Emergence of Frontiers in Networked Schelling Segregationist Models

Philippe Collard

Nice Sophia Antipolis University
I3S Laboratory
Parc Valrose, 06108 Nice, France
and
Institut des Systèmes Complexes—Paris Ile-de-France
57-59 rue Lhomond, 75005 Paris, France
philippe.collard@unice.fr

Salma Mesmoudi

Institut des Systèmes Complexes—Paris Ile-de-France
57-59 rue Lhomond, 75005 Paris, France
salma.mesmoudi@iscpif.fr

Teodor Ghetiu
Fiona Polack

York Centre for Complex Systems Analysis
University of York, UK
teodorg@cs.york.ac.uk, fiona@cs.york.ac.uk

The relation between individuality and aggregation is an important topic in complex systems sciences, both aspects being facets of emergence. This topic has frequently been addressed by adopting a classical, individual versus population level perspective. Here, however, the frontiers that emerge in segregated communities are the focus; segregation is synonymous with the existence of frontiers that delineate and interface aggregates. A generic agent-based model is defined, with which we simulate communities located on grid and scale-free networked environments. Emerging frontiers are analyzed in terms of their relative occupancy, porosity, and permeability. Results emphasize that the frontier is highly sensitive to the topology of the environment, not only to the agent tolerance. These relations are clarified while addressing the topics of frontier robustness and the trade-off between its capacity to separate and allow exchange.

1. Introduction

The general context of this paper is formulated by Thomas C. Schelling himself [1]: "The […] subject that occupied me in the seventies was the ways that individual behavioral choices could aggregate into social phenomena that were unintended or unexpected. One

https://doi.org/10.25088/ComplexSystems.22.1.35

part of this work involved modeling spatial 'segregation,' the ways that people who differ conspicuously in binary groups [...] get separated spatially, in residence, in dining halls, at public events." In Schelling's segregationist model, there are two types of individuals and a *tolerance* level that denotes the threshold under which an agent is satisfied according to the type of its neighbors. Spatial dynamics induced by the moves of unsatisfied agents may lead to the emergence of spatial segregationist aggregates.

Individuality and aggregation constitute an important topic in complex systems research, as they represent facets of *emergence*, a core concept deemed a "central and constructive player in our understanding of the natural world" [2]. In order to recognize the unity of an aggregation of individuals, the interior of the aggregation must be distinguishable from the exterior. This is what defines the concept of *frontier*: when several aggregates are formed, the frontier is the set of all locations that allow contact between the members of opposite types.

The primary function of a frontier is to separate; for example, a defense system is aimed at keeping enemies away. Absolute separation is, however, an ideal; in reality, this is complemented by exchange, for example, information flow. More often, an alternation of phases of opening and closure can be observed: porosity varies over time depending on the relations between the entities. Frontiers can thus be described in terms of their varying separation-exchange trade-off; in addition, they can be classified according to the topology that is being considered. The classical scenario sees frontiers as geographical boundaries between two contiguous territorial systems (e.g., river, mountain range), but they can also describe the separation that emerges from social relations expressed in a network (e.g., cultural frontier, linguistic boundary). In the first case, the frontier appears as lines or fronts, while in the second case, the representation as a line is inadequate and we have to find novel modes of representation.

While Schelling's model deals with segregation only, this paper focuses on the concept of frontier and its associated trade-off between separation and exchange. Efforts are directed toward observing and analyzing the types of frontiers that emerge from varying factors such as tolerance and topology of the neighborhood network. We introduce a novel, three-fold way of analyzing the frontier, based on what we define as its *occupancy*, *porosity*, and *permeability*. This paper is of high generality, contributing directly to complex systems research. We consider however that, given a domain-based description of agents and their relations, this paper can be of relevance to sociologists, ethologists, ecologists, and others.

The paper is structured as follows. In Section 2, we propose a generic model of spatial segregation that will serve as a framework

https://doi.org/10.25088/ComplexSystems.22.1.35

for our studies. Section 3 defines the frontier concept. In Section 4, we perform simulations on a grid network. Section 5 is devoted to scale-free networks (SFNs). Finally, we offer our conclusions and hints for future research in Section 6.

2. Models of Segregation

Schelling's checkerboard model of residential segregation has become one of the most cited and studied models in many domains such as economics and sociology, for example, [3–7]. It is also one of the predecessors of agent-based computer models [8]. In this section, we first define a generic model of segregation (GMS); then we show that the standard Schelling model can be seen as an instance of the GMS.

Schelling's initial work details both a spatial and a frequency distribution model of segregation [9]. Starting from Schelling's initial insight that an individual's satisfaction depends on its tolerance and on the size (and social composition) of its neighborhood, a large number of subsequent studies have addressed topics concerning the spatial representation.

Moving beyond quantitative aspects such as the size of the neighborhood, the very type of the spatial structure has been analyzed. Fagiolo et al. [10], for example, study the underlying network in Schelling's segregation model. They examine models of agents interacting locally in a range of more general social network structures: regular lattice, random graph, and small-world and SFNs. Having a rather quantitative focus, their main result is that "the levels of segregation attained are in line with those reached in the lattice-based spatial proximity model." This contribution suggests that the spatial proximity model's explanation of segregation lies in the dynamic part of the model rather than the very rigid topological constraints used in the original model. In this paper, we take a rather qualitative point of view on the emergent structures; for instance, we look at the role of the hubs in a SFN according to the tolerance level. All this leads us to conclude that the proximity network has great influence on the frontier that emerges between opposite agents.

2.1 A Generic Model of Segregation

The generic model we define must be seen as a framework to show that a wide variety of frontiers can emerge despite the fact that they come from the same mold and share some common features.

2.1.1 Proximity Network

In the GMS we consider a set of agents in a world composed of *locations*. The *proximity network* defines the interconnectivity of locations. Two locations are neighbors if they are connected by an *edge* in the network; proximity equates to adjacency. The perception of an agent spans (only) the locations adjacent to it, constituting its local neighborhood.

Let V be the set of vacant locations and A the set of agent-occupied locations (or agent nodes). Let us note that there is one agent per occupied location. Agent density is the ratio $\delta = \frac{\sharp A}{\sharp A + \sharp V}$, where $\sharp V$ represents the number of vacant locations and $\sharp A$ the number of agent-occupied locations (one per location). We define the undirected proximity network N as the ordered pair

$$N = (A \bigcup V, E), \tag{1}$$

where $A \bigcup V$ is the set of all the location nodes and E is the set of edges between neighboring nodes. In a network, the *degree* of a node is the number of edges that connect it to other nodes; in the GMS, we call it the *proximity degree* (pd) of a node. According to their pd, networks can be homogeneous, for example, grid networks with periodic boundary conditions where $pd = 8$, or inhomogeneous network with specific $pd(l)$ to each location. (A grid network with periodic boundary conditions wraps horizontally and vertically.)

2.1.2 Agent Network

Overlying the proximity network is the agent network. Its nodes represent the agents within the world, whereas the connectivity builds on the layout of the proximity network. A simplifying assumption we make is that the agent network does not have explicit edges, but uses the proximity network ones. This way, two agents are in direct contact if they are located on adjacent locations; a proximity network edge can be seen as a shared communication channel.

