Dynamics of interacting information waves in networks

A. Mirshahvalad, A. V. Esquivel, L. Lizana, and M. Rosvall

Integrated Science Lab, Department of Physics, Umeå University, Umeå, Sweden

(Dated: September 5, 2013)

To better understand the inner-workings of information spreading, network researchers often use simple models to capture the spreading dynamics. But most spreading models only highlight the effect of local interactions on global spreading of single information waves, and ignore effects of interactions between multiple waves. Here we take into account the effect of multiple interacting waves by using an agent-based model in which the interaction between information waves is based on their novelty. We analyzed the global effects of such interactions and found that selection between waves makes information that reaches nodes reach them faster: Slow waves die out and fast waves survive. While most spreading models assume uniform accessibility to information all over the system, our results show that accessibility to information, the wave frequency, decays as a power-law function of the distance from the source. Moreover, when we analyzed the model in different spatial geometries, including various synthetic networks and some real spatial road networks, we found that not only the distance from the information source, but also the path redundancy, the effective dimensionality of the system, determines the rate of accessibility to information. When we investigate the scaling results on real spatial networks, we found that these networks provide an infrastructure for information spreading which is between tree-like structure and two-dimensional lattice. Finally, to better understand the mechanics behind scaling results, we provide analytical calculations of the scaling for a one-dimensional system.

I. INTRODUCTION

In today’s society, we are flooded with information. Waves of new ideas, innovations, products, and trends follow each other in quick succession. To better understand the inner-workings of the dynamics, researchers often use simple models to capture important spreading mechanisms [1-7] on a complex network [8-16]. Broadly speaking, there are two classes of such models: threshold models [17-21] and contagion models [22-25]. Threshold models assume that individuals adopt new information or technology if a certain proportion of their friends have adopted. This mechanism leads to cascades that under favorable conditions can propagate throughout the entire system. Contagion models assume that individuals spread information or rumors much like they spread microbial infections through interactions. Also this mechanism can cause spreading across the entire system, provided that the transmission rate is sufficiently high. Both types of models highlight the effect of local interactions on global spreading, but in general they ignore effects of interactions between multiple waves.

However, ideas inherently depend on each other and waves of new information or technology often interact with one another as they propagate through society. In some systems, information waves integrate or hybridize while in other systems they compete and replace one another. Here we focus on the latter type of interaction, when waves replace one another entirely, and analyze the global effects of such interactions. For simplicity, we use novelty as a proxy for quality and key trait in the competition between waves [26, 27]. Relevant systems include news media reporting from a particular event, release of new software versions, and invention of new technology that makes old technologies obsolete. With interaction between multiple waves, some waves will make it across the system and others will not. Therefore, the wave frequency, or, equivalently, the rate of adoption of individuals, will depend on their position in the system. The aim of this paper is to analyze how the access to new information depends on position in a system and the topology of the system.

To analyze the effects of interactions between multiple waves, we use a simple agent-based model introduced in ref. [28] and further analyzed in ref. [29]. In its simplest formulation, a source agent injects multiple waves of new information over time in a given network. At a given rate, neighboring agents adopt the information if it is newer than the information that they already have. We provide analytical results of the wave frequency for a one-dimensional model and use simulations on lattice models between one and two dimensions, as well as on real spatial networks. In this way, we can quantify the effects of interactions between multiple waves and show, for example, that not only the distance from the source, but also the path redundancy, determine the rate of adoption. Moreover, compared to a system with non-interacting waves, new information reaches agents faster, because of selection between waves: slow waves die out and fast waves survive.

We begin by describing the model in detail and then analyze the model in different spatial geometries. In turn, we analyze the model from the perspective of the agents and the information waves, respectively.
II. METHODS

In this section, we first detail the model and then describe the spatial embedding we use to analyze the spreading dynamics.

