

Introducing Serendipity in a Social Network Model of Knowledge Diffusion

Marco Cremonini^a

^a*University of Milan, Department of Computer Science, Crema (CR), Italy*

Abstract

In this paper, we study serendipity as a possible strategy to control the behavior of an agent-based network model of knowledge diffusion. The idea of considering serendipity in a strategic way has been first explored in Network Learning and Information Seeking studies. After presenting the major contributions of serendipity studies to digital environments, we discuss the extension to our model: Agents are enriched with random topics for establishing new communication according to different strategies. The results show how important network properties could be influenced, like reducing the prevalence of hubs in the network's core and increasing local communication in the periphery, similar to the effects of more traditional self-organization methods. Therefore, from this initial study, when serendipity is opportunistically directed, it appears to behave as an effective and applicable approach to social network control.

Keywords:

serendipity, social networks, agent-based model, self-organization

*Corresponding author

Email address: `marco.cremonini@unimi.it` (Marco Cremonini)

1. Introduction

Models of knowledge and information diffusion in networks have been analytically studied since long, mostly in sociology, economics, and information science. It is inherently an interdisciplinary research since its roots, as evident in the classical Simon's study [32] of a class of distribution functions, the Yule-Simon distribution, which often occurs in nature, especially in social phenomena. Interestingly, modern studies of how ideas spread in a network of homogeneous agents demonstrated that the distribution is an extension of the Yule-Simon distribution [22]. Other important cumulative nonlinear effects of knowledge diffusion in networks, like the distribution of popularity of an idea [11, 22], have their roots in works like the de Solla Price's study of Cumulative Advantage distribution [30], which models situations where failure does not breed failure and success breeds success. Other important research have considered the relation between knowledge diffusion and network structure [8, 9], the emergence of collective behavior [6, 15], and self-organization [5, 31].

Another key research area for our work is related to learning and how to model the process of learning in an agent-based network. In Gale and Kariv model of social learning [14], the process of learning among agents depends on the network structure and is subject to local rules of agents behavior. De Laat *et al.* [21] studied how modern Social Network Analysis can improve networked and collaborative learning studies. In particular, they focused on the analysis of interaction patterns and community formation. Recently, Guechtouli *et al.* [16] characterized knowledge in an agent-based model as an array of stockpiles, meaning that a certain agent may have different knowl-

edge with different degree of expertise and agents choose the most competent agent in the population. A similar assumption was already made in the study on growth and diffusion of knowledge by Jovanovic and Rob [17], where heterogeneity of knowledge distribution in a network is key to its diffusion. Our agent-based model of knowledge diffusion [2,3] represents an extended version of Guechtouli’s direct transfer knowledge model by constraining the choice of the most competent agent to the network structure.

In this work we aim at addressing the unexplored relationship between *knowledge diffusion models* and *serendipity studies*. The concept of serendipity has become of interest for agent-based models of knowledge diffusion and network learning because, it is assumed, if properly controlled, it could be an effective tool for influencing the interactions between agents and the diffusion process, and for escaping an excessive tendency of social networks to homophily and homogeneity. Therefore, our motivating research question is: Could serendipity be introduced into a social network as a strategy for influencing the dynamic of the knowledge diffusion process and the network characteristics? The answer seems positive, from this initial study.

Simulations have been run based on different synthetic social networks produced with an extended version of the model fully described in [3]. Serendipity in agent behavior has been modelled as new topics *randomly* inserted in agents’ state during a simulation. Such a solution wish to represent the typical ”unexpected encounter” of serendipity and to modify both the agent’s criterion of choice of who to communicate with and how knowledge among agents is transmitted. Different strategies to induce a serendipitous behavior have been tested.

2. Background on Serendipity

2.1. Definitions of Serendipity

The concept of *serendipity* has a long history of attempts at defining it. Van Andel [35] defines it as "the art of making an unsought finding" and dates back its initial diffusion to literary circles of the 18th century. Defining serendipity as an "art" is a clear indication of the perceived elusive nature of the concept. More recently, with the diffusion of serendipity studies in some academic fields, definitions of serendipity have lost that "artistic nuance" in favor of more pragmatic descriptions for which chance is one important ingredient of a process involving the individual enjoying a serendipitous experience and the environment in which the individual lives.

It was only when serendipity has been recognized as interesting for Information Seeking research that more structured and less anecdotal studies appeared [4, 13]. Serendipity as "an unexpected experience prompted by an individual's valuable interaction with ideas, information, objects, or phenomena", as recently described by McCay-Peet *et al.* [27], is an example of more meaningful and elaborate definition. In that case, like in similar ones, the personal and unexpected nature of the experience is stressed, as well as the (perceived) value/benefit obtained and the many possible sources of serendipitous encounters (ideas, information, people, etc.)

