Impact of the Interaction Network on the Dynamics of Word-of-Mouth with Information Seeking*

John Doe

removed for blind review
john.doe@gmail.com

Abstract. Word-of-Mouth refers to the dynamics of interpersonal communication occurring during the diffusion of innovations (novel practices, ideas or products). According to field studies, word-of-mouth is made of both information seeking and proactive communication: individuals first become aware of the existence of an innovation, then start actively seeking out for the expert knowledge required to evaluate the innovation; when they hold the expert knowledge, they might start promoting it pro-actively. Successful diffusion of innovation requires the individuals to hold both awareness and expert knowledge, so they can evaluate the innovation and use it properly. In this paper, we investigate the dynamics of the computational model "USA/IPK" designed to study the role and impact of information seeking on the dynamics of word-of-mouth. The experimental results demonstrate the dynamics of the model are similar across tested networks, with the noticeable exception of the efficiency of the diffusion which varies between networks having similar densities and sizes; this difference is likely due to the core-periphery structure of the network and/or the diameter of the network.

Keywords: word-of-mouth · information seeking · complex networks · social networks · social simulation · information dynamics · computational science

1 Introduction

1.1 Word-of-Mouth: Evidence on Information Seeking

When individuals discuss an innovation (a novel product, practice, idea) [23], they *spread the word* about its existence and qualities. More people become aware of a product through word-of-mouth than traditional advertisement [25]. Most consumers attribute a higher importance to interpersonal influence than

* The author would like to thank Sarah Thiriot for the numerous improvements she suggested in the presentation of this research. This research is based on a simplification of a detailed model of innovation diffusion [28] which was partly funded by the french National Association for Research and Technology (ANRT) grant 993/2005 and by the Orange Labs company.

other sources [13]. As a consequence, word-of-mouth is consistently said to determine the success or failure of innovations [23] and facilitate the diffusion of products [13].

Word-of-mouth was often reduced to an epidemic process in which individuals "contaminate" each other with information [9]. Yet interpersonal communication about innovations or products [3] does not only include the proactive emission of information, but also communications initiated by people who seek out information about an innovation [8,23]. When an individual discovers the existence of the innovation, that is when he receives awareness knowledge from advertisement or another individual, he might engage (or not) in information seeking depending on his characteristics [8,23]. Expert knowledge covers "how to" use the innovation (know-how knowledge) [23], "why" the innovation works (principles-knowledge) [23], product category or product-class knowledge or brand knowledge . This expert knowledge might be gathered from individuals who hold it prior to the diffusion of the innovation because of their education or training, because they read specialized press, had experience with another product of same brand, category or class, or because they received this information from another individual. Once they hold the expert knowledge, people might engage into pro-actively passing the word around about the innovation, for instance because they are willing to help others [8] or because they are satisfied or dissatisfied after adoption [2].

Information seeking stands as a step required for most individuals to be able to decide to adopt or reject a product [13]. In the case of the diffusion of disruptive innovations such as vaccination or contraceptives [23], information seeking is even seen as a mandatory step for individuals to adopt the innovation, as it enables people to understand why it works and how to use it. For instance, parents do not accept the vaccination of their children without gathering more knowledge first [30]. Even if innovations might be adopted without expert knowledge, the misuse of the innovation may later cause its discontinuance [23]. As a consequence, *a company or organization promoting an innovation attempts to maximize the proportion of the population which is not only aware, but also holds expertise on the innovation* [23].

1.2 Word-of-Mouth: Computational Models

Numerous computational or mathematical models [20,21,16] were designed to understand the diffusion of information and innovations [12,11,18], assess the potential diffusion of products [5], and recommend strategies to accelerate or maximize this diffusion [14,17]. The three main types of models related to information diffusion are based on information cascades, social influence and social learning [18,32]. The models developed in these last two categories describe the flow of influence within the population, without explicitly representing the flow of information, and can not be used to study the impact of information seeking.

Marketing models based on *information cascades* [9,10,14,24] (also referred to as the Independent Cascade Model) rely on an analogy with epidemic models [9,7] such as the SIR model [15,4]: every individual is either in state Susceptible (no information), Infective (informed and pro-actively passing the information to others) or Recovered (informed but passive). A Susceptible individual becomes Infective (S $\rightarrow$ I) when he meets an Infective individual. After a given time, Infective agents become Recovered (I $\rightarrow$ R). When enough individuals are passing the word around, information cascades appear in the simulations, as observed in reality [18]. In this case, the cumulated curve of the proportion of people informed in time follows the traditional S-shaped curve. Unfortunately, because they only include the information passing behaviour without any information seeking, these models only capture part of the dynamics of word-of-mouth.