A. Model definition

The model consists of a number of agents that each occupies a node connected to neighboring nodes in a spatial network. The core of the model describes interactions between multiple information waves released at a single source node. At each time $t$, the source in the center $j = 0$ generates a new piece of information tagged by the time when it was generated $a_0(t) = t$. In the same time step, each node $j$ with information of age $a_j(t)$ asks each of its neighbors $k$ with probability $\beta$ if $k$ has newer information. If $a_k(t) < a_j(t)$, $j$ adopts the new information and sets $a_j(t) = a_k(t)$. Throughout the analysis, we use $\beta = 0.5$. Note that this model formulation is equivalent to one in which agents actively transmit information to each of its neighbors with probability $\beta$, and the neighbors update their information if it is newer than the information they already had. Therefore, if there was only one information wave or if the waves did not interact, the model describes the standard SI dynamics of susceptible and infected individuals [20, 29, 31], and an information wave would always spread across the system. In the presence of multiple interacting waves, however, the information waves will compete with each other as they spread through the system and only the fast ones survive and make it across the system.

![Image](a.png) ![Image](b.png)

**FIG. 1.** The spatial and temporal dynamics of the spreading model with multiple interacting waves on a two-dimensional lattice. (a) A snapshot of the dynamics in which each color corresponds to a single information wave. (b) The age landscape of the model in which bright colors (light yellow) represent new information and dark colors (dark red) represent old information.

Figure 1 illustrates the dynamics of the spreading model on a two-dimensional lattice. This figure was produced from the Java applet available in ref. [32]. The source of new information is in the center of the lattice. Close to the information source, the diversity of information waves and the competition between them is high. Consequently, agents in this area become updated with high frequency. But far from the source the wave frequency is lower, because high competition close to the source eliminates waves. Therefore, nodes in distant areas must wait longer between each update of information. For example, for a line source that is located at one edge of a lattice, the density of wave fronts decays as the square root of distance from the source [28]. In this paper, we analytically derive this result for such a one-dimensional system and further show that the information wave frequency also depends on the path redundancy, the number of shortest paths between the source and a given node. Path redundancy corresponds to the effective dimensionality of a system and enables us to analyze systems with various effective dimensions. Higher path redundancy in a system gives nodes better access to new information.

B. Spatial embedding

To analyze the effects of path redundancy, we build synthetic spatial networks with varying degree of path redundancy. The networks range from trees with zero path redundancy to two-dimensional lattices with one path redundancy (Fig. 2). In the two-dimension lattice, the number of shortest paths grow as an exponential function of distance from the source In this paper, we cover a wide range of path redundancy and investigate real spatial networks that their path redundancy is between one and two. We construct the networks in two steps:

1. We start with a two-dimensional structure with nodes connected in horizontal rows and one vertical column through the source node in the center (Fig. 2(a)).

2. We then randomly connect a fraction $R$ of the remaining disconnected pairs of nodes (Fig. 2(b,c)).

Figure 2 schematically shows how we by connecting disconnected pairs in the tree structure in Fig. 2(a) can increase the path redundancy through intermediate values in Fig. 2(b) to high values in the fully connected two-dimensional lattice in Fig. 2(c).

To complement the analysis of synthetic networks, we also analyze two real spatial networks that their effective dimension are between one and two, the road networks of Texas and California [33]. For all of the described networks, we quantify the wave frequency as a function of distance and path redundancy. For comparison, we contrast the results with a null model that consider no interaction between information waves. In addition to wave frequency, we also quantify the speed of waves as a function of distance and path redundancy in one and two dimension lattice.
A two-dimensional lattice with path redundancy $R$ section, a path redundancy of an intermediate value, because, as we show in the next path redundancy. We used a path redundancy of 0.4 as acting waves on three networks with different levels of wave ever reached. The maximum penetration distance of a wave is the longest geodesic distance from the source that the fraction of the nodes in the network that simultaneously adopted the corresponding information as their latest information over its whole life time. The maximum extent of an information wave at its peak is the maximum fraction of the nodes in the network that simultaneously adopted the corresponding information as their latest information. This measure reflects the maximum popularity of a piece of information over its whole life time. The maximum extent of an information wave at its peak is the maximum fraction of the nodes that actually reaches them and (ii) the frequency at which new information arrives. In the next section we take the perspective of the individuals and in turn investigate these two effects in detail.