The most important innovation with respect to the understanding of serendipity's nature introduced by Information Retrieval and Information Seeking studies was to consider it as a phenomenon and an experience to be pursued [13, 23]. If serendipity could be somehow facilitated or enabled, then the issue of *control* becomes relevant, although it shows some paradox-

ical elements [24]. Several scholars commented that if serendipity could be controlled, then an event is no longer serendipitous, but predictable or reproducible. Others have argued against this apparent contradiction noting that even though serendipity cannot be fully controlled being characterized by chance, the perception of serendipitous encounters could be enhanced as well as the odds of an unexpected positive event [4, 13, 18, 24, 33, 34].

Erdelez first contributed to serendipity research by introducing the notion of *information encountering* as a fortuitous way of acquiring information different from browsing or information seeking [12]. Her second contribution was to introduce four categories of *information users*: *non encounterers*, people that very seldom perceive to have experienced an information encounter; *occasional encounterers*, those that perceive to have information encounters from time to time, but credit just luck for those events; *encounterers*, those which often experience information encounters, and finally *super-encounterers*, people that recognize in information encountering a suitable and effective strategy for acquiring information, a fruitful complement of active information seeking and browsing.

2.2. Serendipity in Networks and Knowledge Diffusion

Still few studies have directly addressed the relation between serendipity and a digital environment: How are serendipitous events supported by the environment and which are the most relevant features of the context? The interest in these research questions has grown in recent years, in particular in the area of Networked Learning [1, 19, 20], where serendipity as a possible strategy for improving the diffusion of knowledge or ideas has become an active research topic.

In Networked Learning studies, the focus recently shifted from predetermined learning systems to "the act of learning as a response to changes in the learning environment" [7]. Interestingly, Kop noted that "people will first and foremost find information from people with whom they have a strong relationship" [20], which is the narrative definition of the typical community formation of social networks based on proximity and friends-of-friends relationships. Another relevant observation by Kop is that not all nodes are equal when network learning is considered, therefore it is likely that environments better supporting serendipity would handle nodes with different characteristics in different ways. Finally, crucial with respect to our work on *agent-based models of social network* is Kop's conclusion that "facilitating more randomness in our information stream" is a challenging goal in search strategies. Correspondingly, also Makri *et al.* [23], discussing serendipity in the field of Information Science and Technology, draw a similar conclusion: "A new way of thinking about how to support serendipity in digital environments involves moving away from trying to serve up serendipity itself and towards empowering users to create their own personal recipes for it by supporting strategies that may increase its likelihood.". *Randomness in information stream*, as we will see, is precisely the mechanism we have introduced and studied in our agent-based model.

Works on *Social Network Analysis* (SNA) with a focus on serendipity are almost absent in the literature. To our knowledge, ours is the first attempt at explicitly modeling the concept of serendipity in an agent-based social network based on previous works in the Networked Learning. We are aware of just a single relevant work, by Creixell *et al.* [10], which adopted an ele-

mentary definition of serendipity, missing its tie with complexity science, but nevertheless interesting as the first attempt to explicitly model the relation between serendipity and a social network. That work presents some results about the likelihood of serendipitous events and network properties of agents. Their most relevant conclusion is that the agents for which the likelihood of serendipitous information encounter is higher are those whose *node degree is high* and the *clustering coefficient is low*. This combination means that an agent is very active in his/her social network, but his/her peers are not much connected among themselves. In our work we have considered this result and used for a specific configuration of our model, to test its effectiveness in our setup.

3. Original Model

In this section we provide a summary of the characteristics of our original agent-based model, before the extensions we made to test serendipity strategies. A more detailed description and analysis can be found in [3].

We assume a set of agents and a set of *topics* to be given. Each agent has a certain level of *interest* and skill (*quality*) on each topic, both change through interactions with other agents. In more detail, we consider a set of N agents, n_1, n_2, \dots, n_N , each one characterized by a *Personal state* PS_{n_i} (what n_i knows) and a *Friend state* FS_{n_i} (who n_i knows). The *Personal state* has the form $PS_{n_i} = (\bigcup_{j \in T_i} (\text{topic}_j, \text{quality}_{i,j}, \text{interest}_{i,j}))$, where T is the set of topics that the population knows; each agent n_i knows a variable subset of them $T_i \subseteq T$. The *Friend state* has the form $FS_{n_i} = (\bigcup_{j \in N_i} (n_j, \text{answers}_{i,j}))$, where n_j are the identifiers of agents connected with n_i and $\text{answers}_{i,j}$ is a

counter to keep track of the number of interactions with each peer. The setup has been defined to be the most neutral, with topics T_i assigned to each agent and associated qualities selected randomly, interests distributed uniformly and no connection.