1.3 Word-of-Mouth with Information Seeking

To study the impact of information seeking, we describe in USA/IPK two dimensions of knowledge: *awareness knowledge* which refers to knowing the innovation exists, and *expert knowledge* which allows the actual understanding of the innovation. On the *awareness dimension*, the agent is either *Unaware* (U) (does not know the innovation exists), *Seeking* (S) (discovered the innovation exists and seeks out for expert knowledge) or *Aware* (A) (knows the innovation exists but does not actively search for information). On the *expertise dimension*, the agent is either *Ignorant* (I) (does not holds the expert knowledge), *Proactive* (P) (holds the expert knowlege and shares it around him) or *Knowledgeable* (K) (passively holds the expert knowledge, without sharing it pro-actively). The agents start by default in state (U, I), meaning they are Unaware of the existence of the innovation and Ignorant of the expert knowledge required to understand it. A given proportion of agents $k \in [0:1]$ is initialized in state (U, K), meaning they are Unaware of the innovation but hold the expert knowledge to evaluate its benefits, which they are supposed to know because of their education, training or experience with another similar innovation. A low proportion of expertise k in the population represents a disruptive innovation; a higher proportion of expertise represents an incremental innovation that most people can understand as soon as it is advertised.

In order to reflect the impact of personality on information behavior identified in literature (cf 1.1), the authors introduced personality variables which are randomly initialized at the beginning of the population based on ratios provided as model parameters and constant during the simulation. "*Curious*" individuals refer to agents which, when they receive awareness, shift to the "Seeking" state; non curious agents would transition to "Aware" instead. "*Enthusiastic*" agents refer to agents which, when they received the expert knowledge, promote it pro actively around them; non enthusiastic agents would transition directly to Knowledgeable instead. "*Supporter*" agents are those who, when they discover awareness when they are already holding expert knowledge, start promoting the information. The transitions between the knowledge states are described in the state diagram in Fig 1.

During N steps of the simulation, an advertisement campaign dispatches the awareness knowledge to a small proportion of the population: each of the agents

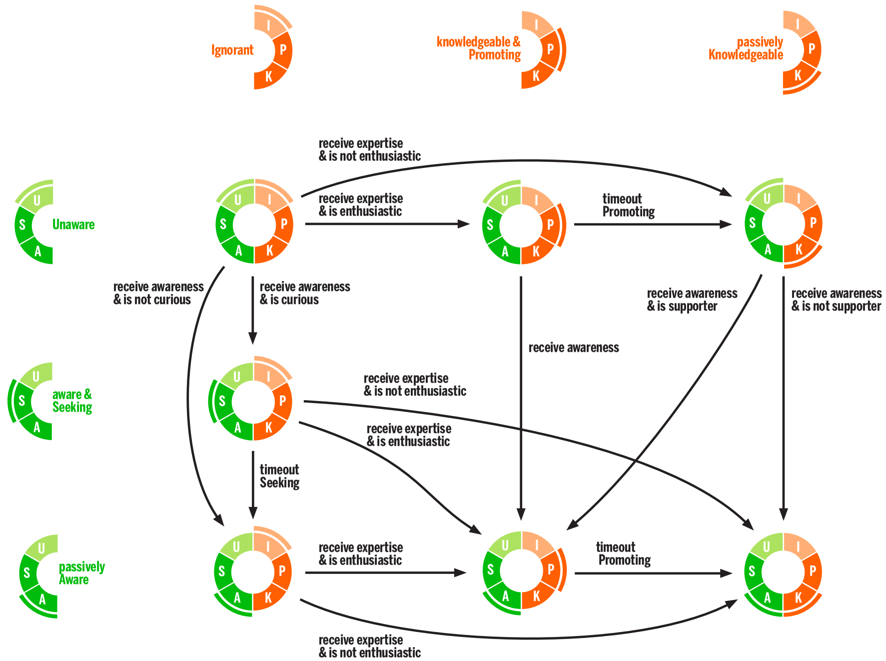

Fig. 1. State diagram representing the evolution of the awareness and expertise states of agents in the USA/IPK model depending on the characteristics of the agent (curious, enthusiastic, promoter) when the agent receives awareness or expertise. The vertical and horizontal dimensions contain respectively the states related to awareness and expertise. The combined state is read as: (U, I) "the agent is Unaware and Ignorant". Arrows depict the conditions for an agent to switch from one state to another one. For instance $(U, I) \rightarrow (S, I)$ is read as: "an agent starts Seeking out information after receiving awareness if it is curious, but does not change its level expertise which remains Ignorant". An agent holds complete knowledge when he is (A, P) or (A, K). An agent instigates conversations either when it is (S, I), (U, K) or (A, K).

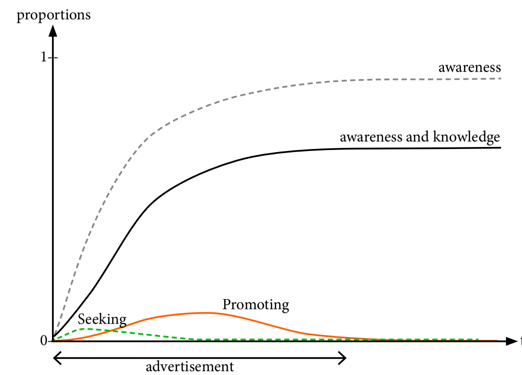

Fig. 2. Typical successful diffusion of awareness (all the agents in state $(A, *)$) and complete information ((A, K)). Dashed curves depict the propagation of awareness, and the underlying proportion of agents Seeking out information during the simulation.