We call *agent degree* ($ad_i(t)$) the number of connections an agent node a_i has to neighboring agent nodes at time t; it is a measure of the influence of the agent in the network. The proximity degree of the node is the maximum number of agents to which it can potentially be connected, that is, $\forall\, a_i \in A$, $(ad)_i \leq pd(l_i)$, where l_i represents the location in which agent a_i resides.

Apart from their location, agents are described by a *type* attribute, which remains constant. For convenience, we consider the type as the agent's color, which can be either blue or green. Let B (resp. G) be the set of agents in the blue type (resp. green type).

The satisfaction of an agent depends on its own type and on the type of its neighbors. For an agent a_i at time t, $o_i(t)$ (resp. $s_i(t)$) represents the number of neighbors with the opposite type (resp. same type). The value $O_i(t)$ (equation (2)) measures the ratio of the number of neighbors of opposite type to the agent degree.

$$O_i(t) = \begin{cases} \dfrac{o_i(t)}{ad_i(t)} & \text{if } ad_i(t) \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

We assume that the tolerance τ is a constant number in the range [0.. 1]; τ denotes the threshold under which an agent a_i is satisfied according to the value O_i. More precisely, we define the Boolean indicator *satisfied* as, for each agent a_i at time t:

$$satisfied_i(t) = (O_i(t) \leq \tau). \tag{3}$$

For instance, a tolerance τ of 0.5 means that each agent accepts at most half of its neighbors to be different from itself.

2.1.3 Micromotive versus Macrobehavior

Agent behavior is oriented on achieving and maintaining satisfaction (equation (3)): an unsatisfied agent is motivated to move toward another location, whereas a satisfied one has no incentive to move. To find a new place, an unsatisfied agent uses a simple rule (what we call the *eulogy to fleeing* (EF) rule): a location is randomly chosen from the world and the agent moves into it if and only if the location is vacant [11]. This rule is in the spirit of the complex systems paradigm, because agents do not need access to global (complete) information in order to make decisions. Consequently, they may move at random toward new locations by allowing utility-increasing or utility-decreasing moves. As the moves do not equate to immediate benefits, it is challenging to predict the overall emerging effect. Over time, movements generate new satisfied or unsatisfied agents through a chain reaction, until an equilibrium may be reached. Equilibrium denotes here a situation where the system (of agents) does not evolve anymore. At a time t, if all the agents are satisfied, the EF rule has no effect and then such a configuration is a *fixed point*. In this paper, we do not discuss the conditions that guarantee that the system converges toward equilibrium; we select system conditions in which equilibrium is reached. The EF rule has already been used within Schelling's model, leading the system toward equilibrium; furthermore, results confirm the paradoxical micro-to-macro link where a high level of tolerance can nonetheless induce a significant level of segregation [11].

2.1.4 An Aggregate Index to Measure Segregation

To have some insight into the aggregation level, it is necessary to measure the global state of aggregation of the world. We reformulate measures proposed by Schelling [12], Carrington [13], and Goffette-Nagot [14]. First, we define a global measure of *similarity* as:

$$s(t) = \frac{1}{\sharp A} \sum_{i}^{\sharp A} (1 - O_i(t)). \tag{4}$$

We then define the *aggregate* index by:

$$\text{aggregateIndex} = \begin{cases} \dfrac{s - s_{\text{rand}}}{1 - s_{\text{rand}}} & \text{if } s \geq s_{\text{rand}} \\ \dfrac{s - s_{\text{rand}}}{s_{\text{rand}}} & \text{else} \end{cases} \tag{5}$$

where s_{rand} is the expected value of the measure s implied by a random allocation of the agents in the world. A zero aggregate index corresponds to a random positioning of the agents. The maximum value of 1 corresponds to a configuration with two homogeneous patterns (complete segregation into two same-color groups), whereas negative values point toward highly mixed populations. (The aggregate index metric is closely related to the dissimilarity index used in the demographic literature [15].)

2.2 Schelling's Model of Segregation

Schelling's model can be viewed as a particular case for the GMS, where the proximity network is a two-dimensional grid and the neighborhood of an agent is composed of the eight nearest cells surrounding it ($pd = 8$). For instance, if $\tau = 0.37$, agents are rather intolerant; if moreover an agent has exactly eight neighbors, it cannot suffer more than two opposite neighbors. Similarly, for the value $\tau = 0.63$, all the individuals are rather tolerant and if moreover an agent has exactly eight neighbors, it can suffer at most five opposite agents in its vicinity.

Because of the grid layout with constant $pd = 8$, $O_i(t)$ can take only 21 values; these represent a nonuniform sample for the range [0.. 1], where values around 0.5 are the least represented. Any other value for the tolerance in the range [0.. 1] can be used, but it is equivalent to the closest inferior meaningful value.

3. Frontier

A frontier is a generic concept that has different instantiations depending on the context in which it is considered. A common class of frontiers is found in the geographical domain, where they appear as

https://doi.org/10.25088/ComplexSystems.22.1.35

fronts, while for social frontiers the representation as a "line" is reductionist and we have to think about new shapes. The interest we take in the concept of frontier lies on two aspects: from a static standpoint, a frontier enables the separation of incompatibilities; however, as it allows at the same time some form of communication between them, a dynamic perspective is also relevant. We consider a frontier as a structure which both determines the "borderland" between two aggregates of opposite types and allows communication between them.

The emergence of frontiers in Schelling systems has been studied in [16]: the authors have introduced a variant of the model in which agents are allowed to leave or enter the "city." Within this *open system*, both types of agents coexist in the city, but the dynamics lead to a segregation into clusters with a variety of frontiers between them. The authors classify these interfaces according to two features: their type (two opposite agents may be in contact or separated by vacancies), and their shape, which can be rugged or smooth. Their most important conclusion is that vacancies have a functional role, allowing weakly tolerant agents to be satisfied. As the authors say: "This is not the case in Schelling's original model where the vacant places are only 'conveyor of moves'."

Our study differs from this effort through the following aspects: (*i*) as in Schelling's original model, we consider a *closed system* with a fixed number of agents moving on a network; (*ii*) we look at frontiers on both a grid and a SFN; and (*iii*) we do not take a physicist's point of view, that is, the existence of a correspondence between Schelling's segregation model and spin-1 models.

3.1 Definition

A frontier is composed of the locations (nodes in the proximity network) where contact occurs between two agents of opposite types.

We consider contacts as being of two types: direct or indirect. In the GMS, a direct contact refers to agents being directly linked in the agent network (through one edge), whereas an indirect contact is mediated through a vacant location. In the real world, a direct contact can be exemplified through the contact of a healthy person with a person having a communicable disease, whereas an indirect contact is achieved through some intervening medium, for example, air.

Let D (resp. I) be the set of direct (resp. indirect) contacts:

$$D = \{(a_i, a_j) \in B \times G \mid a_j \in N(a_i)\}$$
$$I = \{(a_i, a_j, v) \in (B \times G \times V) \mid v \in N(a_i) \cap N(a_j)\}$$

where $N(a_k)$ is the neighborhood of the agent a_k. (Note that a pair of agents may be invoked both in D and I. A pair of agents may be invoked many times in I.) Then we define the frontier F of the network N as a subnetwork of N composed of all nodes that are in contact:

$$F = (A_F \bigcup V_F, E_F) \tag{6}$$

where A_F is the subset of agent nodes that are at least one coordinate of an element of D or I, V_F is the subset of vacant locations that are at least one coordinate of an element of I, and E_F is the set of links between neighboring nodes of F.

