Model-checking Information Diffusion in Social Networks with PRISM

Louise A. Dennis[1] and Marjia Slavkovik[2]

[1] University of Liverpool, UK
L.A.Dennis@liverpool.ac.uk
[2] University of Bergen, Norway marija.slavkovik@uib.no

Abstract. In this paper we present an agent-based approach to formalising information diffusion using Markov models which attempts to account for the internal informational state of the agent and investigate the use of probabilistic model-checking for analysing these models. We model information diffusion as both continuous and discrete time Markov chains, using the latter to provide an agent-centred perspective. We present a negative result - we conclude that current model-checking technology is inadequate for analysing such systems in an interesting way.

Keywords: verification of agent behaviour in a social network, information diffusion

1 Introduction

Interest in social networks research is increasing, following the rise of *social network services* in the last decade. Various aspects of networks have been studied, see *e.g.,*[13].

Social network analysis is concerned with the structures of social relations and the graph they form, as well as how that structure influences, and is influenced by, the spread of information in the networks *e.g.,* [12,27]. Recently we see also see work that brings together social network research and logic. Informational and motivational states of the agents in the network are modelled, not just the relations between agents, *e.g.,* [11,4,21,28,25]. Network phenomena are also being given formal models and specifications *e.g.,*[21,4,20].

Diffusion is the process of spreading information through a network of agents. A social network is given as a graph of agents (vertices) and there exists an edge between two agents if they communicate/share information with each other. Depending upon a number of factors, an agent that has received information might be socially influenced to adopt it as true (believe it) and share it further in the network. There are several social influence models proposed in the literature, each describing different conditions under which the agent spreads the received information. An example of a question typically studied in diffusion is: will a point be reached where the information is adopted by all agents? Following early work studying social networks as part of epidemiology, this is referred to as full contagion. We are particularly interested in how the internal informational state of the agent affects its decisions to spread information, and so have developed an agent-based model for information diffusion which explicitly uses this information. We adopted a Markov chain formalisation for this model since we were secondarily interested in using formal verification to analyse the network.

Formal verification involves proving or disproving that a system is compliant with a formally specified property [10]. Arguably the most practical method of formal verification is model-checking [7], in which all possible executions of a system can be examined automatically based on a model of the system.

Diffusion has been extensively studied in the social network analysis literature, see for example [13,19] for an overview, in particular the impact of the social network graph on the diffusion process has been studied. Social networks of communication have physically changed.In particular aspects of these networks, such as the distance between two nodes, and the speed of communication, have changed drastically. This observation has revived interest in the study of information diffusion, including work that represents the phenomena using Markov chains (*e.g.,* [3]) as we do here.

In our work we have built formal specifications of social networks and diffusion properties using the input language of a probabilistic model-checker (PRISM). Unfortunately even simple models that take account of both network structure and an agent's informational state proved largely intractable for model-checking on networks of any significant size. This essentially gives us a negative result for the use of probabilistic model-checking for information diffusion on social networks. However the formal Markov chain framework for studying the effect of an agent's informational state on information diffusion should also be amenable to study using simulation based techniques which we leave for further work.

Contribution. We provide a framework for formalising information diffusion in a way that takes account of an agent's internal informational state as Markov chains which focuses on the broadcasting of opinions as a key feature of study.

Model-checkers can be used as simulation systems, but their value is in their ability to exhaustively explore all possible system states and produce highly accurate results, exploring best and worst outcomes. Their weakness is that such exploration is computationally expensive and necessarily limits the size of the systems that can be examined. We determine experimentally that the PRISM model-checker – arguably the state of the art in terms of probabilistic model-checkers can not be used to analyse networks of any significant size – a negative result and a challenge to the developers of such tools.

2 Information diffusion in social networks

Several models of information diffusion through influence have been proposed, although the task of finding a good model remains challenging [6]. The social influence models used to define processes of diffusion can broadly be classified into two classes: infection models and threshold models, with the possible exception of the recent Simmelian model [25]. In an infection model, each node is assigned a probability of being influenced [20]. In threshold models, an agent is influenced when the number of her influenced neighbours passes a certain threshold [29].

The SIS model One of the classic infection models is the SIS model [1]. In this infection model each of the nodes in the graph can be in one of two states: infected or susceptible to infection. At time t, $s(t)$ represents the susceptible proportion of the total population N, $i(t)$ represents the infected proportion, and λ represents the daily contact rate, which means the proportion of the susceptible users infected by infected users in the total population, where $s(t) + i(t) = 1$. There will be $\lambda s(t)$ susceptible users infected at time t. At time $t = 0$ the proportion of infected nodes is i_0.