of the simulated population has a given probability to receive the awareness message. Then the interpersonal interactions in the model drive the propagation of knowledge. At each simulation step, every edge of the interaction network (taken in random order) offers its vertices the opportunity to interact. Interactions actually take place if one of the agents is Seeking out information or Proactively speaking. Agents who became aware of the innovation, and have the trait curious, start to seek out for expert knowledge around them; by doing so, they propagate awareness around them; when they discover the expert knowledge, they stop seeking out information; the individuals having trait enthusiastic start promoting the innovation, whilst the others become passive. A typical simulation exhibits two S-shaped curves as displayed in figure 2: the higher is the curve of awareness, and the lower one the curve of both awareness and expert knowledge. This model enables to experiment on the dynamics of information seeking: what are the relative impacts of curious people and enthusiastic ones? On which conditions can we enhance the retrieval of expertise in the population? Published results [26] highlight a surprisingly high efficiency of information seeking, which enables most of the population to achieve gathering an expert knowledge held by as few as 0.1% of the population. Simulations experiments also suggest three different regimes for different proportions of initial knowledge, suggesting different communication strategies for disruptive and incremental innovations.

1.4 Research Question & Outline

Each agent in a simulation of the USA/IPK model is located over an interaction network which determines with which agents it interacts at each time step. The dynamics of the model published so far [26] were only simulated over Watts-Strogatz networks [31]. Yet existing knowledge on real social networks (i.e [1]) suggests they might have different characteristics than WS ones, including a skewed distribution of degrees, a strong modularity or a core-periphery structure. As the structure of interactions usually plays a central role on the dynamics of a model, we question here the impact of the network of interactions on the dynamics of the IPK/USA model: do different networks lead to different qualitative dynamics? We first describe (section 2) our experimental protocol to tackle these questions: space of parameters explored for the USA/IPK model, selection of random network generators, actual implementation oriented towards reproducible research. We then present the Experimental Results (section 3 p. 9) at both mesoscopic and macroscopic scales. We then discuss (4 p.14) the implications of these findings on the methodology to compare computational models over networks, the practical findings for the USA/IPK model, and novel questions arising from these findings.

2 Experimental Protocol

2.1 Space of parameters for the USA/IPK model

Despite of having been designed as a minimalistic model of word-of-mouth dynamics with information seeking, the USA/IPK model still has a large space of parameters to explore. These parameters include parameters related to the innovation on which information diffusion is simulated, the advertisement campaign, the social behavior, and the network of interactions.

The *parameters related to the innovation under study* include: the initial proportion k of agents possessing the expert knowledge required to understand the innovation; the proportions of individuals who would start being curious if they receive awareness; the proportion of agents who would start proactively spreading the word about the innovation when they understood it, and the proportion of agents who would spread the word about the innovation when they are already able to understand it when they become aware. All of these parameters are meant to be systematically explored for the USA/IPK model: the proportions of curious, proactive and supporters have to be finely explored (every 0.05 between 0 and 1) in order to quantify the threshold effect; such a fine exploration might be used to recommand whether the innovation and communication should be oriented towards creating surprise or achieving to make people promote it. For each network generator, we explore the same space of parameters of the model as in the original study to facilitate comparisons: the proportion of curious and enthusiastic people is explored in $[0 : 1]$ by steps of 0.05. The proportions of supporters include 0.0, 0.1 and 0.5. The proportions of initial knowledge were identified as leading to 3 different regimes in the original publication; identifying three values in this space for these three regimes is sufficient. As in the original study we explore k with $k = 0.01$ (1%, disruptive innovation), $k = 0.1$ (incremental innovation), $k = 0.5$ (standard product understood from 50% of the population).

The *parameters related to the advertisement campaign* include the duration of the advertisement campaign N (in steps) and the probability to reach individuals. These two parameters have straightforward results. High values of exposure and/or proportions lead to quicker diffusion of awareness. We fix to values which enable the comparison of different dynamics, and reuse the values of the initial publication: $p^{\text{exposure}} = 0.04$ and $N = 6$ (4% of the population of agents is reached during each of the 6 first steps).

The *parameters related to the social behavior* include the timeouts for active seeking and proactive communication. As we will discuss in section 3.1, these parameters impact the dynamics in a way which can be explained and understood, so we will set these values to the average path length of the network: 5.

The last set of parameters, of a highly specific nature, is the *network of interactions* detailed in the next section. Note the implemented model USA/IPK enables the usage of probabilities of interactions which are left to 1.0, so a link between two agents mean they will always be offered to interact at every simulation step, whilst no link means they will never interact directly.

The interest of the dynamics of the initial study can not be reduced to one unique epidemic threshold; the interesting elements were mostly qualitative, as the observation of three different regimes depending on the proportion of initial expertise, or the asymmetry in the role of information seeking and proactive communication. As a consequence, we drive an extensive exploration of the space of parameters described above and interpret the dynamics of the model in this space as in the original study. We will come back on every qualitative finding of the original study and compare the result with different networks.

For each point of the space of parameters, at least 10 simulations are started with a different random seed. When the measures are too unstable, additional simulations are added, with up to 100 simulations in the areas where the model is less stable. We measure the final proportion of the population holding both the awareness and the expert knowledge. The results quantify the average values of the indicators over the several simulations.

