Modeling and simulating the emergence of Internet communities: impact of the spread of memes and agent memory

Aida Huerta-Barrientos[1], Rémi Lamarque[2], Stephen Leese[3]

1 National University of Mexico, e-mail: aida.huerta@comunidad.unam.mx

2 Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France, e-mail: remi.lamarque@lpl-aix.fr

3 Davenport, Iowa, e-mail: whatcroft@gmail.com

Abstract

The spreading of memes on the web played an important role in the emergence of Internet communities. The principal purpose of the study is to implement a simulation model to analyze the process of emergence of Internet communities. The model shows the importance of factors such as the interactions between online agents as well as their propensity to adopt and remember new memes. In addition, it explores the threshold between isolated cultural short-lived trends and the viral spreading of many cultural features. The simulation model is an agent-based model built using NetLogo™ software, designed such that agents represent Internet users and memes are represented by features appearing randomly on each agent. The model illustrates the spread across the whole network as shown through interactions of agents indicating further, that the structure of the network, especially the number of indegree and outdegree links between agents, has a crucial influence on how many memes are shared among agents in the long run. In other words, a greater connectivity leads to the quick sharing and sustainability of several cultural features, which is the basis for the emergence of a community.

Key words: memes, agent-based modeling, Internet communities, complex adaptive systems.

1. Introduction

A meme was firstly defined by Dawkins (1976) in his book *The Self-ish Gene* as any cultural entity that appears to exhibit self-replication and mutation but using human culture as their medium of propagation, such as words, melodies, recipes, thoughts and ideas, so they can be transmitted from person to person via social interactions. From the evolutionary perspective, memes are subject to a form of natural selection, which allows them to change over time and even evolve to become more effective replicators (Wang and Wood, 2010). Memes thus have their own reproductive fitness, just like a virus and can spread to fixation in a single decade, and become nearly extinct in the same time frame, an evolutionary process much too swift to be discernible in any increase or decrease in human offspring (Dennett, 2017).

On the other hand, as Bauckhage (2011) explains: Internet meme refers to the phenomenon of content or concepts that spread rapidly among Internet users by means of email, instant messaging, forums, blogs, or social networking sites. Following Bauckhage (2011), Internet memes usually consist on offbeat news, websites, catchphrases, images, or video clips and typically evolve through commentary, imitations, parodies, or even through related news in other media. Similarly, Shifman (2013) defined Internet memes as "units of popular culture that are circulated, imitated, and transformed by Internet users, creating a shared cultural experience". More recently, Segev *et al.* (2015) conceptualized Internet memes as families of texts that share a similar quiddity, each meme family also shares content and form characteristics with other meme families. Segev *et al.* (2015) suggested that meme families are organized together into larger networks, shaping the digital culture at large. The most important aspect that Segev *et al.* (2015) take into account in the meme propagation process is the role of social context. As consequence, the social network structure plays a crucial role for understanding the diversity of memes (Weng et al., 2012). The main properties of memes are based on their frequency of occurrence and the degree to which they propagate (Kuhn et al., 2014).

From the social identity approach, the spread of memes is affected by the flow of influence (e.g. high status versus low status) in a group (Mazambani *et al.* (2015). Thus, memes that are introduced by high-status members of an Internet community are more likely to be spread because low-status community members are seeking information and approval from higher-status in-community members (Dino, Reysen, & Branscombe, 2009). According to Blackmore (1999), an Internet meme's survival depends on how easily it can be transmitted from one brain to another. It means that memes that can be easily expressed through the Internet network will have a higher chance of surviving since spread depends on transforming the memory of the idea of Internet users into a form that other Internet users can imitate, such as speech, writing, or actions (Blackmore, 1999).

Research on memes has shown that the evolution of memes can be exploited effectively for inferring networks of diffusion and influence (Gomez Rodríguez *et al.,* 2010) and that information contained in memes is evolving as it is being processed collectively in online social media (Simmons *et al.,* 2011). More recent investigations into memetic communication have focused on Internet memes to understand cultural trends and digital culture, participation in communities and spreadability (Mazambani *et al.,* 2015).

Several top-down approaches have been used to characterize the dynamics of memes, mainly focusing on quantifying the topological centrality, content similarity, or user behaviors based on large-scale datasets (He *et al.,* 2015). Leskovec *et al.* (2009) developed a framework for tracking short, distinctive phrases that travel relatively intact through on-line text. Developing scalable algorithms for clustering textual variants of such phrases, they identified a broad class of memes that exhibit widespread and rich variation on a daily basis. Leskovec *et al.* (2009) showed how such a meme-tracking approach provided a coherent representation of the news cycle — the daily rhythms in the news media that have long been the subject of qualitative interpretation, but have never been captured accurately enough to permit actual quantitative analysis. Similarly, several bottom-up approaches attempt to develop diffusion models using analogy to behavior replication, epidemic contagion or competitive gaming (He *et al*., 2015). For instance, Wang and Wood (2010) adopted an epidemiological approach to develop a model of viral meme propagation, based on the modified SIR model of infectious diseases. The implementation in the modelling of meme spread as reflected in Internet search data showed that memes may be treated as infectious entities when modelling their propagation over time and across societies. Bauckhage (2011) investigated the epidemic dynamics of 150 famous Internet memes. The analysis was based on time series data that were collected from Google Insights, Delicious, Digg, and StumblerUpon. Bauckhage (2011) found that differential equations models from mathematical epidemiology as well as simple log-normal distributions gave a good account of the growth and decline of memes. Mazambani *et al.* (2015) examined the influence of meme consistency and intragroup status on the spread of memes in virtual communities. They analyzed social interactions among members of four online forums from January 2010 to February 2014. Their results showed that memes initiated by low-status members spread faster than memes started by high- or moderate-status members.

According to Leskovec *et al.* (2009), tracking new topics, ideas, and ‘memes’ across the web has been an issue of considerable interest in the last years. Internet memes as pieces of digital content that spread around the web in various interactions, have recently sparked scholarly across several fields (Segev *et al*., 2015; Bennet and Segerber, 2012). However, research on the spread of memes on Internet communities traditionally has used an epidemiological approach. Although considerable progress has been made in the study of memes propagation, many important issues remain unexplored as for instance, the impact of the agent memory on the rate transmission of memes over different Internet communities. The principal purpose of the study is to model the process of emergence of Internet communities from the spread of memes through the network, that underlines the importance of factors such as the interactions between online agents, which generate a specific social network structure that is crucial for understanding the diversity of memes, as well as their propensity to adopt and remember new memes, generating competition among them that can bring the network at the brink of criticality. We consider that this study may contribute towards a better understanding of the spreading process of memes on Internet, represented by features appearing regularly on any random agent.

This study is divided into five main sections. Firstly, a conceptual model of Internet communities as a complex adaptive system is developed. Secondly, an agent-based simulation model, implemented using Netlogo^{TM} software, illustrates the spread of memes considering agents with memory. After that, three simulation scenarios are designed considering three different configurations of the network structure in order to analyze the impact of the network structure on the spread rate, specifically the competition among memes to determine their diversity. The concluding remarks are drawn and some ideas for future research about modeling and simulation of spread of memes through Internet are shown.

