using the converged state as an initial state, we increase the utility threshold τ_u to 55%. At this moment, some of the agents at the edges of color blocks becomes unsatisfied and start relocation. These agents represents a small portion of the entire agent population, so that it is easier to find ideal neighborhood for them under a higher utility threshold, than the entire population. This explains why the model converged quickly for $\tau_u = 55\%$, as Figure 7b. Lastly, we increase τ_u to 60%,

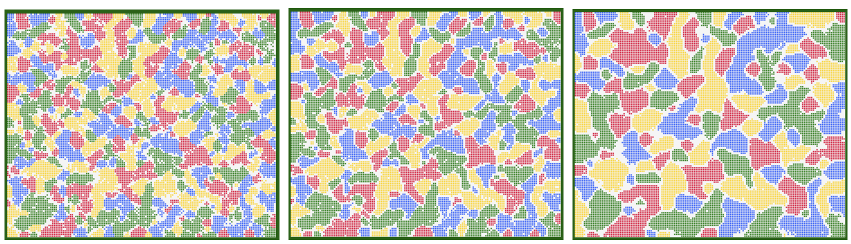


(a) 2-attribute, $\tau_s = 100\%$



(b) 3-attribute, $\tau_s = 100\%$

Figure 3: Iterations it took for models to converge. Iteration changes with the utility threshold τ_u, which is analogous to the threshold in Schelling's original model. In the figure, means are shown of 100 runs using different random seeds. Confidence intervals are too narrow to see if plotted in the figure, so they are omitted. (a) shows the changing of iterations of 2-attribute models. When the threshold is above 60%, the model does not converge within 1000 iterations. (b) shows the changing of iterations of 3-attribute models. When the threshold is above 60%, the model does not converge within 3000 iterations.



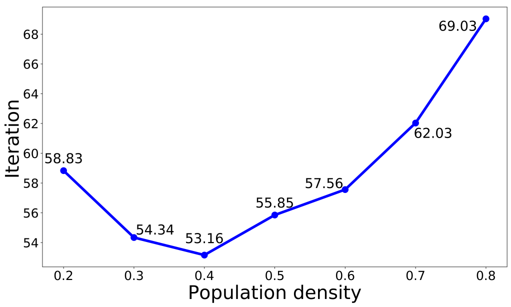
(a) 2 attributes, $\tau_u = 50\%$ (b) τ_u increased to 55% (c) τ_u increased to 60%

Figure 4: Illustrative results of the steady state of 2-attribute models, where $\tau_s = 100\%$ and population density is 80%. Boundaries between communities become more clear as τ_u increases.

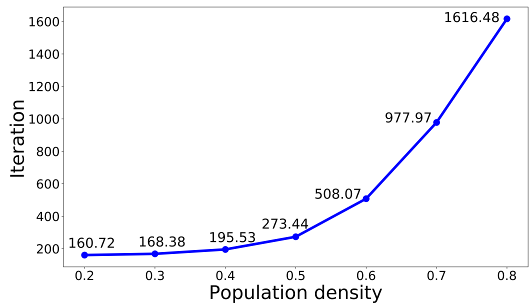
using the converged state of the last model ($\tau_u = 55\%$) as the initial state. The model converges within several iteration (shown in Figure 7c).

Decreasing population density accelerates the convergence process under a high value of τ_u. Mean value of convergence iterations of 100 replications for each instance were recorded and plotted in Figure 5, for 2-attribute and 3-attribute models. Population density changes from 20% to 80%. Surprisingly, the 2-attribute model shows a "U" shaped curve as population density increases. When population is quite scarce, although there are more vacant locations, it appears that agents are taking even longer to find a suitable neighborhood, because of a decrease in the number of identical agents. However, for the 3-attribute model, since we have eight classes of agents rather than four classes in the 2–attribute models, convergence iterations increases as population density goes up, with out the "U" shaped behavior. Extra vacant locations give unsatisfied agents more space to formulate larger color blocks before jumping to other areas.