2.2 Selected Random Network Generators

As for an epidemic model and other models in which the dynamics depends on cascades, some statistical properties would obviously impact the dynamics of the model. If the density of the network increase, each agent seeking out expert knowledge is more likely to find it out, and agents pro actively transmitting information would have more impact. In the same way, networks having different sizes, diameters or average path lengths might bias the results because the increased steps required for information to flow in the entire population. In order to avoid these trivial effects of networks on the dynamics, we select network generators which are all compliant with the small-world effect (low density, short average path length) and all have a high clustering (or transitivity rate). We then set the parameters of these random network generators so that they generate networks having close statistical properties for density, average path length, count of vertices and diameter. We retain for these experiments three network generators that are able to reach similar statistical properties.

[WS] The famous Watts and Strogatz β-model [31] requires as parameters N the size of the network, *nei* the neighbourhood of the original lattice and p^{rewire} the rewiring probability. This algorithm starts with a regular lattice of N nodes in which nodes are connected with their *nei* neighbours (thus having $2.nei$ degree). It then rewires each link with probability p^{rewire}. We define here $n = 1000$, $nei = 5$, $p = 0.055$; this leads to networks having a density of 0.01, average path length 4.4 ± 0.3 and a clustering rate of 0.47 ± 0.02.

[FF] The Forest Fire model was proposed by Leskovec [19] as an algorithm that creates networks having most of the properties observed in real networks, including communities, skewed distribution of degrees, a core-periphery structure. The network is grown step by step, each new node A being attached to m old nodes. Moreover, each time a new link is created between A and B, A explores the outgoing and incoming neighbours of B. A create links with outgoing nodes of B with a *forward probability* p, and also creates links with incoming nodes of B with probability $p.r$, with r the *backward burning ratio*. As this step is ran recursively, A is said to "burn" all the possible links. Here FF is set up with $n = 1000$, fw.prob $= 0.37$, bw.factor $= 0.9$, $ambs = 1$. The resulting networks have a density of 0.01 ± 0.005, an average path length of 4.13 ± 0.2, and a clustering rate of 0.26 ± 0.2.

[SII] We also test a simple model that creates networks composed of several communities (sets of nodes having a strong density). This model, later named SII for Simple Interconnected Islands, starts by creating n islands of identical size $size$. Each island is a random graph in which links exist with probability $p.in$. Each island is connected with all the other islands with $n.inter$ links, each being created between two nodes randomly picked from each island. Density and transitivity in SSI networks can easily be tuned by varying the n, $size$ and $p.in$ parameters, while the average path length may be tuned with $n.inter$. Its distribution of degree is nearly a Poisson-like one (as each island is a random network). This average path length remains low, because all the islands are interconnected. For this study we use $n = 24$ islands, size $= 42$ nodes per island, p.in $= 0.235$ of wiring probability within islands, n.inter $= 1$ link between each pair of island. The resulting networks count 1008 vertices, have a density of 0.0101 ± 0.002, an average path length of 4.36 ± 0.02, and a clustering rate of 0.21 ± 0.1.

2.3 Implementation

These experiments require the chaining of the generation of complex networks and the agent-based simulation over this network. In order to facilitate the exploration of this experimental design, facilitate reproducibility and avoid manipulation errors, we use the OpenMole scientific software [22] which enables the description of scientific computation workflows in which tasks are chained with each other. In our workflow, for each simulation, a different random network is generated in R using the reference igraph package [6]; it is simplified to remove the potential loops or double links; its statistical properties are analyzed (average path length, clustering rate, diameter, average degree of connectivity, connectedness); if the network is not compliant with the experimental conditions (such as connectedness), another random network is generated. The network is then written to a file in graphml format. OpenMOLE ensure this file and the statistical properties of the network are transmitted to the multi-agent simulation engine and starts the simulation using Netlogo [29] version 6. The sourcecode of the USA/IPK model is the original one shared in the github repository[1]. The simulation is stopped when there is no possible evolution of knowledge in the population anymore, that is when the advertisement campaign is finished and there is no more agent in state Seeking nor Proactive. Once the simulation is finished, OpenMOLE accumulates the result of the simulation (duration of the simulation, proportion of the agents being aware, and proportion of agents being both aware and knowledgeable) and the properties of the interaction network in a CSV file. We then use R to analyze the result of simulations, and gnuplot to plot the graphics displayed in this article.

[1] `https://github.com/samthiriot/model-wom-USA-IPK`

3 Experimental Results

3.1 Typical Micro and Meso Dynamics

During a simulation of the USA/IPK model, three distinct dynamics occur at the mesoscopic scale which explain the results obtained at the macroscopic scale: cascades of awareness, chains of information retrieval and cascades of expertise.

An agent in state (U, I) which receives awareness, either from the advertisement campaign or thanks to a peer transmitting the information, might also become actively Seeking out information if it is "curious". Therefore *cascades of awareness* appear in the population as soon as the required conditions are met. We illustrate in Fig 3 an example of chain of awareness. As can be expected in such a dynamics typical of epidemic dynamics, a threshold effect appears, with an epidemic threshold depending on the proportion of agents being initialized as "curious" and on the degree of connectivity in the network. During the advertisement campaign, several cascades of awareness occur in the network of interaction which increase the proportion of awareness with a S-shaped curve. Yet the diffusion of awareness is not sufficient for a target population to decide to adopt or not an innovation; they first have to achieve to gather expert knowledge from on agent in the network who holds it.