III. RESULTS AND DISCUSSION

Before we take the perspective of the individuals and show the results for the speed and wave frequency as a function of distance and path redundancy, we begin by taking the perspective of the information waves.

A. Spatial spreading profile

Unlike non-interacting information waves, many interacting information waves will die out long before they reach the boundaries of the system. To better understand the spreading dynamics and to get a general idea of how far they spread, we analyzed the spatial spreading profile by measuring the probability distribution of the information waves’ total extension, maximum extension, and maximum penetration distance. The total extension of an information wave over its entire life time is the fraction of nodes in the network that at some point adopted the corresponding information. This measure captures the aggregated popularity of a piece of information over its whole life time. The maximum extent of an information wave at its peak is the maximum fraction of the nodes in the network that simultaneously adopted the corresponding information as their latest information. This measure reflects the maximum popularity of a piece of information over its whole life time. Finally, the maximum penetration distance of a wave is the longest geodesic distance from the source that the wave ever reached.

Figure 2 shows the spatial spreading profiles of interacting waves on three networks with different levels of path redundancy. We used a path redundancy of 0.4 as an intermediate value, because, as we show in the next section, a path redundancy of $R = 0.4$ corresponds to the average path redundancy of the road networks of Texas and California. As Fig. 2 shows, the dynamic behavior of this intermediate path redundancy is the same as for the maximum path redundancy of the two-dimensional lattice. For all topologies, the competition between waves is most intense close to the source, such that most waves die out small before they have covered much ground. Except for boundary effects, which are significant for the extensions sizes, a power-law with exponent 1 approximately captures the scaling for all topologies. While the total extension size scales similarly for different degree of path redundancy (Fig. 3(a)), higher path redundancy increases the overall probability of spreading across the system. Since in a topology with high path redundancy, e.g., a two-dimension lattice, there are many possibilities for a wave to escape, many waves could reach to the system boundary. As a result, the fraction of waves that die before hitting the boundary is smaller than a system that has low path redundancy, e.g., a tree. Therefore, the probability density of maximum penetration distance has a vertical shift depending on the path redundancy of the system (Fig. 3(c)). Similar reasoning also applies to the maximum extension size. With no path redundancy in a tree-like topology, there is only one direction to expand into and the chasing waves follow immediately after. Therefore, in a tree-like topology, the waves could not expand as much as a system with high path redundancy, and there is a cut-off just above 200 (Fig. 3(b)). In contrast, with higher path redundancy, a fast wave can expand quicker in multiple directions and reach higher maximum expansion size.

Interactions between information waves prevents slow waves from reaching distant parts of the network. For the individuals’ that propel the spread of the information, this competition affects (i) the age of the information that actually reaches them and (ii) the frequency at which new information arrives. In the next section we take the perspective of the individuals and in turn investigate these two effects in detail.

B. Information wave speed and frequency

To investigate the extent of novelty for arriving information to a node, we calculated the age distribution of waves that reach a certain area. That is, we measured the average age of hitting waves for nodes at a given distance $d$ to the source. For comparison, we did the same experiment for multiple non-interacting waves. We run our experiment on a network with 3600 nodes (ordered in a 60 by 60 square) and quantified the probability distribution of information age for two groups of nodes: those that are close to the source ($d = 10$) and those that are far from the source ($d = 28$). In Fig. 4 we compare the probability distribution of the information age between interacting and non-interacting waves on two networks: a tree with the lowest possible path redundancy, $R = 0$, and a two-dimensional lattice with the highest possible path redundancy, $R = 1$. In both networks, information packages that reach a node have traveled shorter time in systems with interacting waves, because the interaction between waves form a selection process in which only fast waves survive.

Information that reaches a node is newer with than
FIG. 3. Spatial spreading profile of competing information waves on different networks. The networks have $1.2 \cdot 10^6$ nodes and each point corresponds to averaging over more than 5 independent runs. Each run simulates competition between 15,000 different information waves. While for the total extension size in (a), the probability densities are normalized to highlight the similar scalings, for maximum extension size in (b) and maximum penetration distance in (c), the probability densities are normalized by the absolute spreading profiles because ...