When we relax the preference of agents by reducing similarity thresholds, convergence rates change. For 2-attribute model, τ_s ranging from 0% to 50% suggests exactly the same thing: if two agents has only one out of two identical attributes, they consider each other similar; τ_s from 50% to 100% suggests both attributes should be identical for two agents to consider each other similar. For 3-attribute model, the ranges are 0%-33.33%, 33.33%-66.67% and 66.67%-100%. Thus, we only choose two values (25%, 75%) of τ_s for 2-attribute model and three value (25%, 50%, 75%) for 3-attribute model. Table 3 shows mean and confidence interval (CI) of convergence iterations summarized from 100 runs per instance. Utility threshold



(a) 2-attribute models



(b) 3-attribute models

Figure 5: Iterations it took for models to converge. Population density changes form 20% to 80%, with 10% step increase. For 2-attribute models, τ_u is set to be 60% and τ_s is set to be 100%. To acquire faster convergence, we set τ_u to be 50% and τ_s to be 100% for 3-attribute models.

Table 3: A comparison of convergence iterations of 2-attribute and 3-attribute models, where population density is 80%, $\tau_u = 50\%$ and τ_s changes.

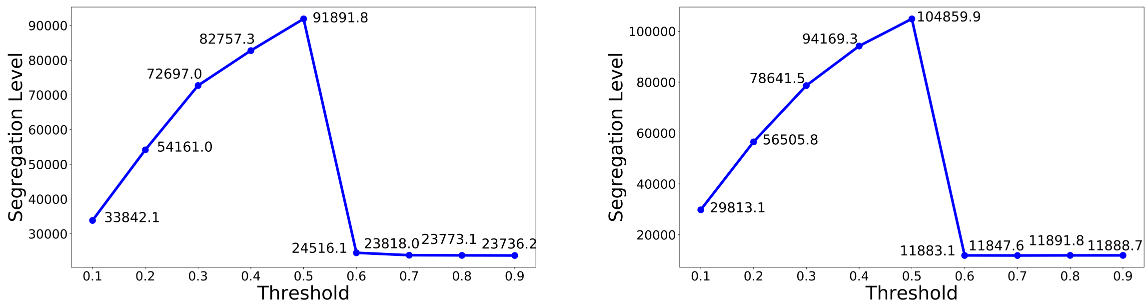
2-attribute models			3-attribute model		
τ_s	Mean Iter.	CI Iter.	τ_s	Mean Iter.	CI Iter.
25%	6.39	(6.23, 6.55)	25%	3.20	(3.12, 3.28)
75%	69.03	(67.85, 70.21)	50%	19.62	(19.03, 20.21)
-	-	-	75%	1616.48	(1576.21, 1656.75)

is set to be 50% for all instances and population density is 80%. Models with lower similarity thresholds converge faster. Under a low similarity threshold, agents have more similar neighbors. For example, in the 3-attribute model with $\tau_s = 100\%$, agents with attributes $\{a_1 = 2, a_2 = 1, a_3 = 1\}$ would only consider the exact same type of agents to be similar. However, if $\tau_s = 50\%$, 2 out of 3 same attributes makes a similar neighbor, suggesting 4 types of agents, with attributes $\{a_1 = 2, a_2 = 1, a_3 = 1\}$, $\{a_1 = 2, a_2 = 1, a_3 = 2\}$, $\{a_1 = 2, a_2 = 2, a_3 = 1\}$ or $\{a_1 = 1, a_2 = 1, a_3 = 1\}$. This expanded pool of similar neighbors makes it easier for agents to find a satisfying neighborhood. Thus the overall convergence rate drops dramatically.