When an agent in state Seeking (S, I) interacts with an agent which holds the expert knowledge (that is an agent either actively Promoting $(*, P)$ or passively Knowledgeable $(*, K)$), this agent discovers the expert knowledge. In the next simulation step, an agent connected in the interaction network which is also Seeking out would then be able to retrieve this expert knowledge. We observe in the simulations *chains of information retrieval* (illustrated in Fig 3) which appear when a cascade of awareness is reaching a Knowledgeable agent. Thanks to these chains, a population with a low proportion of initial expertise in the population is able to retrieve it through the social networks with a good efficiency.

Chains of information retrieval depend on a relationship between the average path length and diameter of the graph of interactions, and the timeout after which agents in state Seeking or Proactive become passive. Each step of a chain of information retrieval requires a simulation step; so a chain of information retrieval involving 4 agents requires 2 steps for the cascade of awareness, then 3 steps for the expertise to backpropagate. Yet the expertise only backpropagates if each agent is still in state curious. As a consequence, the maximum length of a chain is half of the Seeking timeout.

A last dynamics which occur in the population is the apparition of *cascades of expertise*. When an agent discovered expert knowledge and is supporter or promoter, he shifts to state $(*, P)$ and starts spreading expert knowledge to its neighbors in the graph of interactions. If enough agents around are also supporters, proportionally to the degrees of connectivity, a cascade of expertise occurs. This is also typical of an epidemic process with an epidemics threshold which depends on the topological characteristics of the network.

The three dynamics we described at the mesoscopic scale pile up and interleave at the macroscopic scale, thus highlighting three different thresholds on the proportion of curious, supporters and promoters for a given degree of connectivity. Because the impact of the degree of connectivity is trivial, we analyze in the next session what happens over networks having similar characteristics for the average degree of connectivity.

3.2 Impact of Interaction Networks

Figures 4, 5 and 6 depict the simulation results for initial proportions of expertise of 1%, 10% and 50%. These figures represent the results of a total of 276676 simulations.

Authors highlighted **threshold effects** in the original study: a small increase in the proportion of curious or enthusiasts leads to a significant shift in the proportion of informed people. These effects are also visible here. On every network when $k = 0.01$ (Fig. 4), a small increase of the proportion of curious between 0.0 to 0.05 leads to diffusion rates as different as 0 or 100%.

Original results highlighted **three distinct regimes** of the model depending on the proportion of initial expertise k, with information seeking leading the dynamics when k is low, symmetric roles of information seeking and supporters with k higher, and first importance of supporters when k is very high ($k >= 0.5$). Thee distinct regimes also appear in our experiments depending on the value of k. For $k = 0.01$ (Fig. 4), there are vertical dark areas at the left side of the figures which traduce an absence of diffusion when there are no curious individuals in the population. In this part of the space of parameters, information seeking thus stands as a mandatory step for the communication to start. For $k = 0.1$ (Fig. 5), the same effect is visible when there are no supporters (left figures), less visible with a few supporters, and not visible when there are many supporters (right figures). In this regime, curious and enthusiast agents play a similar role: at least one of them is required for the diffusion to start. For $k = 0.5$ (Fig. 6), the success of diffusion is less sensitive to the proportion of curious agents; the supporter parameter here stands as a key element along with the proportion of enthusiasts; it would mean that when expert knowledge is widely available in the population, information seeking is less important than proactive communication. We confirm with these results the existence of three regimes depending on the proportion of initial expertise k on WS and SII networks.

The original study points out the **strong asymmetry** in the impact of the proportions of curious and proactive agents: while diffusion can occur with only a few curious and no enthusiastic, it is not possible with only a few enthusiastic and no curious. This effect has a potential strong impact for policies, as it suggests it is more important to create campaigns of information that make people seek out for information than make people spread the word when they are knowledgeable. We also observe this result here: a minimal proportion of curious being required on any network and any proportion of initial expertise when there are no supporters. No other parameters plays this role in the dynam-

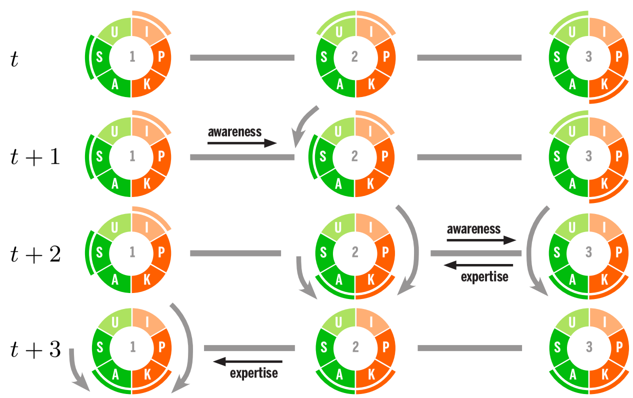

Fig. 3. Example of a *cascade of awareness* and of a *chain of information retrieval.* At time t, agent (1) is Seeking out expert knowledge; this expertise is only available from agent (3) to which (1) is not directly connected with in the social network. When it seeks for information, agent (1) makes agent (2) aware as well. If agent (2) has property curious, it also enters Seeking. Then at time $t+2$, this agent sends awareness to agent (3). There was a cascade of awareness between (1), (2) and (3). Agent (2) also retrieves expertise from (3), which can later be retrieved by (1). There was a chain of information retrieval from (3) to (1). These chains enable the population to retrieve expertise even if it is rare in the population.