FIG. 4. Probability density of information age. (a) On a tree with 3600 nodes and $R = 0$. (b) On a two dimensional lattice with 3600 nodes and $R = 1$. We quantified the age distribution of arriving waves for nodes that are close to the source ($d = 10$) and for nodes that are farther away from the source ($d = 28$). Results are obtained by averaging over more than 10,000 different competing waves.

FIG. 5. Wave frequency as a function of distance for interacting and non-interacting information waves on different networks. The results on the tree and two-dimensional lattice are fitted to a power-law with exponent 0.5 and 0.3, respectively. For the California and Texas road networks, the results are very close to each other and similar to the network with path redundancy $R = 0.4$. All synthetic networks have more than $10^6$ nodes. The results on these networks are achieved by simulating more than 5000 competing waves. The results on the road networks are averaged over more than 25 runs and each run includes more than 30,000 competing waves.

nodes far from the source will only be reached by a fraction of all pieces of information that spread from the source. We quantified this effect with the wave frequency, the rate at which new information waves arrive at a node. In Fig. 5 we show how the wave frequency scales as a function of the distance from the source. We quantified the wave frequency as a function of geodesic distance on multiple networks: a tree ($R = 0$) with 1,210,000 nodes (1100 by 1100 square), a two-dimensional lattice ($R = 1.0$) with the same number of nodes, and two synthetic networks with the same number of nodes and path redundancies in between the tree and the lattice: one
with $\mathcal{R} = 0.4$ and one with $\mathcal{R} = 0.6$.

For non-interacting waves the wave frequency is the same for any node at any location and equal to the transition probability $\beta$ (gray line). For interacting waves, the wave frequency decays as a power law of the form $f(d) \sim d^\gamma$ with an exponent $\gamma$ that depends on the path redundancy. In general, for nodes at similar geodesic distance, a topology with higher path redundancy results in higher wave frequency. For example, the wave frequency decays similarly fast for the two road networks and close to the synthetic network with path redundancy $\mathcal{R} = 0.4$. Moreover, the exponent $\gamma$ is around 0.2 for the two-dimensional lattice and around 0.5 for the tree structure with no path redundancy.

Historically, spreading of innovation, information or ideas take place on spatial networks such as road networks. For example, trade roads like Silk Road were the backbone of information spreading for centuries. To investigate spread of information on spatial networks, we need to capture effective dimensionality of these networks. In this paper, we used path redundancy as a proxy for the effective dimensionality and showed that real spatial networks, such as California or Texas road network, have an effective dimensionality that is neither zero(tree-like topology) nor one(two-dimension lattice), but something in between. We also showed that for all networks with effective dimensionality between zero and one, accessing to new information decays as a power-law of the distance from the source. To highlight the mechanic of this power-law behavior, in the next section, we drive an analytical calculation for wave frequency in a one-dimensional system.

C. Analytical derivation of the wave frequency

In this section we provide an analytical derivation of the wave-frequency scaling in a one-dimensional system. We derive the wave frequency as a function of distance from the source by working with two quantities: the wave size $s$ and the wave position $r$. The wave position $r$ corresponds to the outer fringe of the wave, the farthest point of the wave from the source. Table I shows the corresponding transition probabilities of the wave’s size as a function of the transmission probability $\beta$. In particular, whenever the size reaches zero, the wave dies: $s = 0$ is an absorbing point.

In the model, the origin is special because a new wave starts there at each time step, no other points share that property. To make the analysis simpler, we will consider an alternate origin one step outwards. We will consider that each wave starts there with size one, and after we obtain the expressions we are interested in, we will correct the results to consider a new wave for each time step.