3.1. Network Construction

A network is dynamically formed according to the following steps:

1. At each tick, an agent $n_{i'}$ is selected, then a topic ($topic_{j*}$) is selected from its Personal state. The choice of the topic is a weighted random selection with values of the associated interests ($interest_{i,j*}$) as weights, this way topics with higher interest are more likely to be selected;
2. Among $n_{i'}$ "friend" agents and their "best friend" holding topic ($topic_{j*}$), select agent $n_{i''}$ with maximum value of topic's quality ($quality_{i,j*}$);
3. If $quality_{i'',j*} > quality_{i',j*}$ then the communication takes place and agent $n_{i'}$ increases $quality_{i',j*}$ of $topic_{j*}$;
4. Otherwise, if none holds $topic_{j*}$ or exhibits a topic's quality greater than that of agent $n_{i'}$, then select an agent $n_{i'''}$ randomly among the whole population;
5. if $n_{i'''}$ holds $topic_{j*}$ and $quality_{i''',j*} > quality_{i',j*}$, then the communication takes place and $quality_{i',j*}$ increases, otherwise the communication fails.

Best "friend-of-friends". Given agent $n_{i'}$, and a selected $topic_{j*}$, for each of its friends, the "best friend" agent is the one owning $topic_{j*}$ and the higher value of the attribute *answer*. The reason for this solution is that we consider unrealistic in a social context to scan all agents with a distance of 2

from the one selected. The selection based on the *answer* attribute represents a basic form of transitive trust. It is worth noting that the inclusion of “best friends” fosters network transitivity and the formation of triads, two key characteristics of social networks.

Start up. At start up, agents have no connection (i.e., Friend state is empty). When, for an agent, the 5-steps algorithm is executed, a topic is selected in *Step1*, then *Step2* and *Step3* fail and in *Step4* a random agent is selected. If *Step5* succeeds, then the connection is established. This mechanism triggers the network formation at start up.

In Figure 1, an example of topic and peer selection is presented.

1. for node n_{10} , $topic_{15}$ is selected;
2. friends (n_4, n_{17}, n_{20}) own $topic_{15}$ (depicted in the grey circle);
3. best friends-of-friends (n_{16}, n_{24}, n_{30}) own $topic_{15}$ and have maximum number of answers (depicted in the white circle);
4. node n_{i^*} with higher quality associated to $topic_{15}$ is selected. $n_{i^*} = \{n_i : \max(quality_{i,15}; i \in (n_4, n_{17}, n_{20}, n_{16}, n_{24}, n_{30}))\}$;
5. **IF** $quality_{i^*,15} > quality_{10,15}$
THEN

- case 1* (n_{10} and n_{i^*} were not friends):
 Update n_{10} 's Friend state by adding the pair $(n_{i^*}, 1)$ and n_{i^*} 's Friend state by adding $(n_{10}, 0)$;
 Update n_{10} 's Personal state.
- case 2* (n_{10} and n_{i^*} were already friends):
 Update only n_{10} 's Friend state by increasing $(n_{i^*}, answers_{10,i^*} + 1)$;
 Update n_{10} 's Personal state.

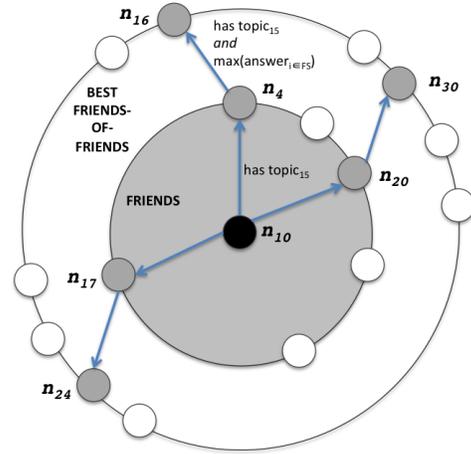


Figure 1: Example of topic and peer selection.

3.2. State Update

After a successful interaction, the agent that started the communication is updated. For model simplicity no change in the respondent's state is produced, because knowledge, being an intangible good, does not decrease when shared, and we assume no cost for the transmission. The *quality* and the *interest* of the topic for which the communication took place increase and the other interests, associated to topics owned by the agent, are decreased. We decided that the topic's quality increases with increasing marginal increments, according to the assumption that an agent distrusts another one when they interact for the first time and this distrust progressively diminishes as interactions occur. The discount starts at a given value (i.e., ρ) and goes to zero exponentially. Motivations for this assumption could be found in the literature about information aggregation [6] and collective behavior [15] and refers both to the prevalence of egocentrism in assimilating new information and to trust dynamics. The *quality gain* obtained by agent $n_{i'}$ is:

$$\delta quality_{j*} = \frac{quality_{i'',j*} - quality_{i',j*}}{\gamma + \rho e^{-\frac{x}{\theta}}} \quad (1)$$

with: $\gamma \geq 1$ setting the nominal fraction of $\delta quality$ that $n_{i'}$ could learn from another agent; x the value of the attribute *answers* representing the number of past interactions that agent $n_{i'}$ had with agent $n_{i''}$; ρ the initial discount; θ the factor that controls the rate at which agents increase their trust towards the others.