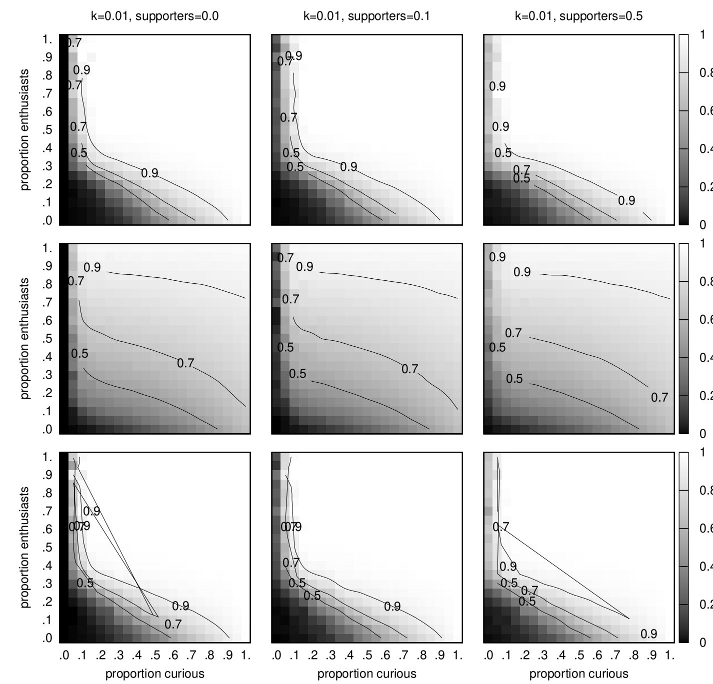

Fig. 4. Final proportion of the population holding both awareness and expertise when 1% of the population initially holds the expertise, over different networks: *(top)* WS *(middle)* FF *(bottom)* SII. On these graphics white and black respectively mean 100% and 0% of the simulation population hold both awareness and expertise.

ics. This suggests, in line with field observations (section 1.1), that information seeking is a mandatory step for diffusion of information.

In the original study, the **impact of supporters** (that is individuals who spread the word when they discover awareness *after* holding the expert knowledge) only have an important impact when the initial amount of knowledge is high. We observe the same impact here: the proportion of supporters is essential for $k = 0.5$, visible for $k = 0.1$, and barely noticeable for $k = 0.01$. This observation stands on all the networks.

The original study also highlighted how it is surprisingly **easier to retrieve expertise when there is fewer initial expert knowledge in the population**. This finding was counter-intuitive: if more expert knowledge is disseminated in the population, then any agent seeking out information is likely to find it; the more initial expert knowledge, the more efficient the diffusion result should be. We find here the same results as in the initial study: with $k = 0.5$, the final diffusion appears lower than with $k = 0.1$ on every network. As in the initial study, this observation is explained by the diffusion process: if there are few individuals holding expertise, the people initially seeking out will create long chains of information seeking; when one agent of this chain gathers expertise, they will all collect it back (chain of information retrieval). If most people are initially knowledgeable however, as soon as an agent seeks out for information, he will find out expertise, and stop seeking out; he will not have raised the attention of many others during the process. Our experiments replicate this phenomenon on every network.

In the original experiment, authors noted the **high efficiency of word-of-mouth to gather the expert knowledge scattered in the population**. The same result is obtained here on WS or SII: when there is 1% of expertise in the population (Fig. 4), having 30% of curious and 30% of enthusiasts leads to more than 90% of success. On FF however, the efficiency is noticeably lower: diffusion to the entire population is very rare. We can reproduce the original results on WS, but we also identified a counterexample with the FF networks. As these networks have similar density and the same order of magnitude of average path length, these characteristics are unlikely to explain those differences. Other statistical indicators of the dynamics of the model also do not explain why the diffusion would be more difficult over FF networks. The duration of the diffusion is longer over WS than FF and SII, but those last two share the same distribution of duration. Observations of the dynamics of diffusion over the network suggests the core-periphery structure of FF might explain this phenomenon: the initial expert knowledge scattered in the periphery is more difficult to retrieve from the network (less dense areas). Moreover, despite the FF networks having an average path length being slightly lower than the ones of SII and WS networks in our settings, their diameter remains considerably higher, with an average of 14 instead of 7 for WS and SSI. This longer diameters makes the information less accessible in the network.