3.2 Characteristics of a Frontier

To enable consistent comparisons of frontiers, we must take into account their importance relative to the entire world, their openness, and their ability to pass a signal from one side to the other. Thus, we define what we mean by a frontier's occupancy, porosity, and permeability. These three criteria are chosen to address the ambivalence between separation and exchange as the main characteristic of a frontier.

Occupancy. We define the occupancy of a frontier F as the ratio between the number of locations forming the frontier and the total number of locations in the world:

$$o(F) = \frac{\sharp (A_F \bigcup V_F)}{\sharp A + \sharp V}. \tag{7}$$

For instance, if each agent is placed on a checkerboard according to its type, all the agents are on the frontier and so the occupancy is equal to 1. (On a checkerboard, a green agent is on a black square and a blue agent is on a white square, so there are no vacant places.) The occupancy is related to the size of the frontier and measures to some extent the cost to build, to operate, or to maintain a frontier in order to separate communities, for example, building a wall to separate two countries or operating and maintaining paths linking the countries.

Porosity. In material physics, porosity is a measure of how much of a rock is open space in between spores or within cavities of the rock [17]: it is defined as the ratio of the occupancy of voids in a material to the occupancy of the whole. By analogy, we consider the elements of D as representing the voids in a material (a lack of communication impediments). We then define the porosity of a frontier as the proportion of direct contacts:

$$\pi(F) = \frac{\sharp D}{\sharp D + \sharp I}. \tag{8}$$

Permeability. In material physics, permeability is the ability of porous material to allow the passage of a fluid. By analogy we define the permeability of a frontier as a measure of ease with which a "signal" can cross the frontier. Consequently, permeability measures the availability of communication paths from one community of agents toward the opposite community. As this is related to a dynamic property, we cannot provide a formula for permeability. To measure this capacity of the frontier we conduct a percolation test that provides a number in [0.. 1].

3.3 Simulation Framework

Experiments are performed via the NetLogo multi-agent programmable modeling environment [18]. The pseudocode for simulating the GMS is defined in Algorithm 1; by instantiating the network N with a grid or a SFN, we obtain two different simulators that we experiment with.

Algorithm 1. Simulation of the GMS.

```
t ← 0, density ← δ, tolerance ← τ
create a network N and position at random the agents on it
for each agent a_i do
  initialize its satisfaction satisfied_i(0)
end for
while not (all the agents are satisfied) do
  for each agent a_i do
    if not (a_i satisfied) then
      choose a node location at random on N
      if the location is vacant then
      a_i moves to this location
      end if
    end if
  end for
  t ← t + 1
  for each agent a_i do
    if required then
      update satisfied_i (t)
    end if
  end for
end while
```

Ensure: all the agents are satisfied

To measure permeability, we assume that (*i*) all the agents are satisfied, and (*ii*) an agent has two possible states: informed and susceptible. (We use the term susceptible in reference to the SI epidemic

https://doi.org/10.25088/ComplexSystems.22.1.35

model [19], meaning "susceptible to become infected"—a susceptible agent is not infected, but has the potential of being so.) Only an informed agent can transmit the signal, whatever its nature, for example, a virus, a rumor, or an opinion. This happens as soon as it comes into direct contact with a susceptible agent, and transmission occurs irrespective of the type (color) of the agents. On the other hand, a vacant place constitutes an impervious barrier. To measure the capacity for one signal to cross the frontier between the two communities, we run a percolation test (Algorithm 2). First, we assume that all the agents are susceptible with the exception of a single blue agent who is informed. Then, during several iterations, the signal is locally propagated. The number T of iterations used in the algorithm must be large enough so that the signal can reach any nodes connected to the source by at least one path through direct contacts. At the end of the run, we compute the proportion of informed agents among green agents. As this proportion can be greatly influenced by the initial choice of the blue source, we repeat the test 3000 times with, for each run, a different random seed. Finally, the permeability of the frontier is the average of the proportion of informed green agents calculated over all the tests.

Algorithm 2. Percolation test.

Require: all the agents are susceptible

1. inform a single blue agent
2. **for** $i = 1$ to T **do**
3.   **for** each informed agent a_i **do**
4. a_i spreads its own information to all its direct neighbors
5.   **end for**
6. **end for**
7. **return** the proportion of informed agents among green agents

In the following we conduct experiments to establish correlation between the frontier, the topology of the network, and the level of tolerance.

4. Frontier on a Grid Network

4.1 Generic Model of Segregation and Grid Network

The world is a grid with periodic boundary conditions; as a consequence, the degree distribution in the proximity network is uniform, that is, $pd = 8$. Simulations are performed on an $L \times L$ lattice of locations, with L set to 100. The agent set is positioned in a random initial configuration, such that the vacant locations and the two types of agents are well mixed and the aggregate index is close to 0.

https://doi.org/10.25088/ComplexSystems.22.1.35

The frontier will be represented graphically by drawing on the grid all the contacts between the two communities. A direct contact is represented by its couple of opposites agents. For ease of reading, an indirect contact is represented by its vacant location only. An agent node in A_F with type blue (resp. green) is a circle colored in blue (resp. green) whereas a vacant node in V_F is a black square.

4.2 Experiments

4.2.1 From Low to High Tolerance

We conduct several simulations to show that dynamics lead to the emergence of different types of frontiers. The main parameter that controls the model is the tolerance (τ) of the agents. To ensure that the size of the largest frontier component is large enough to obtain convergence toward a state of complete satisfaction, we fix the density of agents to 90%. Figure 1 shows different kinds of frontiers for three levels of tolerance: 0.25, 0.50, and 0.80.

First we choose the smallest tolerance value ($\tau = 0.25$), which leads the system to converge toward equilibrium. Figures 1(a) and 1(b) represent the frontier at convergence for a representative run: we see that the dynamics lead to the emergence of spatial homogeneous patterns. The frontier is essentially built with vacant locations (black squares), so there are many indirect contacts and few direct contacts. Therefore homogeneous patterns are isolated by a *no man's land* of vacant nodes. Vacancies are used to allow the agents first to move, and eventually to isolate aggregates by neutral zones. The frontier can be viewed as some Great Wall of China.

As tolerance increases (Figures 1(c) and 1(d)), we observe that the no man's land shape becomes more and more complex: as in a real landscape when roughness dictates many meanders to the edge of a lake, the complexity of contours increases. Finally the frontier looks like a *Peano line* [20], that is, a space-filling curve whose range contains the two-dimensional square (Figures 1(e) and 1(f)). Qualitatively, we observe that both the occupancy and the porosity increase with tolerance.