2. Internet communities modeled as a complex adaptive system

The concept of complex adaptive systems (CAS) was first introduced by Walter Buckley in 1967 and refers to systems that are composed by multiple interrelated fundamental elements interacting in a non-linear way whose structure is based on hierarchical levels. The interrelated elements are complex in nature and are characterized by significant internal properties that are highlighted as being part of a total system (Buckley, 1967). Internet communities, conceptualized as CAS, present the following characteristics:

- **Multiple key components**. Internet communities are formed by Internet users that interact in a socio-cultural way. The socio-cultural interactions are based on interchanging digital content by means of email, instant messaging, forums, blogs, or social networking sites, that we call, in some cases, Internet memes.

- **Different structural levels**. As suggested by Spitzberg (2014), we can distinguish two structural levels on the analysis of the spread process of Internet memes: the meme and the societal level. At the meme level - the initial stage - the meme's popularity relies on having distinctive content, which corresponds with the concepts of quiddities, on a daily basis. At the societal level - the emergent stage - another set of parameters is added for success. It is important to note that at the societal level, an Internet meme needs to follow the mindset of its enveloping society to meet less resistance or total rejection.

- **Intrinsic diversity among the key components**. Internet communities are formed by a great diversity of users with experiences within a socio-economic, cultural, and environmental context.

- **Functional dynamics**. Internet communities are open systems that exchanges information with the complex environment that surrounds it. In order to survive, Internet communities need to adapt themselves to new environmental conditions, adjusting their functional units through modification and selections of Internet memes.

- **The impact of the social structure**. The socio-cultural interactions among Internet users are not random, but rather are limited by social networks, both internally and externally. The social structure and socio-cultural interactions among Internet users have a crucial effect on the process of Internet meme evolution.

3. An agent-based simulation model of spread of memes using agents with memory

In this section, we present an agent-based simulation model which explores the emergence of communities through the diffusion of cultural features in a network and illustrates the spread of memes using agents with memory. This simulation model was implemented using Netlogo^{TM} software which was developed in the late 1990s by Uri Wilensky. This simulation software is a general-purpose agent-based modelling language used worldwide that provides a graphical modelling environment (Dorner, 1997), freely available on the Netlogo^{TM} website (https://ccl.northwestern.edu/netlogo). As Wilensky (1999) explains: it is an extension of the Logo language in which user controls a graphical turtle by issuing commands, it includes a grid of patches, each patch is a cell computationally active. Turtles and patches are self-contained objects with internal local state. In Netlogo^{TM} models, time passes in discrete steps, called ticks. Additionally, Netlogo^{TM} allows to scale space and time, such as 1 square meter = 1 patch and one tick can represent a minute or a day, etc.

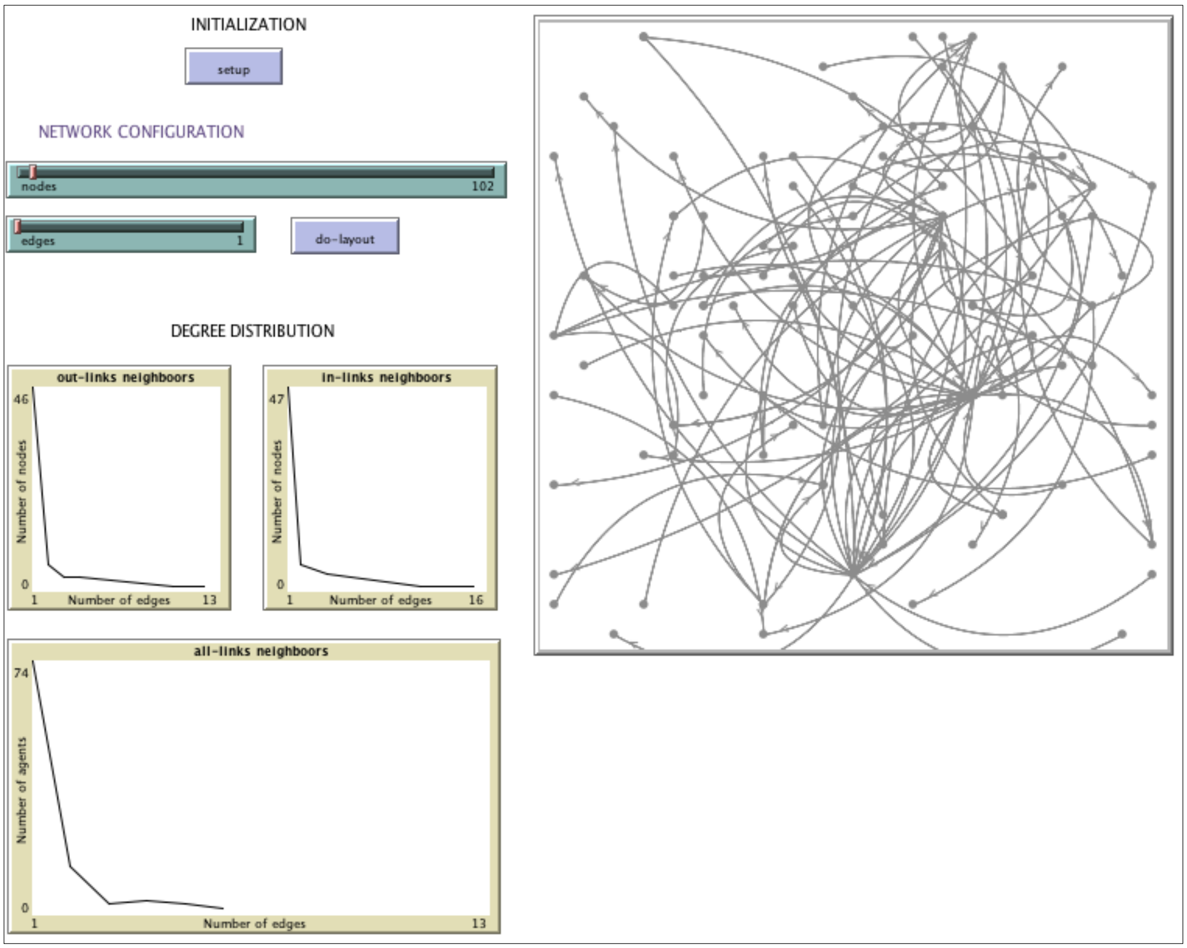
3.1 The network

The agents, modeled as nodes with random distribution of indegrees/outdegrees, represents users interacting on the Internet. The network links are oriented (indegrees/outdegrees) and represent the relationships among users. The general structure of the network is determined by three modifiable parameters:

1. Total population of agents.
2. Network links (edges).

In order to build a network that reflects the social organization of people online, the simulation model first creates the population of agents. Then, for each agent, the simulation model generates a random number of links, from 1 to edges slider value, among those agents and randomly picks other agents. The direction of all of these links is determined randomly.

Figure 1. Netlogo™ interface implemented to create a network of agents with a random number of links



2.2 The memes

In this model, memes are represented as a feature that spreads on Internet, starting from a user selected randomly. The creation of memes is influenced by two modifiable parameters:

1. *memeticks*: this determines the rate (in time-steps) at which new memes are introduced in the network.
2. *rate*: this determines the numbers of memes that are created at each new introduction into the network.