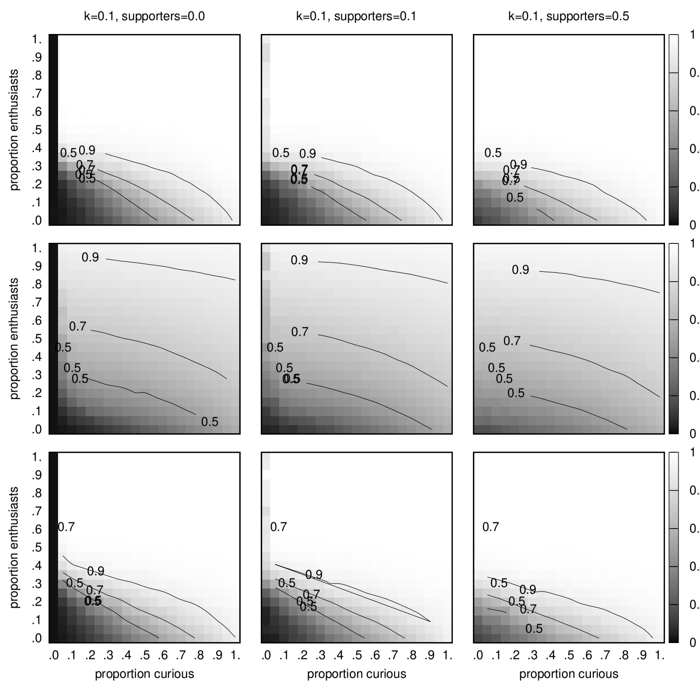

Fig. 5. Final proportion of the population holding both awareness and expertise when 10% of the population initially holds the expertise. *(top)* WS *(middle)* FF *(bottom)* SII

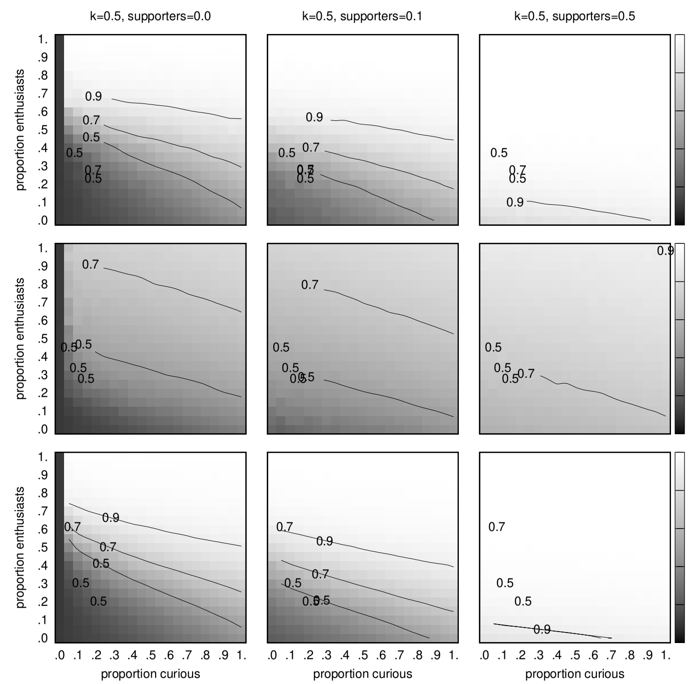

Fig. 6. Final proportion of the population holding both awareness and expertise when 50% of the population initially holds the expertise. *(top)* WS *(middle)* FF *(bottom)* SII

4 Discussion

The experimental protocol deployed for these computational studies enabled us to successfully run hundreds of thousands of simulation, each being run over a different network whom statistical properties were measured. The use of scientific workflows to run these computations appears as a relevant scheme to analyze the impact of networks on simulated dynamics ; the combination of the OpenMole software to coordinate the exploration of the parameters, of R/igraph to generate and analyze the networks, and of a simulation engine to simulate the social dynamics (here, NetLogo) stands as an efficient and reliable solution to drive these simulation experiments and enable their reproduction by peers.

Regarding the USA/IPK model under study, our simulation experiments confirmed the qualitative dynamics of the model are similar over different networks: importance of information seeking to trigger diffusion of innovation, asymetry between information search and proactive transmission, difficult diffusion when too much expertise is available. Any recommendation for policies based on the initial model would still stand after our computational study. Note that these experiments can not prove that the dynamics would be the same on any network; they just increase the likelihood of this statement. These results also do not demonstrate the dynamics of the USA/IPK model over the class of small-world networks, but only on the subset of small-world networks explored by the specific random network generators we selected (as highlighted [27]). However, our results demonstrated the efficiency of diffusion can be significantly different over different natures of networks (as demonstrated by the results over FF networks compared to WS and SII). In order to understand better the dynamics of word-of-mouth with information seeking, further research is needed on the space of parameters of the USA/IPK model. This includes implicit parameters such as the correlation of the various personality traits (for instance, maybe individuals who are curious are also more prone to be enthusiastic), correlation between personality traits and the structural properties of the network (maybe hubs in the network are individuals who have a bias towards curiosity), correlation between initial expertise and structural properties of the network (maybe hubs are more prone to hold initial expertise). In order to use the model to emit recommendations on targeting strategies, we should also explore the impact of the targeting of the communication campaign which is currently assumed to reach anyone in a random and uniform way; the timing of advertisement might also impact the knowledge dynamics.