4.2.2 Quantitative Results

We conduct more simulations to get quantitative results. All the results are averaged over 100 independent runs; they are presented in Figure 2 and summarized in Table 1. We observe that (*i*) the frontier gains in occupancy from 0.11 to 0.93 and the increase is quasi-linear (as soon as the tolerance is over 0.37); (*ii*) porosity increases from 0.20 to reach a high plateau (0.92) as soon as the tolerance is above 0.5; and (*iii*) permeability is always close to 1. (Note that the number

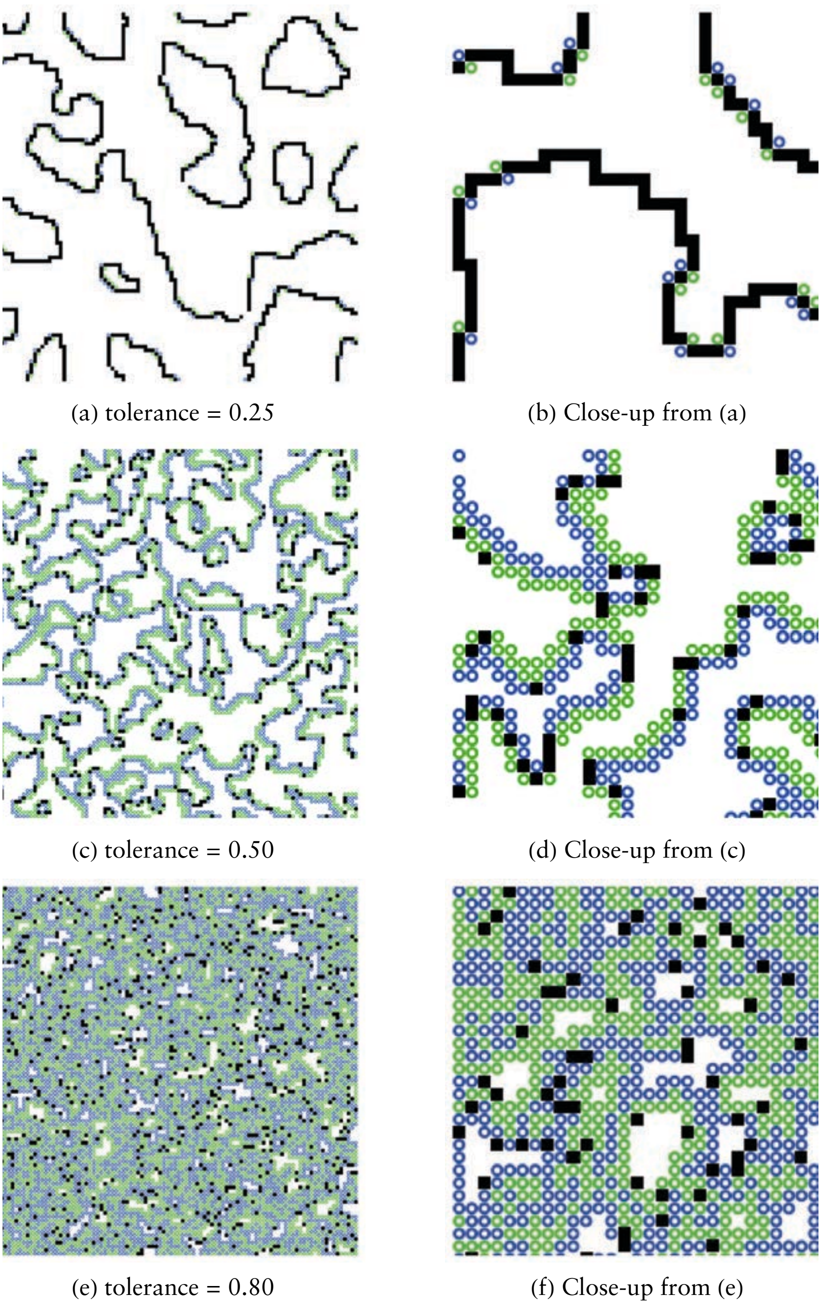


Figure 1. From no man's land (top) to Peano-like lines (bottom). Frontier on a grid network with periodic boundary conditions ($\delta = 90\,\%$).

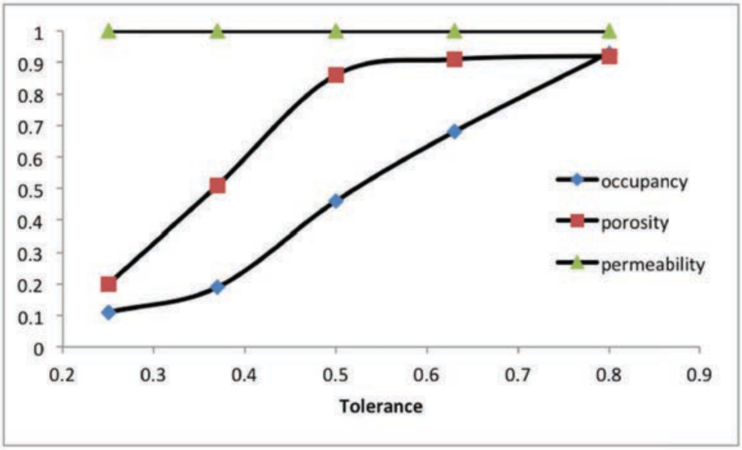


Figure 2. Occupancy, porosity, and permeability versus tolerance. Frontier on a grid network ($\delta = 90\,\%$).

Frontier	Grid τ=0.25 $\nearrow$0.8	Scale-Free τ=0 $\nearrow$0.8
Occupancy	0.11 $\nearrow$0.93	0.447 $\searrow$0.17
Porosity	0.20 $\nearrow$0.92	$\approx$0 $\nearrow$0.87
Permeability	$\approx$ 1	$\approx$0 $\nearrow$0.89
	(Figure 2)	(Figure 6)

Table 1. Synthesis of results with increasing tolerance (τ).

of iterations used in the percolation test [Algorithm 2] remains low compared to the size of the world. We use $T = 200$ for a 100×100 lattice.) So, whatever tolerance, a single blue signal may be propagated via the direct contacts to all the green agents. Especially in the case of low tolerance, although the porosity is low, there are enough direct contacts to allow propagation (Figure 1(b)). Finally, all this shows (*i*) there are no correlations between the permeability and the other characteristics of a frontier, and (*ii*) for intolerant agents, the occupancy and the porosity are correlated.

5. Frontier on a Scale-Free Network

In Section 4 we made the assumption that the proximity network is defined in such a way that every node has the same number of neighbors. This disregards many results showing that the majority of real-

https://doi.org/10.25088/ComplexSystems.22.1.35

world networks do not share this feature. For instance, scale-free or fractal urban networks share more similarities with real cities than grid networks [21]; likewise, some social networks are also scale-free [22].

5.1 Generic Model of Segregation and Scale-Free Proximity Network

In this section, the proximity network is a SFN [23, 24], a structure that is ubiquitous both across natural systems (e.g., brain, cell, social network, and ecosystem) and engineered ones, for example, the internet. In a SFN, the degree distribution $P(k)$ follows a *power law*; that is, the fraction $P(k)$ of nodes in the network having k connections to other nodes goes for large values of k as

$$P(k) \approx k^{-\gamma} \tag{9}$$

where γ is a parameter whose value is typically in the range [2.. 3]. As a consequence, there are some hubs that are highly connected and many nodes that are slightly connected. Another important characteristic is that a SFN can be generated by a random process called *preferential attachment* (PA) [23]; this process simply explains the idea that the rich get richer: in other words, nodes gain new connections in proportion to how many they already have.