The dynamics we have assumed for the *interest* associated to the topic for which the interaction took place is similar to that of the quality, but with two important differences: It only depends on the $\delta quality$ value and, accord-

ingly, all other interests on topics owned by the agent decreases (studies in cognitive science have shown the tendency of people to shift their attention and interest, rather than behave incrementally [15]). The function is:

$$\delta interest_{i',j*} = \alpha(1 - e^{-\frac{\delta quality_{i',j*}}{\beta}}) \quad (2)$$

with $\alpha > 1$ and $\beta > 1$ the two parameters that control, respectively, the scale and slope of the interest increase.

Finally, all interests associated to topics different from $topic_{j*}$ are reduced by $\delta interest_{i',j \neq j*}(t_k, t_{k-1}) = \delta interest_{i',j*}(t_k, t_{k-1}) / (|T_{i'}| - 1)$, that is the value of the interest gain for $topic_{j*}$ at t_k divided by the number of topics $|T_{i'}|$ minus one.

4. A Serendipitous Social Network Model

4.1. Serendipity as a control strategy

In considering serendipity as a viable strategy for our model of knowledge diffusion, we started from the proposed five facets of a potentially serendipitous digital environment introduced by McCay-Peet *et al.* [25–27]. The five facets definition represents indeed a remarkable leap from the early works on serendipity and perhaps the first well-defined model of Serendipitous Digital Environment:

- *Enables exploration: The degree to which a digital environment supports exploration of its information (or resources).* Our model allows the exploration of neighborhood agents through the selection of topics, which drives the following selection of the best peer among friends and friends-of-friends.

- *Trigger-rich: The degree to which a digital environment contains a variety of information that is interesting and useful to an agent.* The ability of our agents to encounter all topics defined within the system depends on the initial setup, especially on the rate $\frac{N}{T}$ (the number of agents over the number of topics), the rate $\%T$ agents can hold and on the network topology.
- *Highlights triggers: The degree to which a digital environment brings interesting and useful information to an agent's attention.* The interest of an agent in a topic is a parameter of the model and is used as a weight in topic selection; there is no mechanism to "facilitate" some topics over others.
- *Enables connections: The degree to which a digital environment makes relationships or connections between information apparent.* Connections are between agents and all topics are independent; opportunities to connect between agents may go unnoticed due to the limited knowledge of all possible topics.
- *Leads to the unexpected: The degree to which a digital environment provides opportunities for unexpected interactions with information.* Originally, randomness was unlikely during the dynamic evolution except at start up; the model extension introduces randomness during the process of network creation.

Given this steps, the meaning of serendipity in our model could be twofold: Either some agents are enabled to communicate with the others *unexpectedly* or obtain *non-deterministic* benefits from successful exchanges, or some

agents systematically get to know topics not already known and by means of them find new peers. The former two cases would mean, respectively, that either the network construction mechanism based on neighborhood (friends and friends-of-friends) should have been relaxed and a different rewiring technique introduced, modifying one of the most basic feature of our social network model, or the functions controlling *quality* and *interest* made non-deterministic. With the latter, instead, it must be relaxed just the assumption holding that the only mechanism to select topics is probabilistic with *interests* as weights. This consideration, together with the fact that by design the selection of a peer is driven by the previous selection of a topic and not vice versa, has motivated the choice of excluding the first two options and concentrate our study on the third one.

4.2. Model modifications

The original model was modified with the aim of introducing a degree of randomness in topics owned and then possibly selected by agents. The choice wants to simulate the systematic encountering of information previously unknown, influencing the weighted probability mechanism implemented so far. To this end, some agents periodically receive new topics, according to different strategies (another way to consider this solution, is to see the addition of random topics as an *exogenous shock* that an agent receives altering its internal state). The combination of random new topics and the mechanism based on interests for establishing new communication implements our notion of serendipity into the model. This is a minimal definition of serendipity that, however, permits to study different strategies based on three features:

- *Randomness*: Two rates of random topics, *low* and *high* (i.e., 1% or 30%)¹ are added to agents selected, according to different criteria.
- *Agent selection*: Three criteria have been considered:
 - *Lower Node Degree-Higher Clustering Coefficient*: Agents are ranked based on the combination of *lower node degree* and *higher clustering coefficient*, according to the result discussed by Creixell *et al.* in [10].
 - *Higher Node Degree*: Agents are ranked in decreasing order based on their node degree value².
 - *Random*: Agents are randomly selected.
- *Topology dependence*: Different network topologies behave differently. Four stylized configurations (i.e., configuration **A**, **B**, **C**, and **D**, described in the following section) have been compared.

As a result of the definition of these features and of its combination, we studied cases defined in Table 1.

¹The choice of 1% and 30% was made for clarity of presentation among the many configurations tested and it represents a very low and a very high rate of topics added at random to selected agents (in simulations, 1% always means that a single topic is added at random per agent).

²Other metrics have been tested (i.e., closeness, betweenness and eigenvector) here omitted for sake of presentation.

Table 1: Case studied characterized by network configurations, number of topics added to agents, and number of agents selected.