3.2.2 Segregation Patterns

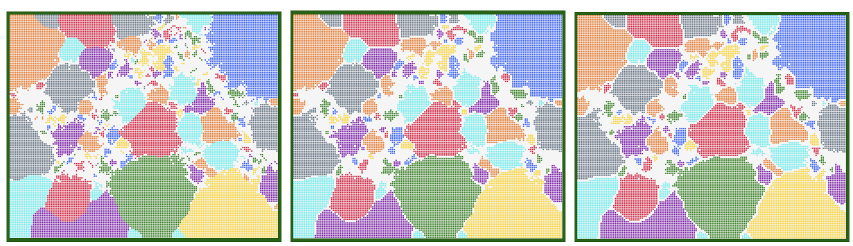
In this section, we investigate how segregation patterns changes if model parameters are tuned. Firstly, we fix similarity threshold τ_s to be 100% and look at model behaviors when τ_u and population density varies. When $\tau_s = 100\%$, segregation patterns appears in steady states of the model. Then, we let τ_s change. For lower similarity thresholds, we observed non-segregated society, i.e., agents no long stick together according to identical colors, rather, agents with different colors lives next to each other and remain satisfied.

We first look at how segregation level changes with utility threshold τ_u. We calculate segregation level of a society in the same way as in Schelling's original model. The number of identical neighbors of every agent in the canvas is added up to represent social segregation level. Figure 6 shows the changing in segregation levels with utility thresholds increasing. For both 2-attribute and 3-attribute model, social segregation becomes more severe under higher τ_u. Agents have strict preferences regarding neighborhood choices under high thresholds, so that identical neighbors are more prone to stay together, causing the segregation level to go up. In the figure, mean values were recorded and plotted in the figure for 100 runs. Confidence intervals are also omitted because they are not significant enough compared to the scale of change of segregation levels in the figure.



(a) 2-attribute, $\tau_s = 100\%$ (b) 3-attribute, $\tau_s = 100\%$

Figure 6: Segregation levels when models are converged into steady states. Segregation levels changes with the utility threshold τ_u. In the figure, means are shown of 100 runs using different random seeds. (a) and (b) shows the changing of segregation levels of 2-attribute and 3-attribute models, respectively. When the threshold is above 60%, models do not converge within time limit, but the segregation level is still recorded. Although the segregation level is low, many agents are unsatisfied and the society is highly unstable.

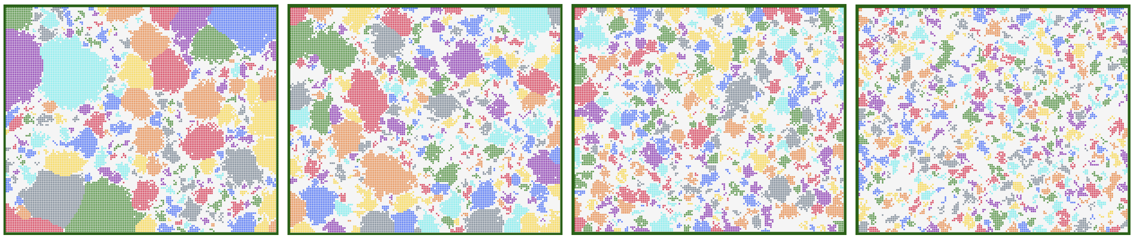


(a) 2 attributes, $\tau_u = 50\%$ (b) τ_u increased to 55% (c) τ_u increased to 60%

Figure 7: Illustration of steady states of 3-attribute model. $\tau_s = 100\%$ and population density is 80%.