Memes are represented by a color. When an agent creates or adopts a new meme, it turns into that color. Also, the index for the number of categories of meme this agent knows increases by 1. This is represented in the model by a number at the center of the node. This number represents the number of colors (memes) that agent has access to and may use in its interaction with other agents. When trying to spread a meme through its outdegrees, the agent's node and outdegree links all turn to the color of this meme. The agents on the receiving end may or may not adopt the meme. If they do adopt it, they turn into that color and their index increases by one, thus repeating the process from the start.

Figure 2. The Netlogo™ interface implemented to create memes and to analyze their spread on the network

3.3 The agents with memory

Upon creation, each agent gets a set of different interests for the 13 different types of memes. To generate those preferences, the model assigns a random value from 1 to 10 (one for each meme type) to each agent. In other words, the interest of each agent for each type of meme is represented on a scale from 1 to 10. The interests of an agent for different types of memes has an influence on 4 processes:

1. Creating new memes: the type of memes that are created by an agent is done randomly.
2. Spreading memes: when an agent knows several memes, the selection process to decide which one is going to spread at each time-step is random.
3. Remembering meme: the rate at which an agent forgets each meme, depends on its own interests.

Table 1. Example of randomly generated interest values for one agent

Agent 1	Scale of interests									
Meme type	0	1	2	3	4	5	6	7	8	9
Red										
Orange										
Brown										
Yellow										
Green										
Lime										
Turquoise										
Cyan										
Sky										
Blue										
Violet										
Magenta										
Pink										

As mentioned before, the agents of this model possess a memory mechanism and can therefore forget memes. When an agent adopts a meme, the memory value for this meme is 15. However, at each time-step, agents forget about the memes they know at a rate of "11 minus the interest value for this meme type"(i.e., 11 – 1~10). If the memory value of a meme falls to 0, the agent forgets the meme altogether (the agent loses the color of the corresponding meme and its index decreases by 1). However, every time an agent is exposed again to a meme, the memory value for this meme returns to 15.

Figure 3. The simulation results of memes spread through a network of 152 nodes

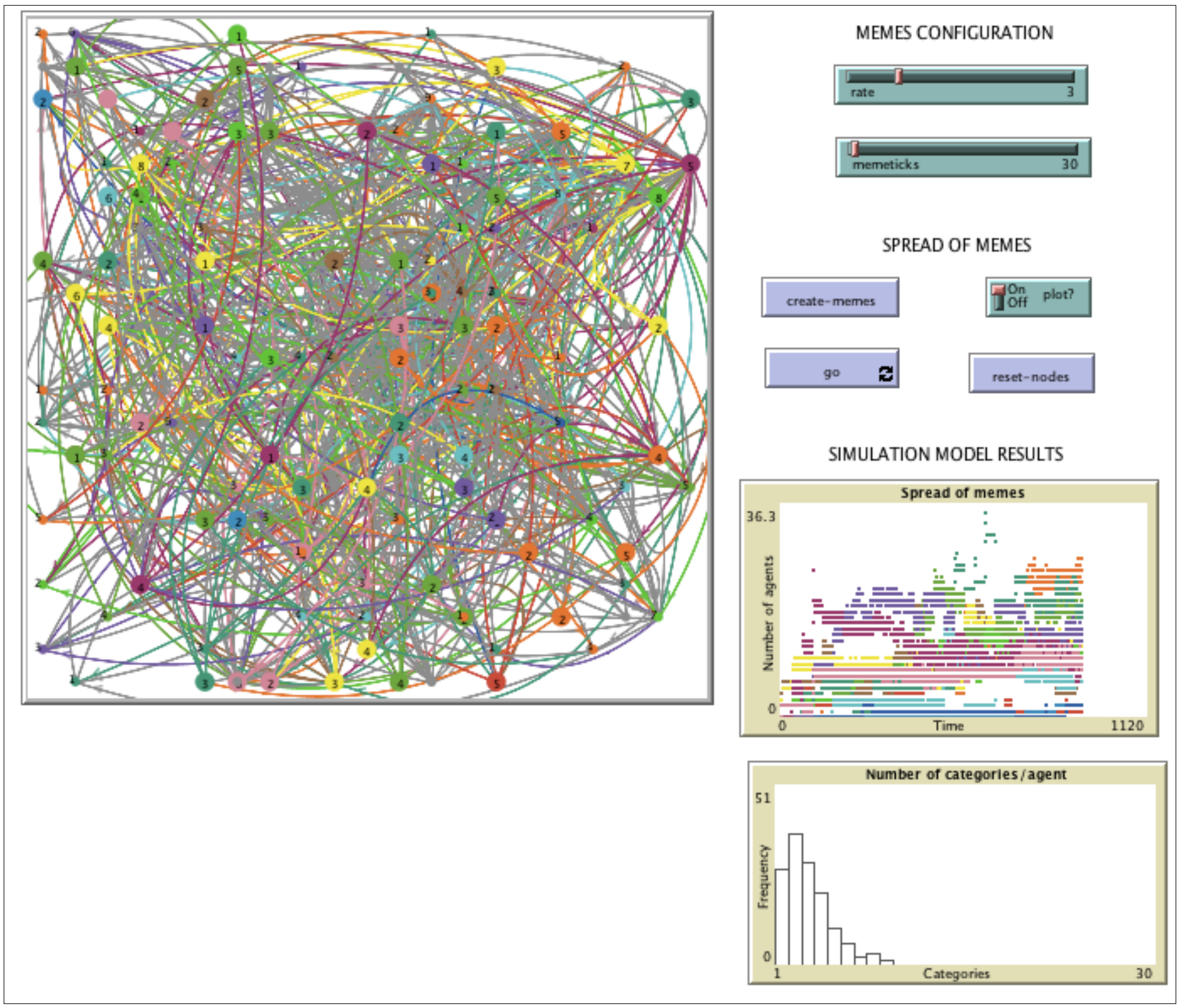
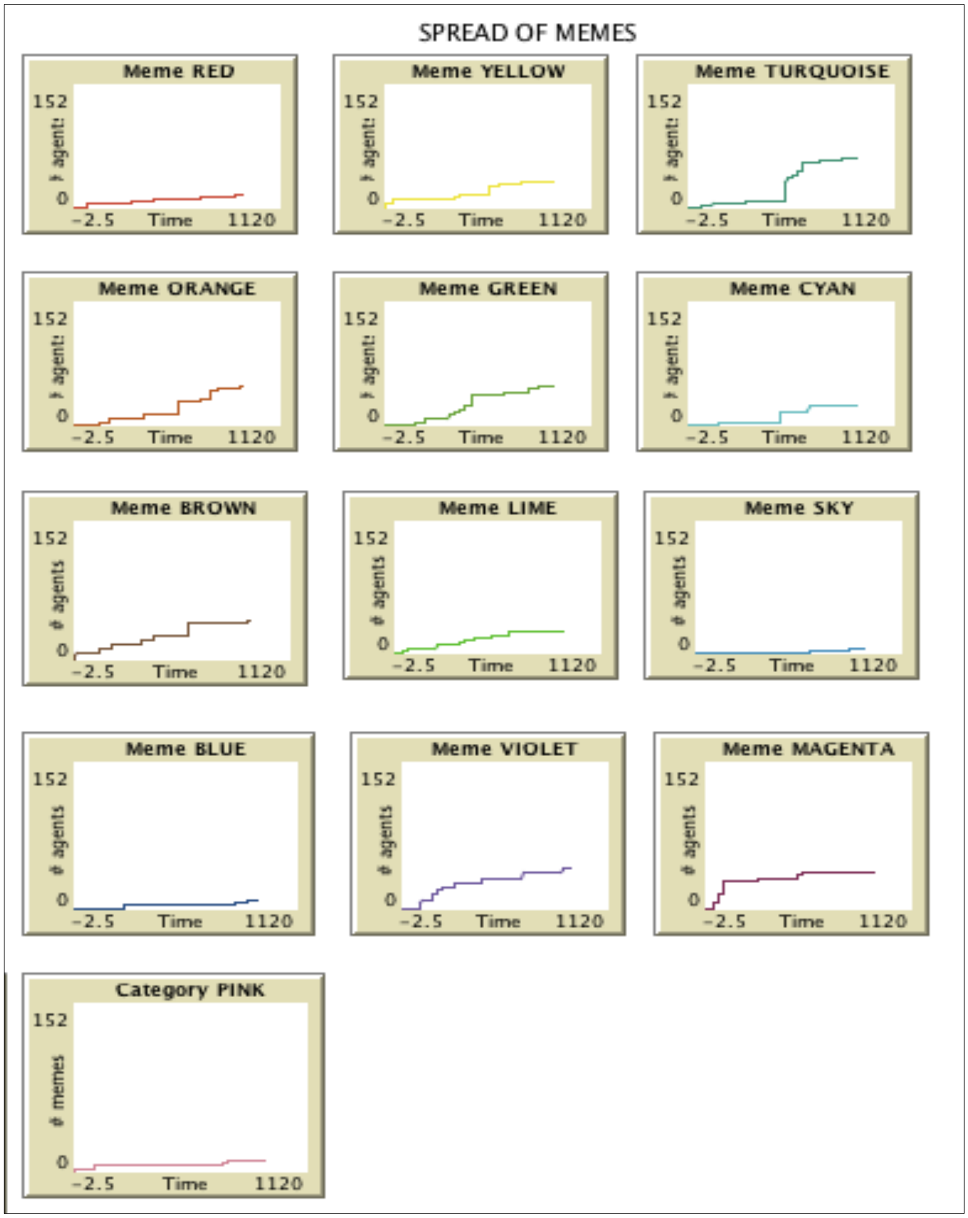