<i>Configuration</i>	<i>Topics</i>	<i>Agents</i>	<i>Description</i>
$A_r/B_r/C_r/D_r$ 100-1	1%	100%	1% of topics held by each agent are randomly added.
$A_r/B_r/C_r/D_r$ 10-30	30%	10% random	30% of topics held by 10% of randomly selected agents.
$A_c/B_c/C_c/D_c$	30%	10% degree	30% of topics held by 10% of agents ranked by decreasing node degree.
$A_{dc}/B_{dc}/C_{dc}/D_{dc}$	30%	10% deg-clust	30% of topics held by 10% of agents ranked with respect to lower node degree and higher clustering coefficient.

5. Simulation Configurations and Results

During simulations, the number of agents $N = 100$ is in the same order of magnitude of a typical online community of collaborating students. The uniform distribution for the introduction of random topics, represented by the fixed number of ticks $k = 1000$, has been chosen being the more neutral hypothesis, given the lack of experiences and references in the literature. The total number of topics is $T = 20$ for configurations A/B, $T = 100$ for configurations C/D. Agents in configurations A/C are assigned with at most 10% of T at setup, agents in configurations B/D could reach 100% of T. All other important parameters were the same used in testing self-organizing strategies described in [3] that define the four base configurations (see Figure 2). Configurations A and B form the typical giant component, while configurations C and D present disconnected communities. From these configurations, the following serendipity strategies have been tested:

Random 10-30/Node Degree/Degree-Clustering: For each base configuration we add 30% of new topics to the 10% of agents respectively selected *randomly* (configurations $A_r/B_r/C_r/D_r$ 10-30), ranked with respect to *decreasing node degree* (configurations $A_c/B_c/C_c/D_c$), or ranked with respect the *lower degree-higher clustering* criterion (configurations $A_{dc}/B_{dc}/C_{dc}/D_{dc}$).

Random 100-1: For each base configuration we apply uniformly to all agents a low level of serendipity (i.e., 1% of new topics added to 100% of nodes every k ticks).

Figure 2 presents the dynamics of *Average Network Degree* and *Average Clustering* comparing the results for the base configurations A/B/C/D and the three strategies selecting 10% of agents and adding 30% of random topics.

For the strategy *Random 100-1*, instead of a general overview of the results, we take a closer look at the particular case of Configuration C that exhibits a peculiar behavior (see Figure 3).

Considering Figure 2, a first result is evident for the clustering coefficient that consistently increases with respect to the original setup for configurations A and C. These are the setup forming the typical giant component, which in social networks tend to have high clustering in the core, almost fully connected, and lower clustering in the periphery. The effect of adding random topics is, in general, to facilitate the transitivity of connections among neighbour agents in the periphery of the network. This effect, in general, improves the characteristic of collaborating networks [29]. The same effect is less evident for configurations B and D producing fragmented networks instead of a giant component. For these networks the clustering coefficient is originally high, so the impact of serendipity on clustering is limited.

With respect to the average degree, qualitative differences could be seen for the different strategies. In particular, we observe that selecting agents with high degree and low clustering (labelled as *Conf_{--dc}*) does not seem to produce any improvement over the random selection (labelled as *Conf_{--r}*). The explanation for our result and the ones described in [10] is that agents more likely to experience serendipitous events are those with low clustering because these belong to the periphery of a social network, where triangle density is lower and then homophily is likely lower, combined with high degree, because they have more collaborations with non-neighbors, likely more heterogeneous. However, stimulating these agents with random topics has mostly the effect of increasing local collaborations, thus higher clustering, similar to a random selection of agents, given that the majority of agents belong to the periphery of the network.

Different is the strategy of selecting agents with higher degree (labelled as *Conf_{--c}*) that in configurations A and C shows an average degree remarkably lower than the others (and correspondingly, a slightly greater value of the clustering coefficient). In this case, agents stimulated with random topics include the hubs of the original configuration that typically lay in the core of the network. Induced serendipity has thus the effect of mitigating the prevalence of hubs in the core while fostering local connections in the periphery. The overall effect of serendipity is then to improve small-world properties (i.e., low average degree and high clustering) [28].

In configurations B and C, all strategies produce a low average degree with respect to original configurations, because due to the limited size of the networks, adding new topics prevent the formation of hubs and foster

triangle closure.

Figure 3 presents the result of a special case that emerged when the *Random 100-1* strategy was applied (the result has been confirmed through several rounds). For configurations A, B, and D the effect of just adding a single random topic to each agent every k ticks (a simplified scenario we used as a theoretical case) do not exhibit interesting properties with respect to previous strategies. Configuration C is a special case, though: The low but diffuse rate of randomness strongly influenced the dynamics that produces the giant component by keeping the average degree of agents at very low values and clustering coefficient at very high values. This means that the communication among agents is local, most of the time. The effect was destructive because the formation of the giant component was prevented and the network remains fragmented in more connected components (see the *Connected Components* chart of Figure 3).