The position and size of the wave are coupled, they can only change in the way shown by the state transition diagram in Figure 6. In the figure, red nodes represent dying states, black nodes represent surviving states, and the only green node corresponds to the initial state. Because time is not present in the diagram, we have to discard the case when the wave keeps both position and size in a time step, and normalize the other probabilities accordingly; this is the origin of the denominator in the expressions that follow. The horizontal arrows correspond to the cases when the size of the wave does not change but the wave moves from a position $r$ to the next position $r + 1$. This situation happens with probability $\beta/(2 - \beta)$, which we denote $b$. The vertical arrows correspond to the case when the size of the wave shrinks while the position of the wave stays unchanged. This case happens with probability $\beta(\beta - 1)/(\beta - 2)$, which we denote $c$. Finally, the inclined arrows correspond to the case when the size of the wave expands and the wave moves from a position $r$ to the next position $r + 1$. This case happens with probability $\beta(\beta - 1)/(\beta - 2)$, also, which for the sake of clarity we denote as $c$. Using these transition probabilities and guided by the diagram in Fig. 6, we can derive a recursive equation for the probability $g(r)$ of having a size-one wave at a given position $r$.

FIG. 6. state transition diagram of position-size The position and size of the waves can change by following this diagram’s arrows. The horizontal axis represents the position $r$ of the right fringe, and the vertical axis represents the size $s$ of the wave.

For each time step, the state can change following the arrows in the diagram. Because this scheme is regular enough, any walk of the wave through this state space that starts and ends at $s = 1$ is equivalent to a walk that starts and ends at a different value of $s$ (say $k$), as long as this second walk is contained in the half-space $s \geq k$.

<table>
<thead>
<tr>
<th>Value transition</th>
<th>Probability</th>
<th>Short name</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s \to s - 1$</td>
<td>$(1 - \beta)\beta$</td>
<td>$a$</td>
</tr>
<tr>
<td>$s \to s$</td>
<td>$\beta^2 + (1 - \beta)^2$</td>
<td>$1 - 2a$</td>
</tr>
<tr>
<td>$s \to s + 1$</td>
<td>$(1 - \beta)\beta$</td>
<td>$a$</td>
</tr>
</tbody>
</table>
Let us suppose that the walk starts at \( s = 1, r = 1 \), which we abbreviate as \((1,1)\) and represent the position with the green point in the diagram. If we fix a value of the position, say \( r \), we want to know the probability \( g(r) \) that the wave will go from the green point to a new position of size 1 in the state space, the point \((r,1)\). This last state can only be reached in two ways: from the left, with a green arrow, or from the state above \((r,2)\), with a gray arrow. If the state is reached from the left, it means that the wave had to go from \((1,1)\) to \((r-1,1)\), and thus we need the value of \( g(r-1) \). The case for the wave coming from above is more convoluted and we handle it in the following way. First, because the state above had to go to \( 2 \), we can safely assume that there was at least one transition where \( s \) changed from 1 to 2. Let us fix the position of the last of such a transition as \( r_1 \), such that \((r_1,1)\) and \((r_1+1,2)\) is a segment of the walk. We can see that the walk from \((1,1)\) to \((r_1,1)\) will have probability \( g(r_1) \). The segment from \((r_1,1)\) to \((r_1+1,2)\) will have probability \( a \). That leaves us with the part of the walk from \((r_1+1,2)\) to \((r,2)\). Because the last transition from \( s = 1 \) to \( s = 2 \) was the one at \( r_1 \), there is no way that the walk could go to \( s < 2 \) in the section from \((r_1+1,2)\) to \((r,2)\). In other words, this part has probability \( g(r-r_1-1) \). With this information, we can write the recurrence for \( g(r) \) as

\[
g(r) = \sum_{r_1=0}^{r-1} [g(r_1) a g(r-r_1-1) c] + bg(r-1). \tag{1}
\]

To solve the above equation, we use a generator function of the form

\[
G(z) = \sum_{i=0}^{\infty} g(i) z^i. \tag{2}
\]

Since Eq. 1 is only valid for \( r \geq 2 \), we first write

\[
G(z) = g(0) + z g(1) + \sum_{i=2}^{\infty} g(i) z^i, \tag{3}
\]

and then apply the recursivity to obtain

\[
G(z) = g(0) + z g(1) + \sum_{i=2}^{\infty} b g(i-1) z^i + \sum_{i=2}^{\infty} \sum_{i_1=0}^{i-1+i} ac g(i-i_1-1) g(i_1). \tag{4}
\]