Figure 4. The simulation results of individual memes spread through the whole network of 152 nodes



3.4 Agent-based simulation model validation

As Wilensky and Rand (2015) explains: simulation model validation is the process of determining whether the implemented simulation model corresponds to some phenomenon in the real world. Our agent-based simulation model was validated using the dynamic technique called *sensitivity analysis*. Through it, values of simulation parameters are systematically changed over some range of interest and the simulation model´s behavior is observed (Banks, 1998). This technique allows identifying the simulation parameters to which the simulation model behavior is very sensitive. The simulation parameter considered to carry out the sensitivity analysis was the rate of memes which determines the numbers of memes that are created at each new introduction into the network. The minimum value was 2 and the maximum value was 10. We simulated the spread of memes considering two rate extreme values and a network built of 100 agents.

Figure 5. The simulation results of memes spread through a of 100 nodes and a memes rate equal to 2, after 1000 ticks

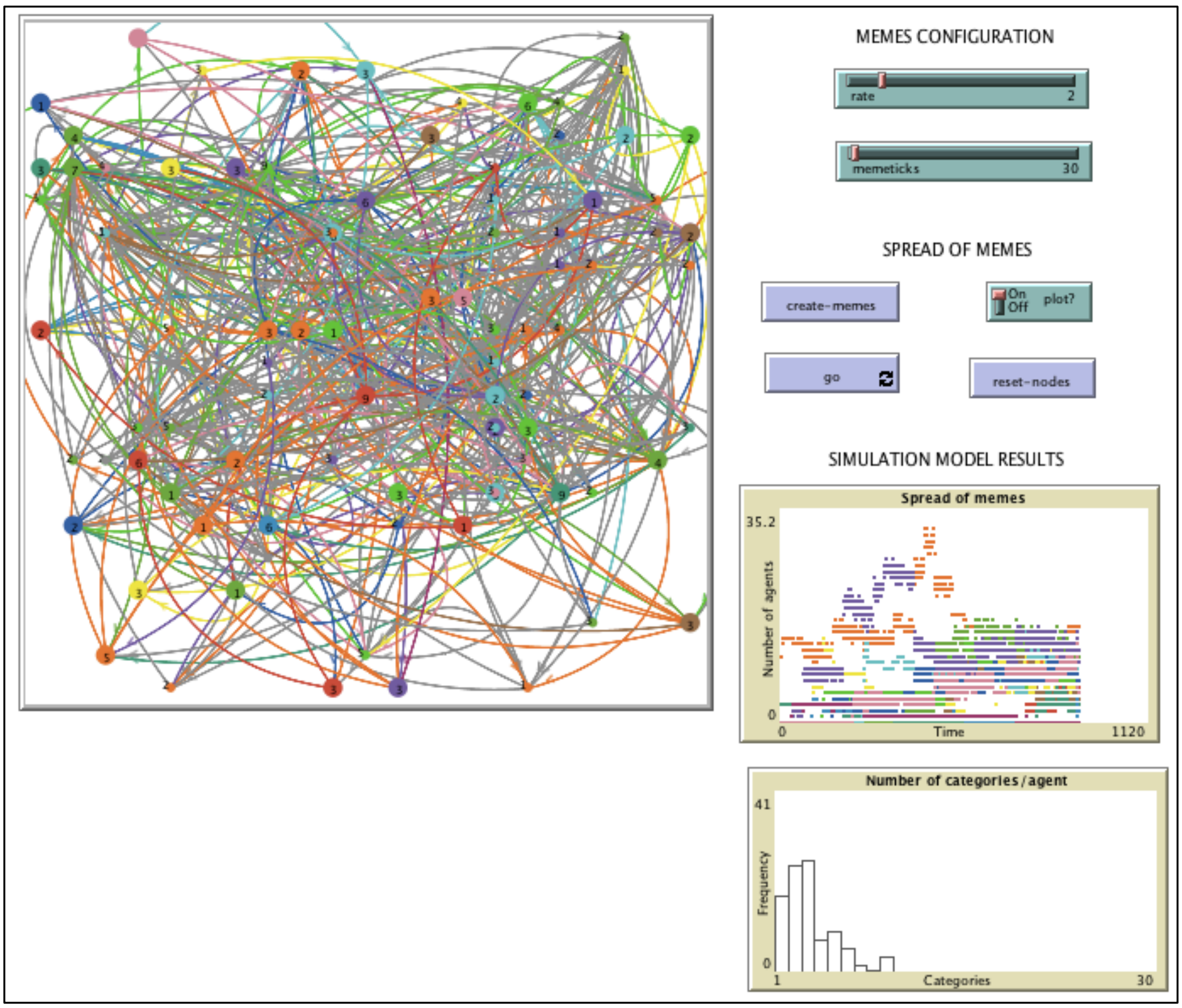


Figure 6. The simulation results of individual memes spread through a network of 100 nodes and a memes rate equal to 2, after 1000 ticks