From these qualitative tests, although not conclusive, we could derive some observations. The first is that forcing a degree of randomness in the mechanism of social network creation stimulates primarily the typical triangle closure in the periphery of the giant component, which is at the base of a more localised transmission of information. This has the corresponding effect of mitigating the prevalence of hubs in the core of the network. This give us a first tentative practical strategy for using serendipity to influence a social network dynamics as a means to reduce the influence of hubs in polarising the diffusion of knowledge. In this case, serendipity may act as a factor contributing to a transmission of information more homogeneous throughout the whole network rather than mainly driven by hub agents.

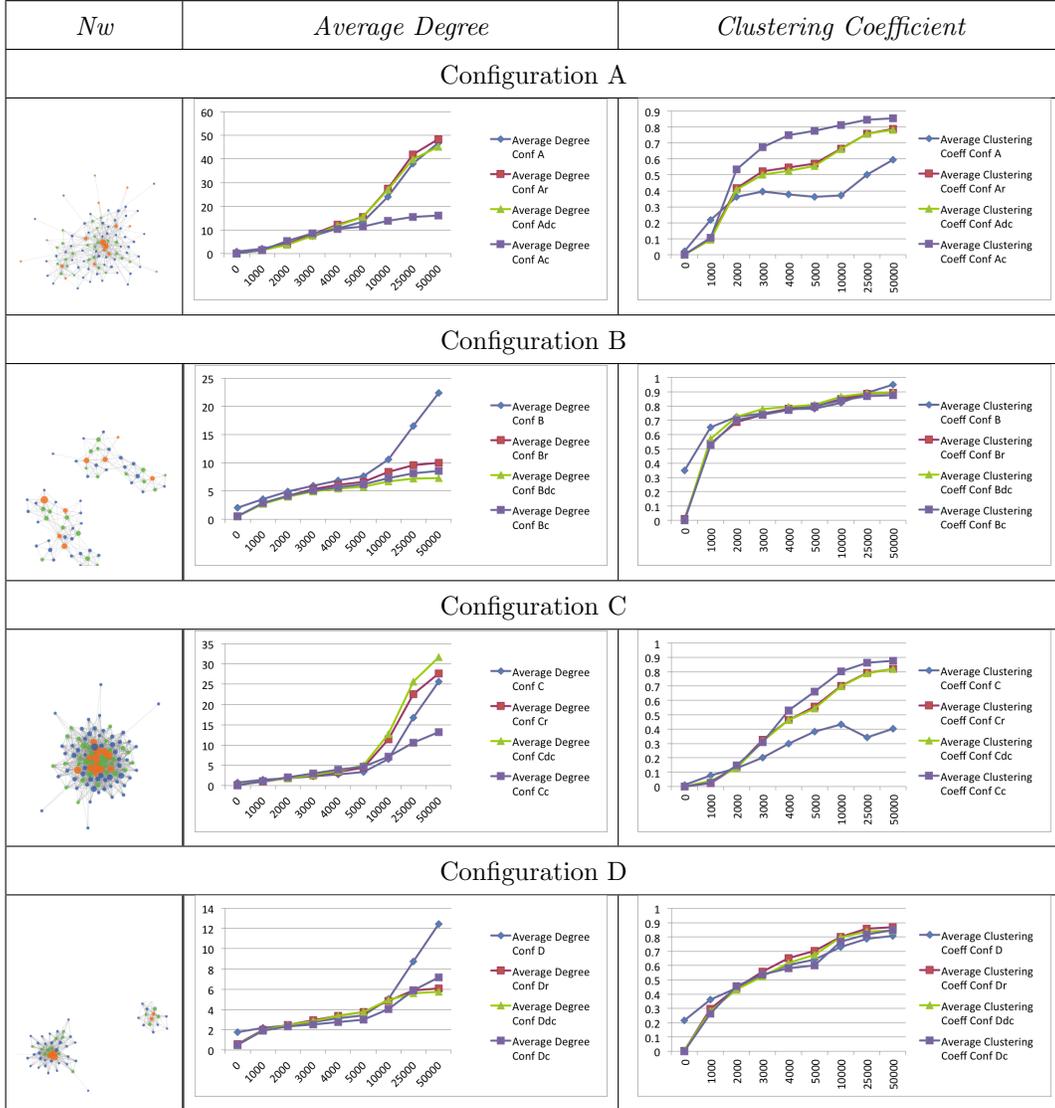


Figure 2: Networks in the first column represent the typical outcome of base configurations. Second and third columns show results of serendipity strategies based on *random selection of nodes*, *centrality metric ranking* or *degree-clustering criterion* compared to base configurations. Duration of the simulation $\Gamma = 50000$ ticks, number of agents $N = 100$, total number of topics $T = 20$ for configurations A/B, $T = 100$ for configurations C/D. Agents in configurations A/C are assigned with at most 10% of T at setup, agents in configurations B/D could reach 100% of T .

However, this does not come without side-effects, as we have seen in Figure 3. An uncontrolled application of randomness to agent behavior has prevented the development of the giant component. Usually this is a severe drawback for diffusion processes and must be avoided. However, we can think to situation where instead it is exactly the formation of a giant component the event to be avoided (e.g., a network of malware, of criminals, or of colluded peers).