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Review of Modern Physics **74**, 47–97 (2002). 5
2. Anderson, E.W.: Customer Satisfaction and Word of Mouth. Journal of Service Research **1**(1), 5–17 (1998). 2
3. Arndt, J.: Role of product-related conversations in the diffusion of a new product. Journal of Marketing Research **4**, 291–295 (1967) 2

4. Bailey, N.T.J.: The Mathematical Theory of Epidemics. London: Griffin, 1st edn. (1957) 2
5. Chatterjee, R.a., Eliashberg, J.: The Innovation Diffusion Process in a Heterogeneous Population: A Micromodeling Approach. Management Science **36**(9), 1057–1079 (sep 1990). 2
6. Csárdi, G., Nepusz, T.: The igraph software package for complex network research. InterJournal Complex Systems **1695**, 1–9 (2006). 8
7. Daley, D.J., Kendall, D.G.: Epidemics and Rumours. Nature **204**(4963), 1118 (1964) 2
8. Gilly, M.C., Graham, J.L., Wolfinbarger, M.F., Yale, L.J.: A Dyadic Study of Interpersonal Information Search. Journal of the Academy of Marketing Science **26**(2), 83–100 (1998). 2
9. Goffman, W., Newill, V.A.: Generalization of epidemic theory: an application to the transmission of ideas. Nature **204**, 225–228 (1964) 2
10. Goldenberg, J., Libai, B., Muller, E.: Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. Marketing Letters **12**(3), 211–223 (2001) 2
11. Goldenberg, J., Libai, B., Solomon, S., Jan, N., Stauffer, D.: Marketing percolation. Physica A: Statistical Mechanics and its Applications **284**(1-4), 335–347 (2000) 2
12. Granovetter, M.S.: Threshold models of collective behavior. American Journal of Sociology **83**(6), 1420–1443 (1978) 2
13. Katz, E., Lazarsfeld, P.: Personal Influence: The Part Played by People in the Flow of Mass Communication. Tech. rep., Bureau of Applied Social Research, Columbia University (1955) 2
14. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03 p. 137 (2003). 2
15. Kermack, W., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. In: Proc. R. Soc. Lond. A. vol. 115, pp. 700–721 (1927) 2
16. Kiesling, E., Günther, M., Stummer, C., Wakolbinger, L.M.: Agent-based simulation of innovation diffusion: A review. Central European Journal of Operations Research **20**(2), 183–230 (2012). 2
17. Kimura, M., Saito, K., Nakano, R., Motoda, H.: Extracting influential nodes on a social network for information diffusion. Data Mining and Knowledge Discovery **20**(1), 70–97 (2010). 2
18. Leskovec, J., Adamic, L.A., Huberman, B.A.: The dynamics of viral marketing. ACM Trans. Web **1**(1), 5 (2007). 2, 3
19. Leskovec, J., Kleinberg, J., Faloutsos, C.: Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD) **1** (2007) 7
20. Meade, N., Islam, T.: Modelling and forecasting the diffusion of innovation–A 25-year review. International Journal of Forecasting **22**(3), 519–545 (2006) 2
21. Peres, R., Muller, E., Mahajan, V.: Innovation diffusion and new product growth models: A critical review and research directions. International Journal of Research in Marketing **27**(2), 91–106 (2010). 2
22. Reuillon, R., Leclaire, M., Rey-Coyrehourcq, S.: OpenMOLE, a workflow engine specifically tailored for the distributed exploration of simulation models. Future Generation Computer Systems **29**(8), 1981–1990 (2013) 8
23. Rogers, E.M.: Diffusion of Innovations. New York: Free Press, 5th edn. (2003) 1, 2

24. Saito, K., Nakano, R., Kimura, M.: Prediction of information diffusion probabilities for independent cascade model. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **5179 LNAI**(PART 3), 67–75 (2008). 2
25. Sheth, J.N.: Word of Mouth in Low Risk Innovations. Journal of Advertising Research **11**(3), 15—-18 (1971) 1
26. Thiriot, S.: Word-of-mouth dynamics with information seeking : Information is not (only) epidemics. Physica A **492**, 418–430 (2018). 5
27. Thiriot, S.: Small world is not enough: Criteria for network choice and conclusiveness of simulations (2020) 14
28. Thiriot, S., Kant, J.D.: Using associative networks to represent adopters' beliefs in a multiagent model of innovation diffusion. Advances in Complex Systems **11**(2), 261–272 (2008) 1
29. Tisue, S., Wilensky, U.: Netlogo: A simple environment for modeling complexity. In: International conference on complex systems. vol. 21, pp. 16–21. Boston, MA (2004) 8
30. Trim, K., Nagji, N., Elit, L., Roy, K.: Parental Knowledge, Attitudes, and Behaviours towards Human Papillomavirus Vaccination for Their Children: A Systematic Review from 2001 to 2011. Obstetrics and gynecology international **2012**, 921236 (2012). 2
31. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 440–442 (1998) 5, 7
32. Young, H.P.: Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. The American Economic Review **99**(5), 1899–1924 (2009). 2