A node is either vacant or contains an agent (with blue or green type). The satisfaction of an agent depends to a large extent on its degree. For example, a leaf node will change its point of view as soon as its neighbor moves away, whereas an agent localized on a hub will probably be insensible to the departure of one of its numerous neighbors. Thus some nodes tend to be breakable, while others tend to be robust.

Previous work on "network effects in Schelling's model of segregation" [21] provides new evidence from agent-based simulations; their aim is to obtain neighboring graphs more or less marked by the existence of cliques, with the authors assuming that the number of neighbors is fixed both for grid and SFNs. To achieve this, each node is connected to its n closest nodes. Consequently, our paper differs deeply from this approach as the proximity degree of a node remains free and depends on the PA rule only.

5.1.1 Beyond the Aggregate Index Measurement

To accurately measure segregation in a SFN, in addition to the aggregate index, we propose measuring the ability of a hub to aggregate agents of the same type. First, for each hub h we define a local measure of segregation $\text{seg}_h(t)$ as the maximum proportion of edge connections to agents of the same type it has: a high value, close to one, means the hub is the center of a cluster where agent nodes share a common type. Then, in order to get a global measure, we compute the mean of this local measure over the set H of hubs

$$\text{segHub}(t) = \frac{1}{\sharp H} \sum_{h}^{s} \text{seg}_h(t). \tag{10}$$

Initially, segHub(0) is close to 0.5.

5.1.2 Experiments

Each simulation is based on a network built by PA: nodes are added one after another in such a way that the degree distributions for the two types of agents, blue or green, are comparable. Then, for each agent, we compute its initial satisfaction according to the tolerance and its neighborhood.

In the following, we conduct experiments using the GMS based on SFNs. Qualitative results are shown on the network itself: a circle is an agent node and a black square is a vacant node. A full circle is a satisfied blue agent while a hollow circle is a satisfied green agent. For ease of reading, the size of a node is proportional to its degree (Figure 3). The frontier is a subnetwork (equation (6)); it is represented as such by drawing all the contacts between the two communities.

To obtain a readable representation we must be willing to accept a relatively small size, thus we use SFNs composed by 1000 nodes. The exponent γ is close to 2.3 (see equation (9)). To give an order of magnitude on the degree distribution, we indicate that initially 80% of nodes have a degree less than or equal to 3. We get a distribution with a long tail that extends up to high levels until it reaches a degree of around 30 for the more connected nodes. A node is considered a hub as soon as its degree is over some threshold; in order to consider only "true" hubs, we fix this threshold to 10, so approximately 1% of the nodes are seen as hubs. We compute the mean, over each type of node, for the degree; in all cases—blue agents, green agents, and vacant locations—this indicator is beyond 2. Because the PA process is stochastic, quantitative results are averaged over 100 independent runs.

5.2 From Zero to High Tolerance

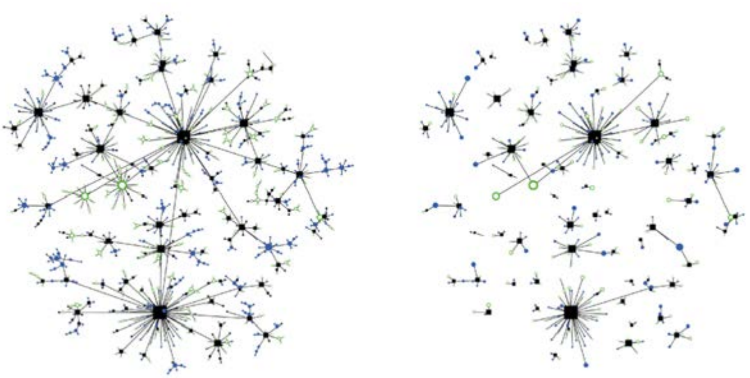
To get preliminary results about the emergence of macrobehaviors within a SFN, we consider three cases: zero tolerance, intolerant, and tolerant agents. In all cases we look at the frontier only when all the agents are satisfied (equilibrium state). To obtain convergence toward a state of complete satisfaction, it is necessary to have at least 15% of free places; that is why we chose a value of 80% for the density of agents. Remember that on a grid convergence occurs even if density is 90%. To better compare the different dynamics, we use the same initial representative network for all the qualitative results.

5.2.1 Zero Tolerance

In this extreme case ($\tau = 0$), an agent will not tolerate any neighbor of the opposite type. Therefore, initially most agents are unsatisfied. Remember that on a grid a tolerance of at least 0.25 is required for convergence. Conversely, even in this extreme case, a SFN is able to reorganize itself in order to converge toward a state where all the agents are satisfied. This is possibly made by a spatial rearrangement where the vacant nodes occupy the vast majority of highly connected nodes. The analysis of the evolution of the mean degree confirms these observations: from a value close to 2, it increases to 4.3 for the vacant nodes and decreases to 1.5 for the sets of agents of each type.

The aggregate index, close to 0 for the initial random configuration, reaches at the end of the run its maximum value of 1. This reveals a situation of extreme segregation where there is no contact between opposite agents. At convergence, a vacant hub can isolate all its neighbors, and so it makes the presence of antagonist agents consistent among its direct neighbors. A value of around 0.49 for segHub shows that there is no agent type drift.

Figure 3(a) represents the network at the end of a representative run when all the agents are satisfied ($t = 361$). We observe that vacant places have become the backbone of the network as the vast majority of hubs are vacant nodes.



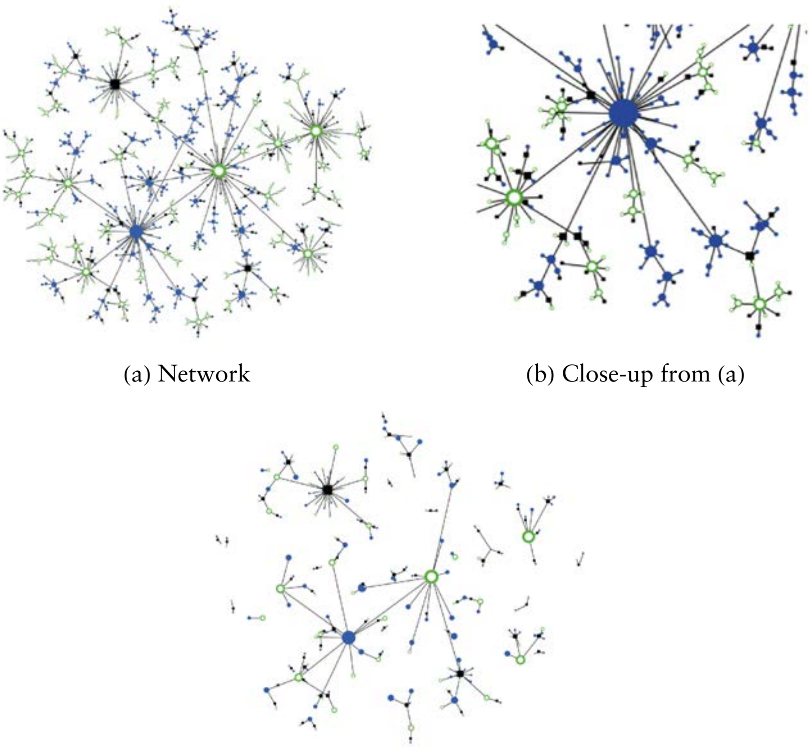
(a) Network (b) Frontier

Figure 3. Equilibrium state ($t = 361$): aggregationIndex $= 1.0$, segHub $= 0.49$, occupancy $= 0.456$, and porosity $= 0$. SFN with zero tolerance: $\tau = 0$ ($\delta = 80\,\%$).