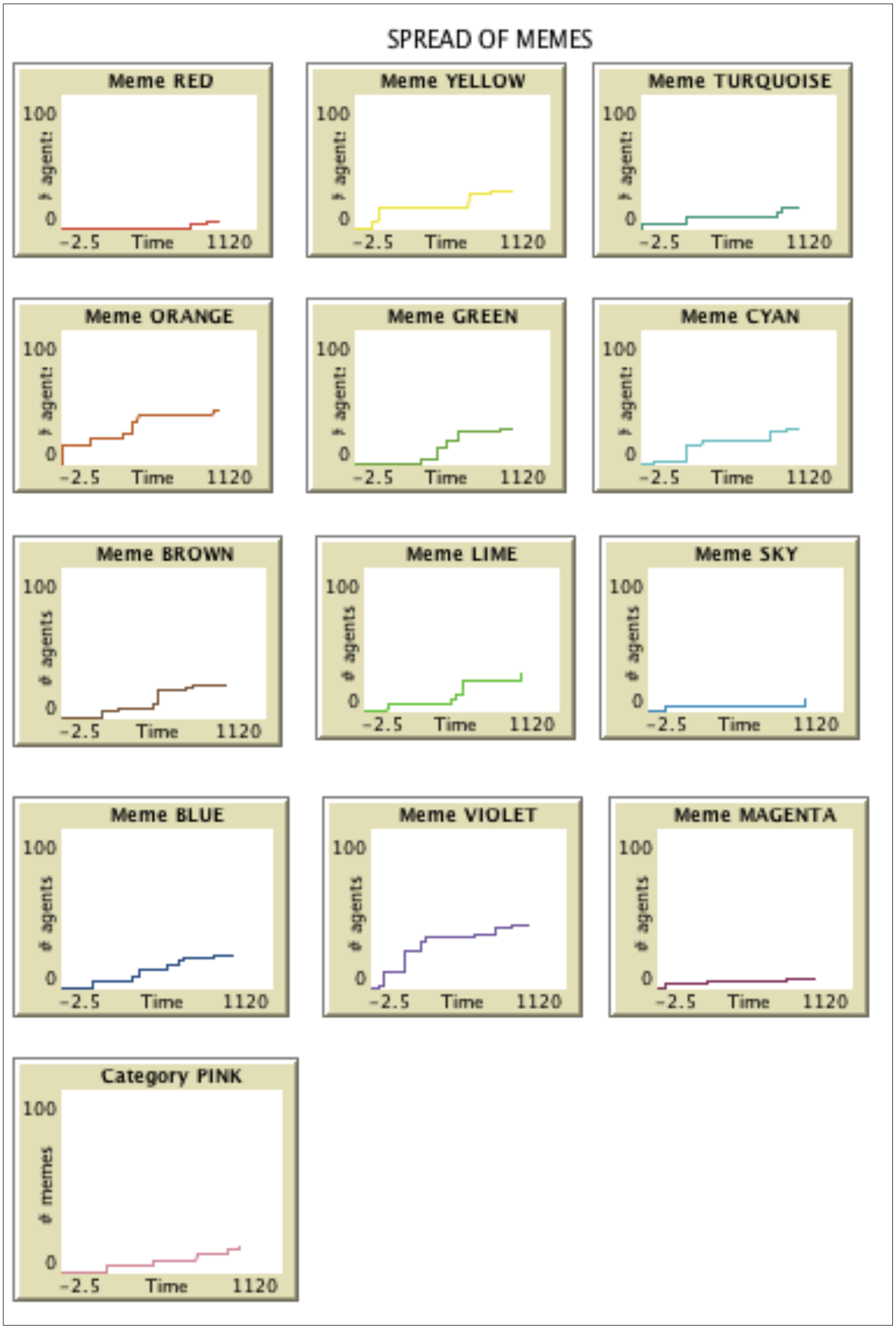


Figure 7. The simulation results of memes spread through a network of 100 nodes and a memes rate equal to 10, after 1000 ticks