With some variable substitutions and algebraic manipulation we can write it as

\[
G(z) = g(0) - b z g(0) + z g(1) + b z G(z) + ac z \sum_{j=0}^{\infty} \sum_{i_1=0}^{j} g(j-i_1) g(i_1) - zac g(0) g(0). \tag{5}
\]

The terms in the sum of the previous expression represent a neat convolution, which can be expressed as the product of generating functions. From there we get the quadratic expression for \( G(z) \),

\[
G(z) = 1 - b z - ac z + (b + ac) z + b z G(z) + ac G(z)^2. \tag{6}
\]

From the two solutions, we select

\[
G(z) = -1 + b z + \sqrt{1 - 2 b z - 4 ac z + b^2 z^2} \over 2 ac z. \tag{7}
\]

As mentioned in the text, with probability \( c g(r) \), the wave is going to die without ever reaching the position \( r + 1 \). Thus, by summing \( c g(x) \) from \( x = 1 \) to \( x = r - 1 \), we can calculate the fraction of walkers alive at a specific position \( r \) of any size as

\[
h(r) = 1 - c \sum_{x=0}^{r-1} g(x). \tag{8}
\]

To calculate the survival probability \( h(r) \), we write down the corresponding generating function as

\[
H(z) = \sum_{i=0}^{\infty} \sum_{j=0}^{i-1} z^i c \sum_{r=0}^{g(r)} g(r). \tag{9}
\]

The first term in the difference is \( \frac{1}{1-z} \) and the second term is again a convolution:

\[
H(z) = \frac{1 - c \sum_{i=0}^{\infty} z^i g(i) + c \sum_{i=0}^{\infty} z^i g(i)}{z - 1}. \tag{10}
\]

such that we get

\[
H(z) = \frac{c z G(z) - 1}{z - 1}. \tag{11}
\]

After substituting Eq. 3 and doing some simplifications we obtain

\[
H(z) = \frac{2}{\sqrt{(-4 + \beta (4 + \beta (-1 + z))) (-1 + z) - \beta (z - 1)}}. \tag{12}
\]

The function \( H(z) \) has only one principal singularity at \( z = 1 \), and we know that the coefficients \( b(r) \) are strictly positive. Therefore, we can apply Corollary 2 in ref. [3] and derive the asymptotic scaling for \( h(r) \). By that corollary, we get that when \( r \to \infty \),

\[
h(r) \sim \frac{1}{\sqrt{1 - \beta \sqrt{\pi}}} r^{-1/2}. \tag{13}
\]

We should remember that \( h(r) \) represents the survival probability of the wave once it takes off at the first position immediately after the origin, and that happens with probability \( \beta \), so that the frequency with which new waves are observed at a given point \( r \) is

\[
f(r) \sim \frac{\beta}{\sqrt{1 - \beta \sqrt{\pi}}} r^{-1/2}. \tag{14}
\]

This expression is valid as long as \( \beta < 1 \), provided that \( r \) is sufficiently large. Figure [4] shows the values of the frequency as obtained by simulation and those obtained by the previous equation.
FIG. 7. The wave-frequency at different positions \( r \). Red dots: the frequency obtained averaging the observation of 2e6 time-steps. Green line: the theoretical prediction according to equation [14].

IV. CONCLUSION

We used a simple agent-based model to capture that waves of new information or technology often interact with one another as they propagate through the system.

In the model, we use novelty as a proxy for quality and key trait in the interaction between waves. We showed that information that reaches agents is newer with than without interactions between waves at the cost of lower arrival frequency of information waves. Moreover, high path redundancy has a positive effect on the wave frequency, such that information more easily spreads in a system with multiple routes to targets. In general, the wave frequency decays as a power law of the distance from the source, and analytically we showed that the scaling goes as one over the square root of the distance in a one-dimensional system. Our analysis on road networks of California and Texas showed that these networks provide an infrastructure for information propagation that correspond to lattice models between one and two dimensions. We conclude that interacting information waves show interesting dynamics that calls for further study.

ACKNOWLEDGMENTS

We are grateful to Akhil Kedia for many valuable suggestions. This research was conducted using the resources of High Performance Computing Center North (HPC2N) and MR was supported by the Swedish Research Council grant 2012-3729.