From the network we have extracted the frontier (Figure 3(b)); a large variety of components can be seen from largest to smallest. The occupancy of the frontier is relatively important as it represents about half of the total space ($o(F) = 0.456$); this is because it contains many indirect contacts via the hubs. As the frontier is built only from indirect contacts between opposite agents connected via a vacant node, the porosity is zero ($\pi(F) = 0$). As a consequence, the permeability is zero ($p(F) = 0$).

5.2.2 Intolerant Agents

We assume that agents are rather intolerant ($\tau = 0.37$). Figures 4(a) and 4(b) represent the network at convergence after 136 time steps. We can observe that many agents have moved on the hubs. Segregation is still very important since the aggregation index reaches 0.97.



(a) Network

(b) Close-up from (a)

(c) Frontier

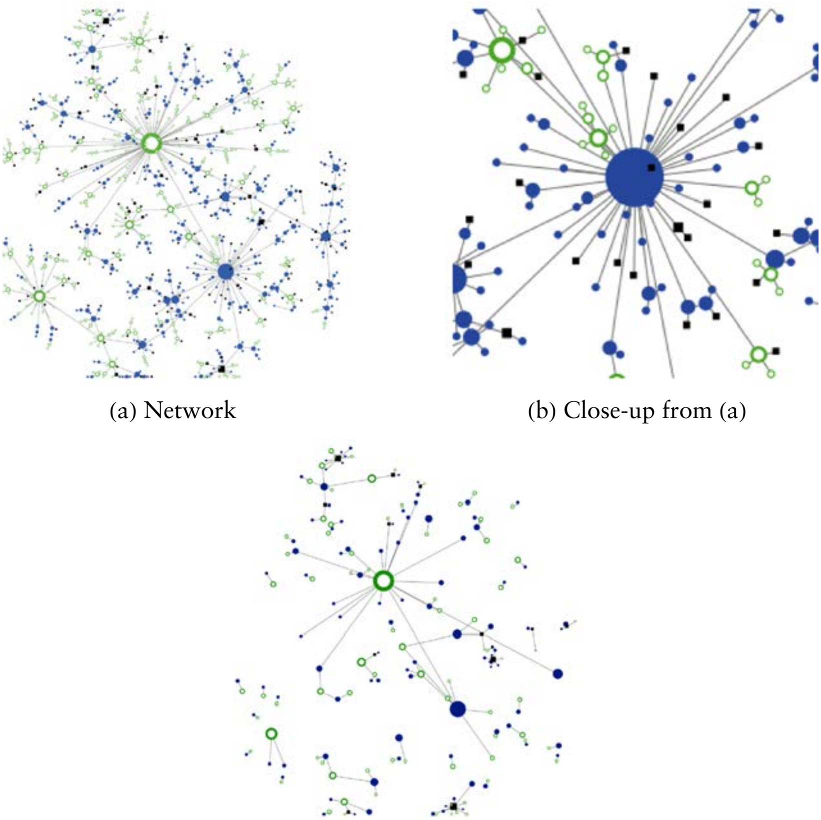
Figure 4. Equilibrium state ($t = 136$): aggregationIndex = 0.97, segHub = 0.67, occupancy = 0.25, and porosity = 0.21. SFN with intolerant agents: $\tau = 0.37$ ($\delta = 80\,\%$).

https://doi.org/10.25088/ComplexSystems.22.1.35

From the network we have extracted the frontier (Figure 4(c)); a large variety of components can be seen from largest to smallest. The occupancy is lower ($o(F) = 0.25$) and the porosity is higher ($\pi(F) = 0.21$) than the ones obtained with zero tolerance; the frontier is now constituted by both direct contacts and indirect contacts. As there are agents on some hubs, the segHub measures how much the hubs promote segregation; for the hubs, 67% of the neighbors belong to the same type (Figure 4(b)).

5.2.3 Tolerant Agents

We assume that agents are rather tolerant ($\tau = 0.63$). It might be tempting to predict that from the "sum" of these individual behaviors will emerge a global configuration where mixing is the rule. Figure 5



(a) Network

(b) Close-up from (a)

(c) Frontier

Figure 5. Equilibrium state ($t = 54$): aggregationIndex $= 0.88$, segHub $= 0.75$, occupancy $= 0.19$, and porosity $= 0.74$. SFN with tolerant agents: $\tau = 0.63$ ($\delta = 80\,\%$).

shows that in such circumstances the system converges quickly in less than 60 time steps. Vacant nodes mainly occupy the leaves of the network. As they are needless in such positions, we can assume that there are supernumerary vacant nodes; this is confirmed by the fact that the system converges with 5% of vacant places only ($\delta = 95\,\%$). Comparison of the results with those obtained with intolerant agents shows that the occupancy is lower ($o(F) = 0.19$) and the porosity is very high ($\pi(F) = 0.74$): the frontier is now largely made up of direct contacts.

Although each agent has a tolerant behavior, a high index of aggregation (0.88) reveals a strong segregation. Moreover, the segHub ratio is very meaningful: for the agent hubs, there are on average 75% of neighboring agents with the same type (Figure 5(b)). Coming from tolerant agents, such global behavior is unexpected. This macrobehavior magnifies results obtained by Schelling on a grid network; when groups of rather tolerant agents organize themselves on a spatial SFN, we may obtain a strong segregation as soon as all the agents become satisfied. Once again, this is an example of the gap that may exist within a complex system between the micromotives and the macrobehavior [25].

5.3 Quantitative Results

We conduct more simulations to get quantitative results on the influence of the tolerance. All the results are averaged over 100 independent runs; they are presented in Figure 6 and summarized in Table 1.