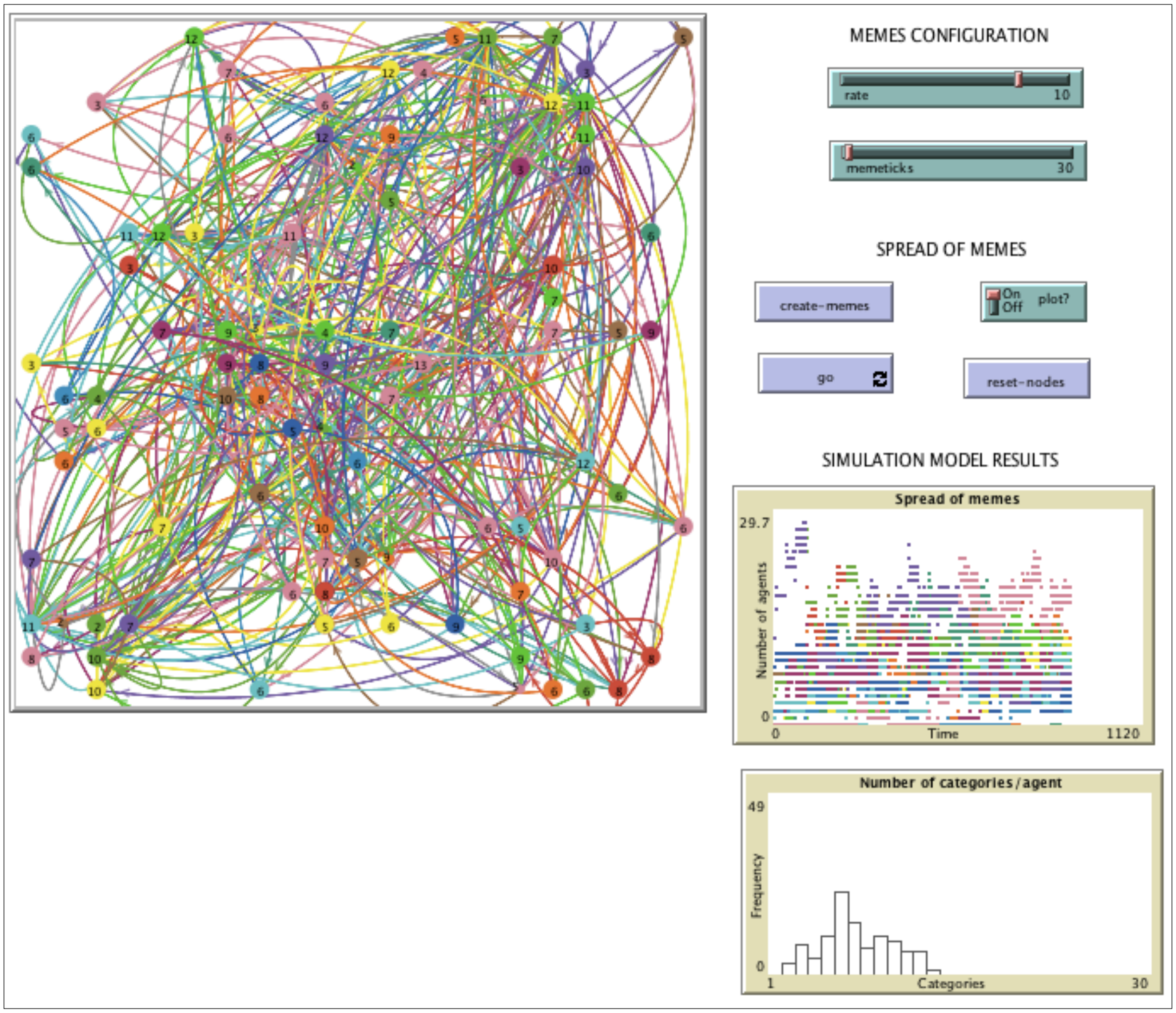
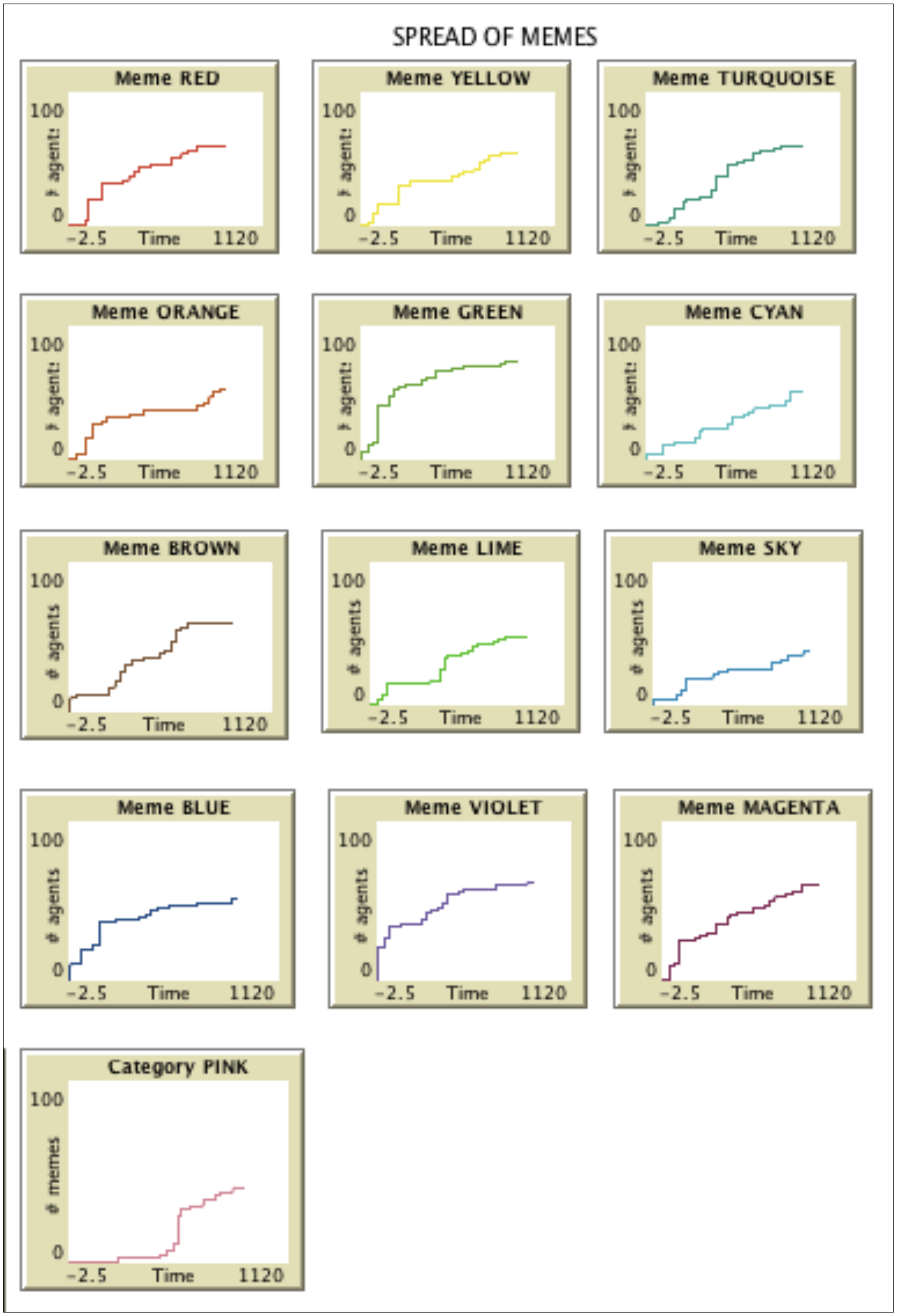


Figure 8. The simulation results of individual memes spread through a network of 100 nodes and a memes rate equal to 10, after 1000 ticks



3.5 Agent-based simulation model verification

As Wilensky and Rand (2015) explains: simulation model verification is the process of making sure that the simulation model has been correctly implemented on computer using simulation software. In Netlogo™ software the compilation and execution process that prepare the model to run happen behind the scenes and require no intervention by the user however, syntax and other coding errors that cause runtime errors during compilation and execution process interrupt the simulation (Stigberg, 2012). In this phase, we eliminated the "bugs" from the code so the model was correctly implemented, free of errors.

4. The impact of the network structure on the memes spread rate and diversity

As previously mentioned, the social network structure plays a crucial role for understanding the diversity of memes. In this section, three simulation scenarios are designed considering three configurations of the network structure in order to analyze the impact of the network structure on the spread rate, specifically the competition among memes to determine their diversity.

4.1 Definition of simulation scenarios based on the network structure

According to Peterson *et al.* (2003), in order to build simulation scenarios, it is important to specify what is known and unknown about the system´s dynamics then, the alternative ways that the system could evolve must be identified. After that, a set of simulation scenarios is built, through which our current thinking about the system should be expanded. The simulation model considers that network structure depends directly on the network oriented links (indegrees/outdegrees) that represent the relationships among agents, and their random distribution. In this case, we suggest three different values for the number of indegrees/outdegrees: 5, 25 y 50. We consider 100 agents and 5 memes created each 30 ticks.

4.2 Analysis of simulation scenarios

The impact of the network structure on the spread rate when the number of indegrees/outdegrees is 5, is shown in Fig. 9. The frequency distribution of different memes categories per agent is presented graphically in Fig. 10.