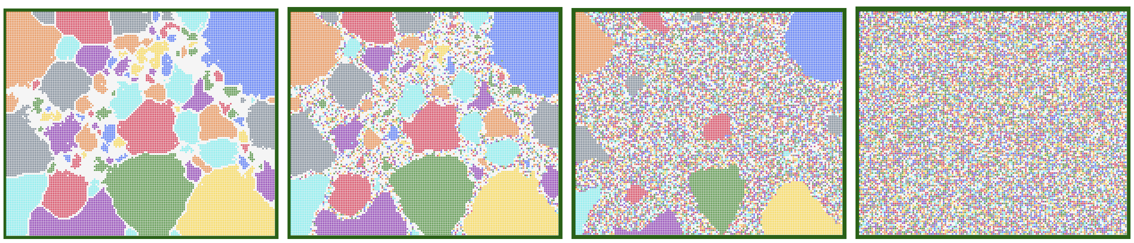
Similar to what was observed for Schelling's original model, in Figure 7, boundaries of the steady state of models with 2 attributes and 3 attribute becomes more clear as utility threshold τ_u increases, due to a limitation of neighbors one agent can have. When $\tau_u = 50\%$, color blocks of the converged model stay adjacent to each other, without any gap between them. With a slight increase of the threshold, an agent desires more to have no neighbors than having different neighbors. So a few agents adjacent to a different color block leaves and a gap gradually forms. When τ_u is high enough, say 60%, a clear white boundary can be observed between all color blocks. From the last section, in Figure 4, similar behaviors can also be observed from the 2-attribute models. Visually, with higher utility threshold, segregation becomes more severe.

With scarce population density, the model tends to converge into smaller color blocks (communities). Figure 9 demonstrates steady states of 3-attribute models when $\tau_u = 50\%$, $\tau_s = 100\%$ and population density ranges from 70% to 40%. We can still observe segregated communities, but as population density decreases, community blocks become smaller in size. A lower population density leaves more locations blank, where unsatisfied agents can stay with few or no neighbors and become satisfied. Thus more communities are formulated centering around those isolated agents. With more communities, each community becomes smaller in size. These communities absorb new members in parallel, accelerating the overall convergence speed.



(a) Population density 70%. (b) Population density 60%. (c) Population density 50%. (d) Population density 40%.

Figure 8: An illustrative case of 3-attribute model, where $\tau_u = 50\%$, $\tau_s = 100\%$ and population density changing form 70% to 40% with 10% interval. (a) to (d) shows that community forms into smaller blocks as population density decreases.



(a) (b) (c) (d)

Figure 9: An illustrative case of how communities disassemble if a high utility threshold τ_u is imposed. In this case, $\tau_u = 50\%$ at the beginning and is increased to 60%. $\tau_s = 100\%$ and population density is 100%. (a) to (d) shows that communities gradually disassembles.

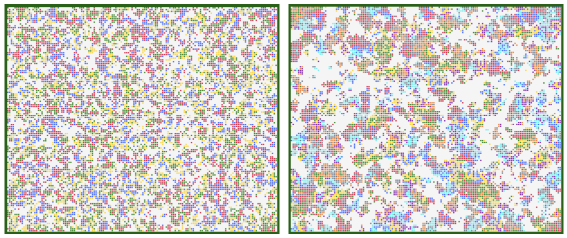
Interestingly, the model also interprets how a society turns from order into chaos. Consider a stable society where everyone is satisfied with its current neighborhood, as Figure 10a shows, where $\tau_u = 50\%$, $\tau_s = 100\%$, with population density 80%. Now, for some reason, agents become more demanding on their neighborhood and asks for more similar neighbors than before. Suddenly the utility threshold is raised to 60%. Firstly, agents at the edge of communities begins to move, because they are more likely to be closer to other classes of agents (Figure 10b). With a high population density of 80%, the moving agents are having a difficult time finding suitable neighborhoods and relocate adjacent to dissimilar agents, causing other agents to leave. Thus communities shrink as more and more agents start to relocate (Figure 9c). Finally no one is satisfied and the entire society turns into chaos (Figure 9).

Table 4: A comparison of segregation level of 2-attribute and 3-attribute models, where population density is 80%, $\tau_u = 50\%$ and τ_s changes.