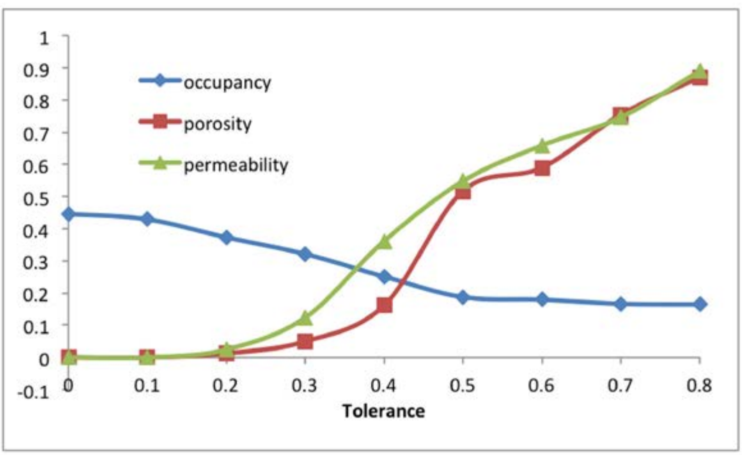


Figure 6. Occupancy, porosity, and permeability versus tolerance. Frontier on a SFN ($\delta = 80\,\%$).

https://doi.org/10.25088/ComplexSystems.22.1.35

As tolerance increases from 0 to 0.8, we can observe that (*i*) the occupancy decreases from 0.447 to reach a plateau (0.17) as soon as the agents become tolerant ($\tau > 0.5$). (*ii*) At the same time, the porosity increases from 0 to high values. Roughly, the porosity follows a logistic growth curve: first there are no direct contacts, then it increases exponentially until growth saturates, and finally the curve levels off at $\pi = 0.87$. (*iii*) The permeability roughly follows the evolution of the porosity; it increases with tolerance from 0 to 0.89. (As in a SFN, the average number of edges between any two vertices is small; the number N of iterations needed in Algorithm 2 for a signal to spread between any connected pair remains low compared to the size of the network. Here we use $N = 30$.) We can observe two phases: below 0.3, there is a weak permeability, but conversely, above 0.6, many green agents are informed. Between these values, there is a smooth transition from the greater tightness to the greater permeability. As a consequence, the effect of increasing tolerance on the ability to propagate a signal is drastic. This can be explained by the degree of the agent nodes in the frontier, which deeply changes according to the tolerance level.

All these results confirm that the nature of the frontier changes as agents become more and more tolerant: gradually, the hubs initially occupied by vacancies are replaced by agents.

6. Summary and Discussion

In this paper we first define a generic agent-based model of segregation. To stay in the spirit of complex systems, the agents follow a simple eulogy to fleeing (EF) rule: unsatisfied agents do not seek immediate benefits, nor do they need absolute knowledge of their environment. Simulations have shown that through using the simple EF rule, global segregation can be observed despite a high level of local tolerance. This confirms unexpected effects observed with Schelling's model on a grid network and extends these results to a scale-free network (SFN), all pointing toward the gap that may exist within a complex system between micromotives and macrobehavior.

The main topic of this paper is the emergence of frontiers between two communities. Considering the geographical and the social cases, we have defined a frontier as the subnetwork of locations where contacts occur between opposite agents. For a frontier, we have defined its occupancy, porosity, and permeability.

6.1 Frontier on a Grid Network

If density is high enough, intolerant agents give rise to a short no man's land frontier, essentially built of vacant locations. At the opposite end, if agents are tolerant, the borderland is essentially built with couples of opposite agents and the frontier resembles a Peano line. This means that both the shape and the composition of the frontier change as agents become more tolerant. As tolerance increases, the frontier gains in occupancy, porosity increases to reach a high plateau, and permeability is always at its maximum. There is no obvious correlation between the permeability and the other characteristics of a frontier.

6.2 Frontier on a Scale-Free Network

If density is high enough, as tolerance increases, the occupancy decreases and both porosity and permeability increase from 0 to a high value. As a consequence, the effect of increasing tolerance on the ability to disseminate information is drastic.

A remarkable result is that even in the case of zero tolerance, the agent network is able to reorganize itself to ensure everyone's satisfaction. In this extreme case, the vacant places become the backbone of the frontier; the vast majority of hubs are vacant. Whereas porosity and permeability are zero, the occupancy reaches its maximum value relatively to any other values for the tolerance. In addition, while there is an extreme global segregation, locally in agent hubs there is no agent type drift.

6.3 Frontier and Network's Topology

According to the topology of the underlying network, we have observed significant differences.

1. If tolerance is zero, there is convergence (resp. no convergence) in a SFN (resp. grid).
2. As tolerance increases, the occupancy decreases (resp. increases) in a SFN (resp. grid).
3. As tolerance increases, the evolution of the porosity and the permeability are correlated (resp. uncorrelated) in a SFN (resp. grid).
4. As tolerance increases, the permeability increases (resp. remains equal to 1) in a SFN (resp. grid).

6.4 Robustness of the Frontier against Attack

According to its ability to favor exchange between the two communities, the frontier may be robust or fragile against attack. This depends on its own structure and on the type of attack, random or intentional. As exchanges pass the frontier through its agents, we assume that attacks reach agents only.

Grid network. Low tolerance leads to a no man's land frontier where direct contacts are in the minority ($\pi = 0.20$); with regard to permeability, the frontier is robust to random damage but vulnerable

https://doi.org/10.25088/ComplexSystems.22.1.35

to malicious attacks targeted against an agent involved in direct contacts. Conversely, high tolerance leads to a frontier like a Peano line where direct contacts are plentiful ($\pi = 0.92$); whether the attack is random or intentional, the frontier is robust.

Scale-free network. If tolerance is zero, the vast majority of hubs are vacant nodes and there are no direct contacts ($\pi = 0$), so attacks have no effect on the permeability. On the other hand, if agents are tolerant, a large majority of hubs belong in the frontier as agent nodes and are largely invoked in direct contacts ($\pi = 0.87$ and $\text{segHub} = 0.75$). If attacks are targeted against agent hubs, the frontier is vulnerable.

6.5 Separation versus Exchange

Occupancy (i.e., the size of a frontier) measures to some extent the cost of separating communities, whereas permeability measures the availability of communication paths between communities separated by a frontier. Is it possible to maximize exchange while minimizing the cost of separation?

Whatever the topology, the occupancy depends on the tolerance. Tuning the tolerance level allows control over the occupancy: low occupancy corresponds to low (resp. high) tolerance in a grid (resp. SFN).

The permeability depends both on the topology of the network and on the porosity. In a grid, high porosity is not necessary to allow exchange through the frontier, and indeed the permeability is always very high. In the opposite way, in a SFN porosity and permeability increase together with tolerance, so we can control the permeability by tuning the tolerance level.

As a consequence, whatever the topology, the permeability can be maximized while minimizing the occupancy. In a grid, the trade-off is obtained with low tolerance; the frontier is like a no man's land essentially built with vacant nodes with many indirect contacts. In a SFN, this is obtained with high tolerance, that is, when the frontier is largely made up of agent nodes with direct contacts. In both cases, the frontier is vulnerable to malicious attack targeted against direct contacts.

6.6 Discussion

The generic nature of Schelling's model makes its interpretation in the real domain a difficult task. We took a first step here by analyzing SFNs, which are more representative of real networks (e.g., social and biological). Another step is reflected by looking at communication aspects, that is, permeability, rather than simply measuring segregation levels. A system is complex also (mainly) due to the interaction of its components. In the real domain, communication is both direct and indirect (humans, animals, plants, information technology systems, etc.—they can all communicate) and in this study we have taken this aspect into account.

All the results show that the underlying topology plays a key role in influencing the dynamics of moves and the emergence of frontiers. In particular, results about permeability suggest that it is essential to study the influence of topology when trying to understand the functional properties of a frontier. This enlightens our understanding of the dual nature of a frontier that both separates different aggregates and allows some form of communication between them.