Figure 9. The spread of memes when the indegrees/outdegrees is 5, 1000 steps

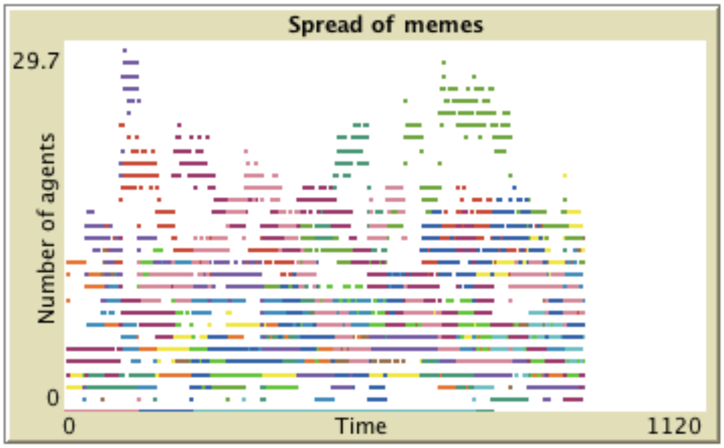
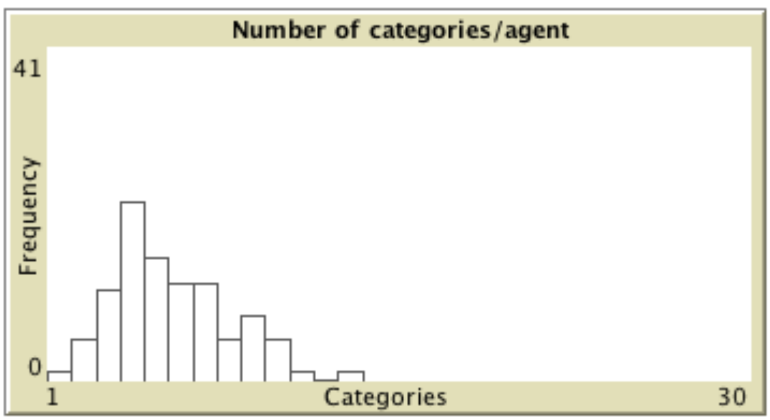


Figure 10. The frequency distribution of different memes categories per agent when the indegrees/outdegrees is 5, 1000 steps



The impact of the network structure on the spreading rate when the number of indegrees/outdegrees is 25, is shown in Fig. 11. The results may also be represented graphically as is shown in Fig. 12.

Figure 11. The spread of memes when the indegrees/outdegrees is 25, after 1000 ticks

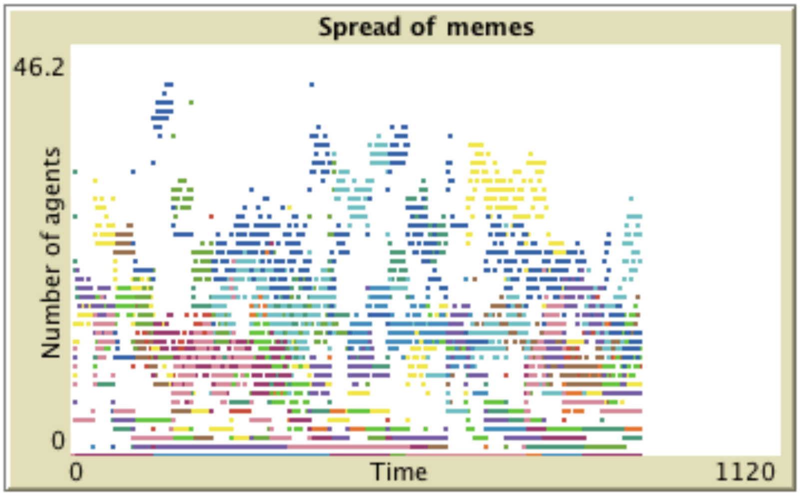
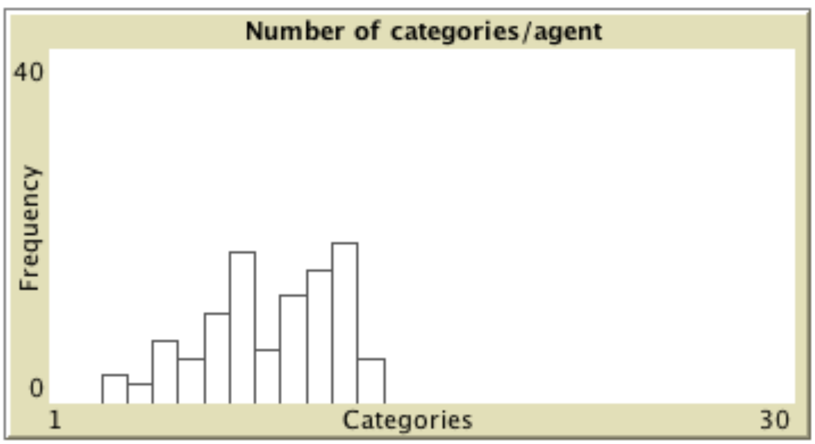


Figure 12. The frequency distribution of different memes categories per agent when the indegrees/outdegrees is 25, after 1000 ticks



The impact of the network structure on the spreading rate when the number of indegrees/outdegrees is 50, is shown in Fig. 13. The results may also be represented graphically as is shown in Fig. 14.

Figure 13. The spread of memes when the indegrees/outdegrees is 50, after 1000 ticks

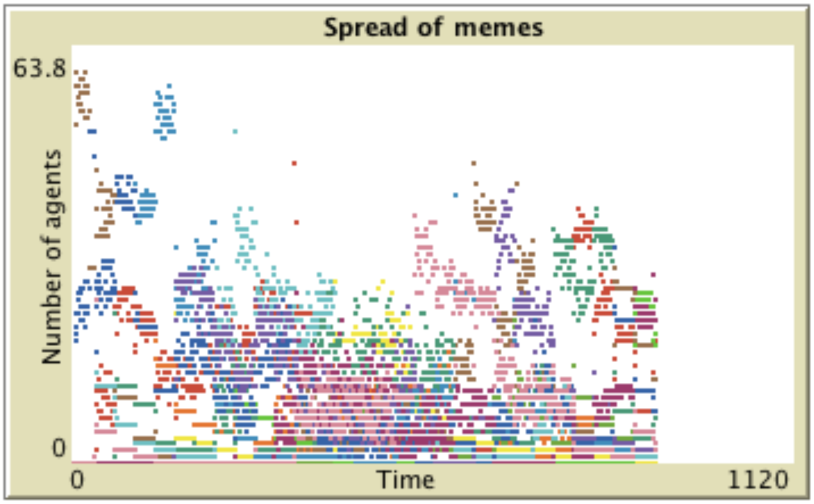
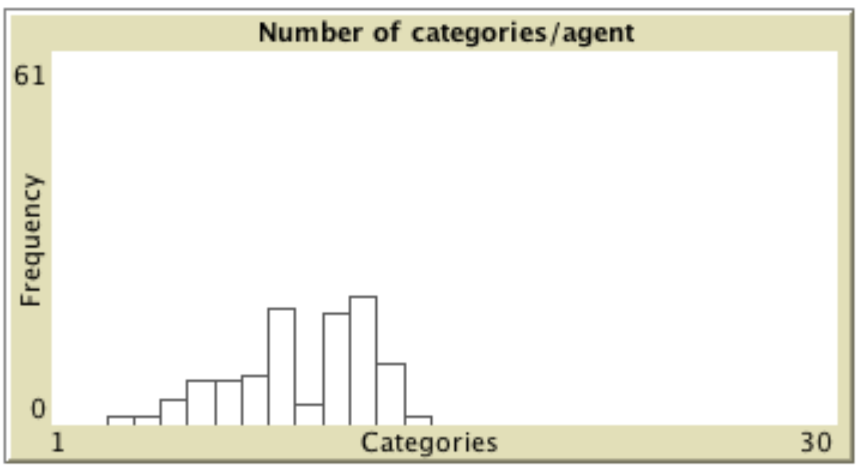