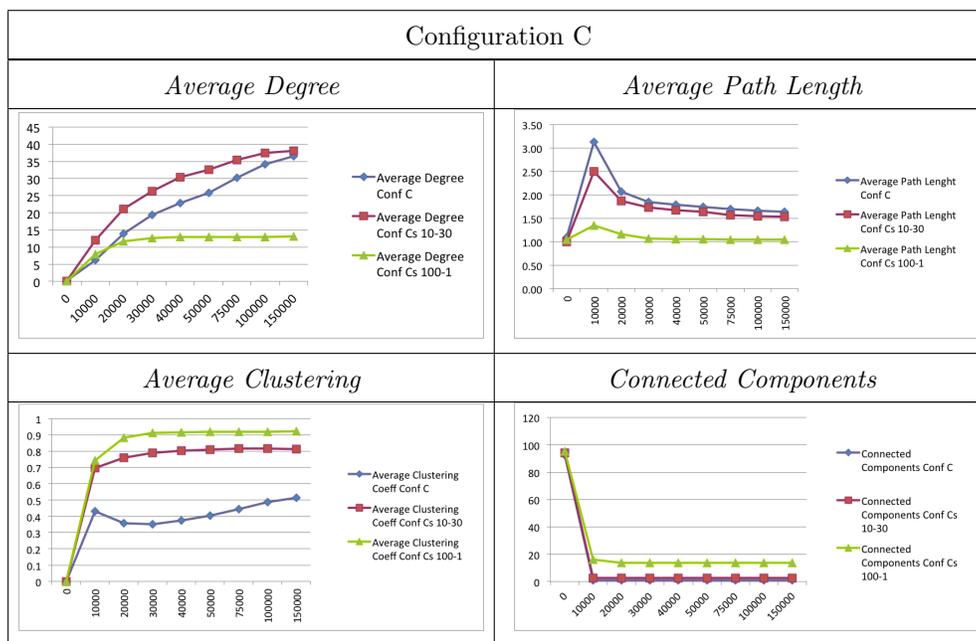


Figure 3: Configuration C results of serendipity strategy with 1% of new random topics added to 100% of nodes compared to strategy with 30% of new random topics added to 10% of randomly selected nodes. Duration of the simulation $\Gamma = 150000$ ticks.

6. Conclusions

In this paper we have discussed the concept of serendipity suggesting that it could have an important role in Social Network Analysis as a strategy for controlling, or just nudging the behavior of a social network.

For the possible application of serendipity strategies to real case studies, we believe that Network Learning is one of the most promising. In particular, we imagine a possible challenging case study in online classes in which students interact by means of forums, for example, posting and answering questions, proposing new issues or discussing about topics related to the course. In such a scenario, the interactions establish a social network among students, with emergent roles depending on individual characteristics such as curiosity and competence on a specific matter.

In such a modern learning environment the role of the instructor is tricky: Should he/she just monitor the correctness of the answers keeping a low profile or should he/she actively engage in discussions and answer to questions or even should he/she propose new arguments aiming at stimulating reflection, critical thinking, or curiosity? The risk is that too much external control of the discussion environment by an instructor may result in reflecting just his/her personal preferences and opinions (he/she becomes a hub), undermining the freedom of exploring of the students. On the contrary, a discussion left on its own between the students may derange in flames or towards out-of scope topics (a fragmented network) or be dominated by the interests of the most vocal or expert among the students (again, a few hubs dominate the diffusion process).

Other possible applications may be found for recommendation systems,

web browsing and also online news platforms, all areas which have been recognized since long the problem of excessive homogeneity, lack of cultural difference or information source heterogeneity.

Finally, as mentioned, we may envision potentially useful applications of the possible destructive effect of serendipity on control processes for contrasting networked threats, coordinated attacks or collaborative criminal networks.

References

- [1] Alisa Acosta, *Using serendipity to advance knowledge building activities*, Ontario Institute for Studies in Education, University of Toronto, Canada (2012).
- [2] Luca Allodi, Luca Chiodi, and Marco Cremonini, *The asymmetric diffusion of trust between communities: Simulations in dynamic social networks*, Proceedings of the 2011 Winter Simulation Conference (WSC 2011), 2011.
- [3] ———, *Self-organizing techniques for knowledge diffusion in dynamic social networks*, Proceedings of Complex Networks Conference 2014 (CompleNet14), 2014.
- [4] Paul André, Jaime Teevan, Susan T Dumais, et al., *Discovery is never by chance: designing for (un) serendipity*, Proceedings of the 7th ACM Conference on Creativity and Cognition, 2009, pp. 305–314.
- [5] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, *Science* **5439** (1999), no. 286, 509–512.
- [6] L. M. A. Bettencourt, *The rules of information aggregation and emergence of collective intelligent behavior*, *Topics in Cognitive Science* **1** (2009), no. 4, 598–620.
- [7] Paul Bouchard, *Network promises and their implications*, RUSC. Universities and Knowledge Society Journal **8** (2011), no. 1.
- [8] R. Cowan and N. Jonard, *Knowledge creation, knowledge diffusion and network structure*, *Economies with Heterogeneous Interacting Agents*, 2001, pp. 327–343.