2-attribute models			3-attribute model		
τ_s	Mean Iter.	CI Iter.	τ_s	Mean Iter.	CI Iter.
25%	25730.74	(25686.05, 25775.43)	25%	11962.64	(11935.25, 11990.03)
75%	91891.78	(91829.31, 91954.25)	50%	22134.58	(22099.77, 22169.39)
-	-	-	75%	104859.92	(103641.17, 106078.67)

However, things starts to change when we change similarity threshold τ_s for agents. Table 4 summarizes segregation levels of 2-attribute and 3-attribute models when $\tau_u = 50\%$, at 80% population density. We let $\tau_u = 50\%$ so that all models can converge into steady states and the segregation levels are comparable.

Obviously, as we decrease the similarity threshold, segregation level of converged societies drops significantly, for both 2-attribute and 3-attribute model. This suggests that agents are more open to neighbors that are not identical to them.



(a) 2-attribute model (b) 3-attribute model

Figure 10: A demonstration of non-segregated societies. 2-attribute and 3-attribute model is shown in (a) and (b), respectively. For 2-attribute model $\tau_u = 80\%$, $\tau_s = 25\%$ and population density is 60%. For 3-attribute model $\tau_u = 70\%$, $\tau_s = 50\%$ and population density is 60%.

Figure 10 demonstrates non-segregated societies for 2-attribute and 3-attribute model. A fixed seed was used to generate this figure. For 2-attribute model, $\tau_u = 80\%$, $\tau_s = 25\%$ and population density is 60%; for 3-attribute model, $\tau_u = 70\%$, $\tau_s = 50\%$ and population density is 60%. High utility thresholds were used to guarantee tighter formation of smaller communities. The figures show that although agents still form into communities, each community is no longer composed of same classes of agents: agents with various colors appear in the same community and all are satisfied with the current neighborhood. The reason behind a non-segregated community pattern and low segregation level resembles that of a faster convergence rate for low similarity threshold models. As we discussed before, an expanded agent pool enlarges the choices of an agent may have on its neighbors. Agents are no longer asking for the exact same type of neighbors as itself, rather, it accepts other types of agent, as long as them have some attributes in common.

4 CONCLUSION AND DISCUSSION

In this paper, we have revisited Thomas Schelling's social segregation model and extend it by replacing the original uni-dimensional agents with multi-dimensional agents. Two types of threshold are defined for the multi-dimensional agents models. The similarity threshold represents the percentage of attributes two agents should have in common to make them similar and the utility threshold represents the percentage of similar neighbors an agent desire to have in its neighborhood.

As a special case of the proposed extended model, the results of Schelling's original model are similar to what was reported in Schelling (1969) and Schelling (1971), where segregated communities are observed in the stable stage, with a moderate threshold. Further, we investigate the impact of the threshold to the model's convergence behavior. A high threshold prolongs the time in which the model reaches the stable state; or even prevents it from converging. This implies that a society where individuals have high homophily preference is prone to become unstable, because most people in such society are "perfectionist", that even a small violation in the neighborhood demography structure can render them unhappy. With a high similarity threshold and a moderate utility threshold, the multi-attributes models converges to a stable state and showed similar segregated result as the original model. However, when we lower the similarity threshold, the model still converges but with non-segregated community patterns. Such results suggest that a society where people are viewed as multi-dimensional beings is more likely to become stable, even when individual homophily preferences are high, as long as people are happy to accept partial differences from their neighbors and focus more on common interests.

REFERENCES

Axelrod, R. 1997. "The Dissemination of Culture: A Model with Local Convergence and Global Polarization". *The Journal of Conflict Resolution* 41(2):203–226.

Clark, W. A. V. 1991. "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model". *Demography* 28(1):1–19.

Clark, W. A. V., and M. Fossett. 2008. "Understanding The Social Context of The Schelling Segregation Model". *Proceedings of the National Academy of Sciences of the United States of America* 105(11):4109–4114.

Gracia-Lázaro, C., L. F. Lafuerza, L. M. Floría, and Y. Moreno. 2009. "Residential Segregation and Cultural Dissemination: An Axelrod-Schelling Model". *Physical Review E* 80(4):046123.