The SIS model assumes that μ represents the daily rates of the "cured" nodes (a mode can now become uninfected). The SIS model can be described by

$$\begin{aligned} &\tfrac{di}{dt} = \lambda i(1-i) - \mu i \\ &i(0) = i_0. \end{aligned}$$

Threshold influence models. Threshold influence models define the choice of whether a node will become infected or not as a function of the degree (or set of neighbours) of the node in question. Given an agent (node) x, let $n(x)$ be the set of agents that are directly connected to x in the social graph. Threshold models define a threshold q. The agent x will become infected if $|n(x)| \geq q$.

Other influence models. In the Simmelian model [25] of influence x gets infected if x is in a clique in which all other nodes are infected.

We will use the SIS model and threshold influence model as our starting point for introducing an agent's information state into models of information diffusion.

3 PRISM Background and Theory

PRISM [17] is a probabilistic symbolic model-checker in continuous development since 1999, primarily at the Universities of Birmingham and Oxford. Typically a model of a system is supplied to PRISM in the form of a probabilistic automata. This can then be exhaustively checked against a property written in PRISM's own probabilistic property specification language, which subsumes several well-known probabilistic logics including PCTL, probabilistic LTL, CTL, and PCTL*. PRISM has been used to formally verify a variety of systems in which reliability and uncertainty play a role, including communication protocols and biological systems [9,18].

In our models we use discrete-time and continuous-time Markov Chains as our probabilistic automata.

Definition 1. *[16] (Discrete-time Markov chain (DTMC)). A discrete-time Markov chain (DTMC) is a tuple $D = (S, s_i, P, L)$, where is S a finite set of* *states, $s_i \in S$ is a distinguished initial state, $P : S \times S \to [0,1]$ is a transition probability matrix such that $\sum s' \in S.\, P(s,s') = 1$ for all $s \in S$, and $L(s) \subseteq AP$ is labelling with atomic propositions.*

A discrete-time Markov Chain describes a set of execution paths through the state space S where P gives the probability of one state moving to the next and L describes propositions that are true in any given state. PRISM explores the state space and can calculate the probability that various logical properties are always true, sometimes true, or true at some time t and so on in the model.

Definition 2. *[16] (Continuous-time Markov chain (CTMC)). A continuous-time Markov chain (CTMC) is $C = (S, s_i, R, L)$ where:*

- *S, s_i and L are defined as for DTMCs*
- *$R : (S \times S) \to \mathbb{R}_{\geq 0}$ is the rates matrix.*

Intuitively a CTMC describes a set of states and the rate at which one state moves to another. It is possible that for any state s there are several states s' such that $R(s,s') > 0$ and whichever state it transitions to first will determine the resulting behaviour of the system. Given a set of rates, $R(s,s')$ for some state s it is possible to infer the probability with which it will transition to each s' for any given time step t. PRISM can then explore this state space.

As well as calculating probabilities, PRISM is able to calculate the *expected reward* in some system. We can specify a rewards function, $\rho : S \to \mathbb{R}$ which assigns some reward value to the states in S. Among other things, this allows us to investigate the expected reward at some time step, t, in the system. This proves a particularly powerful tool in the study of information diffusion since we can model the number of agents adopting some idea as a reward.

4 Model-Checking Infection Models

4.1 Classic SIS Model

We choose as a first example the classic *SIS* infection model. This model takes neither the network structure nor the informational state of the agents into account, beyond the infection and recovery rates. We model this as a continuous time Markov Chain (CTMC). Our main aim in presenting this model is to illustrate the kinds of questions that can be asked and answered using a probabilistic model-checker.

Model 1 *We consider a network with m_a agents. This network contains $m_a + 1$ states, $s_i, 0 \leq i \leq m_a$. There are $m_a + 1$ propositional variables $p_i, 0 \leq i \leq m_a$ where p_i means that i agents in the model are infected. The labelling function is $L(s_i) = p_i$ (*i.e.*, i agents are infected in state s_i). The initial state is s_1 (1 agent is infected at the start).*

λ and μ are the infection and recovery rates from the SIS infection model and these give us the following rate matrix:

$$R(s_i, s_{i+1}) = \lambda(m_a - i) \qquad \textit{if } 0 < i < m_a \qquad (1)$$

$$R(s_i, s_{i-1}) = \mu(i) \qquad \textit{if } 0 < i \qquad (2)$$

We use PRISM to explore the behaviour of this model for different values of λ, μ, m_a and so on. For instance, Figure 1 shows the probability of full contagion for all values of μ given $\lambda = 0.5$ and $m_a = 20$. We can see that where $\mu < \lambda$ there is a high probability that all agents will adopt some information while as soon as $\mu \geq \lambda$ this probability drops.