- [9] ———, *Network structure and the diffusion of knowledge*, *Economic Dynamics and Control* **8** (2004), no. 28, 1557–1575.
- [10] Werner Creixell, Juan Carlos Losada, Tomás Arredondo, Patricio Olivares, and Rosa María Benito, *Serendipity in social networks.*, *NHM* **7** (2012), no. 3, 363–371.
- [11] George Ehrhardt, Matteo Marsili, and Fernando Vega-Redondo, *Diffusion and growth in an evolving network*, *International Journal of Game Theory* **34** (2006), no. 3, 383–397.
- [12] Sandra Erdelez, *Information encountering: It's more than just bumping into information*, *Bulletin of the American Society for Information Science and Technology* **25** (1999), no. 3, 26–29.
- [13] Allen Foster and Nigel Ford, *Serendipity and information seeking: an empirical study*, *Journal of Documentation* **59** (2003), no. 3, 321–340.
- [14] D. Gale and S. Kariv, *Bayesian learning in social networks*, *Games and Economic Behavior* **45** (2003), no. 2, 329–346.
- [15] R. L. Goldstone and T. M. Gureckis, *Collective behavior*, *Topics in Cognitive Science* **1** (2009), no. 3, 412–438.
- [16] W. Guechtouli, *Agent-based modeling of knowledge transfer within social networks*, 2014.
- [17] B. Jovanovic and R. Rob, *The growth and diffusion of knowledge*, *The Review of Economic Studies* **56** (1989), no. 4, 569–582.
- [18] Ben Kirman, Conor Linehan, and Shaun Lawson, *Get lost: facilitating serendipitous exploration in location-sharing services*, *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, 2012, pp. 2303–2308.
- [19] Rita Kop, *Information aggregation in networked learning: The human factor and serendipity*, 8th Int. Conf. on Networked Learning, 2012.
- [20] ———, *The unexpected connection: Serendipity and human mediation in networked learning*, *Journal of Educational Technology & Society* **15** (2012), no. 2, 2–11.

- [21] M. De Laat, V. Lally, L. Lipponen, and R. J. Simons, *Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis*, Computer-Supported Collaborative Learning **1** (2007), no. 2, 87–103.
- [22] François Lafond, *Self-organization of knowledge economies*, Journal of Economic Dynamics and Control **52** (2015), 150–165.
- [23] Stephann Makri, Ann Blandford, Mel Woods, Sarah Sharples, and Deborah Maxwell, *Making my own luck: Serendipity strategies and how to support them in digital information environments*, Journal of the Association for Information Science and Technology **65** (2014), no. 11, 2179–2194.
- [24] Abigail McBirnie, *Seeking serendipity: the paradox of control*, Aslib Proceedings, 2008, pp. 600–618.
- [25] Lori McCay-Peet and Elaine G Toms, *The process of serendipity in knowledge work*, Proceedings of the 3rd Symposium on Information Interaction in Context, 2010, pp. 377–382.
- [26] ———, *Investigating serendipity: How it unfolds and what may influence it*, Journal of the Association for Information Science and Technology (2015).
- [27] Lori McCay-Peet, Elaine G Toms, and E Kevin Kelloway, *Examination of relationships among serendipity, the environment, and individual differences*, Information Processing & Management **51** (2015), no. 4, 391–412.
- [28] Alan Mislove, Massimiliano Marcon, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee, *Measurement and analysis of online social networks*, Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, 2007, pp. 29–42.
- [29] Mark EJ Newman, *The structure and function of complex networks*, SIAM Review **45** (2003), no. 2, 167–256.
- [30] DJ Price de Solla, *A general theory of bibliometric and other cumulative advantage process*, Journal of the American Society of Information Science **27** (1976), 292–306.

- [31] J. J. Ramasco, S. N. Dorogovtsev, and R. Pastor-Satorras, *Self-organization of collaboration networks*, Physical Review E **3** (2004), no. 70, 036106.
- [32] H. Simon, *On a class of skew distribution functions*, Biometrika Trust **42** (1955), no. 3/4, 425–440.
- [33] Jennifer Thom-Santelli, *Mobile social software: Facilitating serendipity or encouraging homogeneity?*, IEEE Pervasive Computing **3** (2007), 46–51.
- [34] Elaine G Toms, *Serendipitous information retrieval.*, DELOS Workshop on Information Seeking, Searching and Querying in Digital Libraries, 2000.
- [35] Pek Van Anandel, *Anatomy of the unsought finding. Serendipity: Origin, history, domains, traditions, appearances, patterns and programmability*, The British Journal for the Philosophy of Science **45** (1994), no. 2, 631–648.