Grauwin, S., F. Goffette-Nagot, and P. Jensen. 2012. "Dynamic Models of Residential Segregation: An Analytical Solution". *Journal of Public Economics* 96:124–141.

Hamming, R. W. 1950. "Error Detecting and Error Correcting Codes". *The Bell System Technical Journal* 29(2):147–160.

Neal, Z. P., and J. W. Neal. 2014. "The (In)compatibility of Diversity and Sense of Community". *American Journal of Community Psychology* 53:1–12.

Pancs, R., and N. J. Vriend. 2007. "Schelling's Spatial Proximity Model of Segregation Revisited". *Journal of Public Economics* 91:1–24.

Paolillo, R., and J. Lorenz. 2018. "How different homophily preferences mitigate and spur ethnic and value segregation: Schelling's model extended". arXiv:1809.01437 [physics.soc-ph].

Schelling, T. C. 1969. "Models of Segregation". *The American Economic Review* 59(2):488–493.

Schelling, T. C. 1971. "Dynamic Models of Segregation". *Journal of mathematical Sociology* 1:143–186.

Spaiser, V., P. Hedstr´om, S. Ranganathan, K. Jansson, M. K. Nordvik, and D. J. T. Sumpter. 2018. "Identifying Complex Dynamics in Social Systems: A New Methodological Approach Applied to Study School Segregation". *Sociological Methods & Research* 47(2):103–135.

Stivala, A., G. Robins, Y. Kashima, and M. Kirley. 2016. "Diversity and Community Can Coexist". *American Journal of Community Psychology* 57:243–254.

AUTHOR BIOGRAPHIES

ZEYU LIU is a Ph.D. student in the Department of Industrial and Systems Engineering at the University of Tennessee, Knoxville. He received his Bachelor's Degree at Southeast University in Nanjing, China. He currently focuses on simulation and modeling of complex systems, optimization, heuristics algorithms and Markov decision processes on healthcare problems. His email address is zeyu.liu@utk.edu.

XUEPING LI is a Professor in the Department of Industrial and Systems Engineering at the University of Tennessee, Knoxville. Dr. Li is the Co-Director of the Health Innovation Technology & Simulation (HITS) Lab (https://hits.utk.edu) and the Director of the Ideation Laboratory (iLab). His research interests include complex systems modeling, simulation and optimization, health information technology and mobile health, data analytics, and smart and connected systems. His research has been sponsored by the National Science Foundation (NSF), National Institutes of Health (NIH), Department of Energy (DOE), Health Resources and Services Administration (HRSA), and so on. He is currently the President of the Modeling & Simulation Division of the Institute of Industrial and Systems Engineering. His email address is xli27@utk.edu. His personal website is http://web.utk.edu/~xli27/.

ANAHITA KHOJANDI is an Assistant Professor in the Department of Industrial and Systems Engineering at the University of Tennessee, Knoxville. She received her Ph.D. in Industrial Engineering from University of Pittsburgh. Her research interests include decision making under uncertainty, data mining and machine learning with applications in a wide array of complex systems. She is a member of INFORMS, IISE, IEEE, and Society for Medical Decision Making. Her email address is khojandi@utk.edu.

SANJA LAZAROVA-MOLNA is an Associate Professor with the Faculty of Engineering at the University of Southern Denmark. Her current research interests include modeling and simulation of stochastic systems, reliability modeling and analysis, and data analytics for decision support. Since 2019, she is serving on the Board of Directors of The Society for Modeling & Simulation International (SCS). Sanja Lazarova-Molnar obtained her PhD in Computer Science from the University in Magdeburg, Germany, in 2005, specializing in the area of Modeling and Simulation, where she was also a member of the Simulation Group at the Institute of Simulation and Graphics. She has successfully collaborated with industry and academia, authoring more than 80 peer-reviewed research publications. Her email address is slmo@mmmi.sdu.dk.