We conclude with a series of suggestions for future research. (*i*) For SFNs, we will establish if the same conclusions hold for different exponents in the power law and for a bigger network. (*ii*) It will be very interesting to study the emergence of frontiers with a model that interpolates between scale-free and Erdos–Rényi networks [26] to see how "scale-freeness" affects the occupancy, porosity, and permeability. (*iii*) Beyond the study of frontiers when all agents are satisfied, we will answer questions on how a frontier evolves over time and, ultimately, how it is built. (*iv*) In a SFN, when agents are intolerant, vacancies play a central role; whereas we have assumed that each vacant place is an impervious barrier, it will be interesting to introduce a nonzero probability to transmit information through a free place. This new parameter would allow accurate control of the permeability. (*v*) In today's world, where social interactions are no longer constrained by geography, it might be of great interest to make clear what is meant by binding together "where you live" and "who you know." This will be a track to consider together several types of networks derived from the generic model of segregation (GMS); this will allow the study of their interaction and their joint evolution toward complementary forms of frontiers.

Although this paper deals with artificial complex systems and is based on a multi-agent simulation, we hope the results will give rise to debate among sociologists, ethologists, geographers, and others.

Acknowledgments

We thank the anonymous referees for useful comments and suggestions. The source code to reproduce the results is available, on request, from the authors.

References

[1] T. C. Schelling. "Thomas C. Schelling - Autobiography." *Thomas C. Schelling - Autobiography*. (2005) http://www.nobelprize.org/nobel_prizes/economics/laureates/2005/schelling-autobio.html.

[2] M. A. Bedau, "Weak Emergence," in *Philosophical Perspectives: Mind, Causation, and World*, Vol. 11 (J. Tomberlin, ed.), Malden, MA: Blackwell Publishing, 1997 pp. 375–399.

[3] R. Pancs and N. Vriend, "Schelling's Spatial Proximity Model of Segregation Revisited," in *Proceedings of the 9th International Conference of Computing in Economics and Finances (CEF03)*, Seattle, WA (C. F. Baum, ed.), 2003.

[4] J. Zhang, "A Dynamic Model of Residential Segregation," *Journal of Mathematical Sociology*, **28**, 2004 pp. 147–170.

[5] S. Gerhold, L. Glebsky, C. Schneider, H. Weiss, and B. Zimmermann, "Computing the Complexity for Schelling Segregation Models," *Communications in Nonlinear Science and Numerical Simulation*, **13**(10), 2008 pp. 2236–2245. doi:10.1016/j.cnsns.2007.04.023.

[6] L. Gauvin, J. Vannimenus, and J.-P. Nadal, "Phase Diagram of a Schelling Segregation Model," *The European Physical Journal B*, **70**(2), 2009 pp. 293–304. doi:10.1140/epjb/e2009-00234-0.

[7] A. Banos, "Exploring Network Effects in Schelling's Segregation Model," presentation given at *S4-Modus Workshop: Multi-scale Interactions between Urban Forms and Processes*, Besançon, France, 2009.

[8] J. B. Rosser, Jr., "On the Complexities of Complex Economic Dynamics," *Journal of Economic Perspectives*, **13**(4), 1999 pp. 169–192. doi:10.1257/jep.13.4.169.

[9] T. C. Schelling, "Models of Segregation," *The American Economic Review*, **59**(2), 1969 pp. 488–493. http://www.jstor.org/stable/1823701.

[10] G. Fagiolo, M. Valente, and N. J. Vriend, "Segregation in Networks," *Journal of Economic Behavior & Organization*, **64**(3–4), 2007 pp. 316–336. doi:10.1016/j.jebo.2006.09.003.

[11] P. Collard and S. Mesmoudi, "How to Prevent Intolerant Agents from High Segregation?," in *Advances in Artificial Life (ECAL 2011)*, Paris, France (T. Lenaerts et al., eds.), Cambridge, MA: MIT Press, 2011. http://hal.archives-ouvertes.fr/docs/00/74/28/59/PDF/ecal11.pdf.

[12] T. C. Schelling, "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, **1**(2), 1971 pp. 143–186. doi:10.1080/0022250X .1971.9989794.

[13] W. J. Carrington and K. R. Troske, "On Measuring Segregation in Samples with Small Units," *Journal of Business & Economic Statistics*, **15**(4), 1997 pp. 402–409. http://gatton.uky.edu/faculty/troske/publish_pap/jbes_oct97.pdf.

[14] F. Goffette-Nagot, P. Jensen, and S. Grauwin, "Dynamic Models of Residential Segregation: Brief Review, Analytical Resolution and Study of the Introduction of Coordination," *HAL-CCSD*, 2009. http://halshs.archives-ouvertes.fr/docs/00/40/44/00/PDF/0914.pdf.

[15] A. J. Laurie and N. K. Jaggi, "Role of 'Vision' in Neighborhood Racial Segregation: A Variant of the Schelling Segregation Model," *Urban Studies*, **40**(13), 2003 pp. 2687–2704. doi:10.1080/0042098032000146849.

[16] L. Gauvin, J. P. Nadal, and J. Vannimenus, "Schelling Segregation in an Open City: A Kinetically Constrained Blume-Emery-Griffiths Spin-1 System," *Physical Review E*, **81**(6), 2010 p. 066120. doi:10.1103/PhysRevE.81.066120.

[17] J. Sperl and J. Trckova, "Permeability and Porosity of Rocks and Their Relationship Based on Laboratory Testing," *Acta Geodynamica et Geomaterialia*, **5**(1), 2008 pp. 41–47.

[18] U. Wilensky. "NetLogo." Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (Oct 25, 2012) http://ccl.northwestern.edu/netlogo/index.shtml.

[19] M. E. J. Newman, *Networks: An Introduction*, New York: Oxford University Press, 2010.

[20] G. Peano, "Sur une courbe, qui remplit toute une aire plane," *Mathematische Annalen*, **36**(1), 1890 pp. 157–160. doi:10.1007/BF01199438.

[21] A. Banos, "Network Effects in Schelling's Model of Segregation: New Evidences from Agent-Based Simulation," *Environment and Planning B: Planning and Design*, **39**(2), 2012 pp. 393–405. http://halshs.archives-ouvertes.fr/docs/00/46/97/27/PDF/Banos-HAL.pdf.

[22] A.-L. Barabási and E. Bonabeau, "Scale-Free Networks," *Scientific American*, 14 Apr 2003 pp. 50–59. http://www.scientificamerican.com/article.cfm?id=scale-free-networks.

[23] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, **286**(5439), 1999 pp. 509–512. doi:10.1126/science.286.5439.509.

[24] L. Li, D. Alderson, J. C. Doyle, and W. Willinger, "Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications," *Internet Mathematics*, **2**(4), 2005 pp. 431–523. doi:10.1080/15427951.2005.10129111.

[25] T. C. Schelling, *Micromotives and Macrobehavior*, New York: Norton, 1978.

[26] J. Gómez-Gardeñes and Y. Moreno, "From Scale-Free to Erdos–Rényi Networks," *Physical Review E*, **73**(5), 2006 p. 056124. doi:10.1103/PhysRevE.73.056124.