Figure 14. The frequency distribution of different memes categories per agent when the indegrees/outdegrees is 50, after 1000 ticks



5. Concluding remarks

This study contributes towards a better understanding of the spread of memes considered as cultural features, represented by thoughts and ideas, appearing regularly on any random user and trying to spread across the Internet. Users are modeled as agents exchanging information. From the interaction among agents a specific social network structure is built and it is crucial for understanding the rate spread and the diversity of memes. An agent-based simulation model was built using Netlogo^{TM} software. We considered as a reference the simulation model proposed by Amaddio (2015). However, in the simulation model proposed in this study we considered multiples memes created randomly by agents in different steps of time; agents possess a memory and their own interests are influenced for instance by the socio-cultural, political, environmental, etc. conditions; the network edges are directional.

We consider that the agent-based simulation model proposed in this study has many potential applications. The spread of memes through a community is similar to the spread of a technology innovation. These can often be spread by word-of-mouth and certainly the uptake of a consumer innovation is reflected in this application. Of course, ‘industrial’ - commercial adoption - type innovations are not adopted in this manner.

In tracking terrorist activity, the application can be used to track messages of motivation from terrorist groups. Messages aimed at highly susceptible candidates could be identified by the strength of the connections, enabling potentially lifesaving intervention to take place. Similarly, messages aimed at recruitment into a terrorist cell could be identified, especially if coming from a lead agent. Counter messaging could take place directly or locally to reduce the threat and minimize recruitment. The system could also potentially track general movement of potential threats from country to country by tracing the messages and establishing the receiving locations.

Also, the simulation model implemented could be used to study the spread of religion. The application could model the spread of religions over time, indicating how such communities have developed. There are potentially two levels to consider. One would be the spread of religious messages through many years and generations, which could identify geographical ‘hot-spots’. The demise of such communities could also be established if seeding of religious messages could be achieved and then traced, enabling the contraction of the network to be established. A second level is to model the more traditional approach whereby children are automatically ‘assigned’ the religion of their parents. This is technically still within a network, albeit a very small one.

A more commercial use entails the tracing of marketing messages. These can be general, brand messages or more specific ‘special offers’ aimed at a particular group of customers. Once the behavioral pattern is understood and can be predicted, a more prescriptive approach could be utilized to target specific customer groups who may be identified as being currently ‘susceptible’ to a certain marketing initiative, such as a special offer on a product they are known to be interested in.

By knowing the spread of information through an online community, political parties can target their messages to certain portions of the community. If the recipients are ‘susceptible’, they will either adopt the message and convert to the sending party or existing converts will feel re-enforced in their political direction.

Considering sex trafficking, the simulation model can be used to see how the notorious underground and often concealed connections between people involved in this 'industry' can be more easily tracked and infiltrated.

Finally, the next steps consist on the addition of a giant agent called Media whose function will be equivalent to websites. We also are interested on the creation of memes by an agent according to their own interests instead of randomly.

Acknowledgements
The author appreciates the partial support by National Council for Science and Technology of Mexico (CONACyT) (SNI number 65477).

References

Amaddio, K. (2015). The Spread of a Meme Across a Social Network. Retrieved from: http://modelingcommons.org/browse/one_model/4424#model_tabs_browse_info

Banks J. Handbook of simulation: Principles, methodology, advances, applications, and practice. Georgia: John Wiley & Sons, Inc; 1998.

Bauckhage, C. (2011). Insights into Internet Memes. Proceedings of the Fifth International AAAI conference on Weblogs and Social Media.

Bennett, L., & Segerberg, A. (2012). The logic of connective action. *Information, Communication and Society,* 15(5), 739–768

Blackmore, S. (2001). Evolution and memes: The human brain as a selective imitation device. Cybernetics & Systems, 32(1-2): 225–255.

Buckley W. Sociology and Modern Systems Theory. Englewood Cliffs, N.J.: Prentice-Hall; 1967.

Dawkins, R. (1976). The Selfish Gene, Oxford University Press.

Dennett, D. C. (2017). From Bacteria to Bach and Back. London: W. W. Norton & Company.

Dorner D. The logic of failure: Recognizing and avoiding error in complex situations. New

York: Basic; 1997.

Gomez Rodríguez, M., Leskovec, J., and Krause, A. (2010). Inferring networks of diffusion and influence. In Proceedings of ACM SIGKDD, 1019-1028.

He, S., Zheng X., and Zeng, D. (2015). A model free scheme for meme ranking in social media. Decision Support Systems, 81: 1-11.

Kuhn, T., Perc, M., Helbing, D. (2014). Inheritance patterns in citation networks reveal scientific memes. ArXiv:1404.3757V1. 14 Apr2014.

Leskovec, J., Beckstrom, L. and Kleinberg, J. (2009). Memetracking and the dynamics of the news cycle, Technical report, Cornell University and Stanford University.

Mazambani, G., Carlson, M., Reysen, S., Hempelmann, C. (2015). Impact of status and meme content on the spread of memes in virtual communities.

Peterson G D, Cumming G S, Carpenter S R. (2003). Scenario planning: a tool for conservation in an uncertain world. Conservation Biology. **17** (2): 358-366.

Segev, E., Nissenbaum, A., Stolero, N., Shifman, L. (2015). Families and Network of Internet Memes: The relationship between cohesiveness, uniqueness, and quiddity concreteness. Journal of Computer-Mediated Communication, 20: 417-433.

Simmons, M. P., Adamic, L. A., and Adar, E. Memes online: Extracted, subtracted, injected, and recollected. In Proceedings of ICWSM, 353-360.

Spitzberg, B.H. (2014). Toward a model of meme diffusion. *Communication Theory,* 24(3), 311–339.

Stigberg D. An introduction to the Netlogo modeling environment. In: Westervelt J D, Cohen G L, editors. Ecologist-developed spatially explicit dynamic landscape models, Modeling dynamic systems. New York: Springer; 2012.

Wang L. and Wood B.C. (2010). An epidemiological approach to model the viral propagation of memes. Applied Mathematical Modelling. 35: 5442- 5447.

Weng, L., Flammini, A., Vespignani, A., and Menczer, F. (2012). Competition among memes in a world with limited attention. Scientific Reports (2).

Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Wilensky U. GasLab: An extensible modeling toolkit for exploring micro-and-macro-views of gases. In: Roberts N, Feurzeig W, Hunter B, editors. Computer Modeling and Simulation in Science Education. Berlin: Springer Verlag; 1999. p. 151–178.

Wilensky U, Rand W. An introduction to Agent-Based Modeling. Cambridge: The MIT Press; 2015.