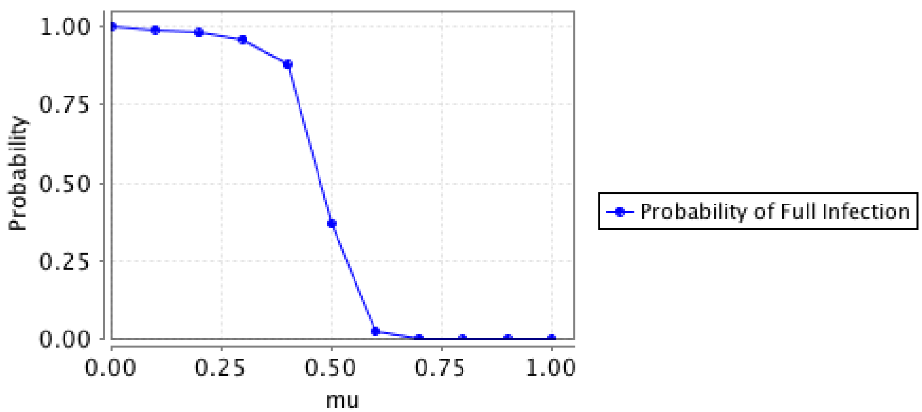

Fig. 1. Probability that all agents are infected in model 1 where $\lambda = 0.5$ and $m_a = 20$

This result doesn't hold for all network sizes. As the network grows the probability of full infection increases. Figure 2 shows that even when $\mu = 1$ the probability of full infection occurring at some point is over 0.9 once there are more than 30 agents.

Figure 3 shows the probability that all agents will be infected before time t (for $t < 100,000$) given various values of μ when $m_a = 20$. From this we can see that when $\mu = 0$ (i.e., when there is no possibility of recovery) or $\mu = 0.2$ the model rapidly reaches a point where the probability that all agents are infected is close to 1. However with $\mu = 0.4$ it takes in the region of 65,000 time steps for the probability of full infection to converge (to a value of 0.92). For higher values of μ the probability of full infection remains very low.

However as the network size increases (to numbers where we know the overall probability of full infection at some point is high for all μ) then this difference disappears. It becomes a more interesting question to ask how many people do we expect to be infected at any point in time. Figure 4 shows the expected number of infected agents at time T for various values of μ in a network of 200 agents. As can be seen this value stabilizes quite rapidly and then remains steady, but

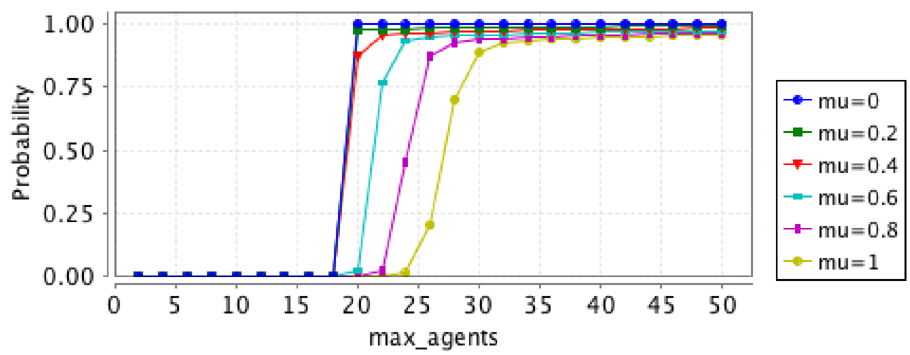

Fig. 2. Probability that all agents are infected in model 1 for $\lambda = 0.5$

as the value of μ increases the expected number of infected agents decreases (as the rate at which agents are recovering from infection has increased).

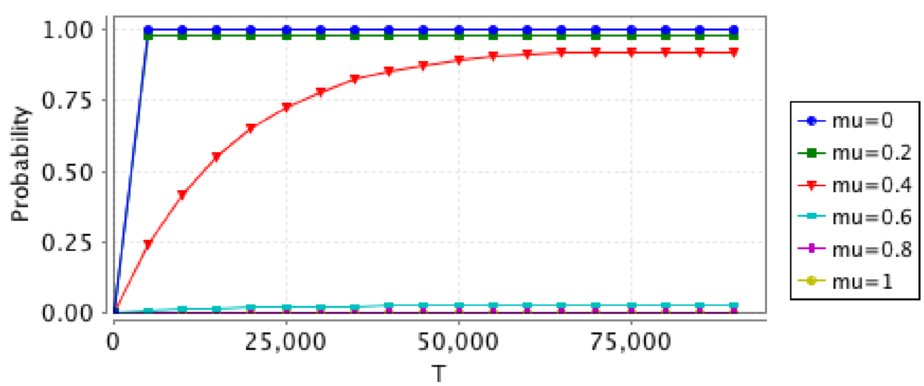

Fig. 3. Probability that all agents are infected by time T in model 1 for $\lambda = 0.5$ and $m_a = 20$

4.2 Taking the Agent View: Informational State and Opinion Broadcast

As has been noted in the literature [4,5] the transmission of information around a social networks may depend both on the features of the specific agents in the network and on the structure of the social network itself. We are interested in the possibility of using model-checking to explore traditional social analysis aspects of how network structure affects the spread of ideas. Further we want to see how

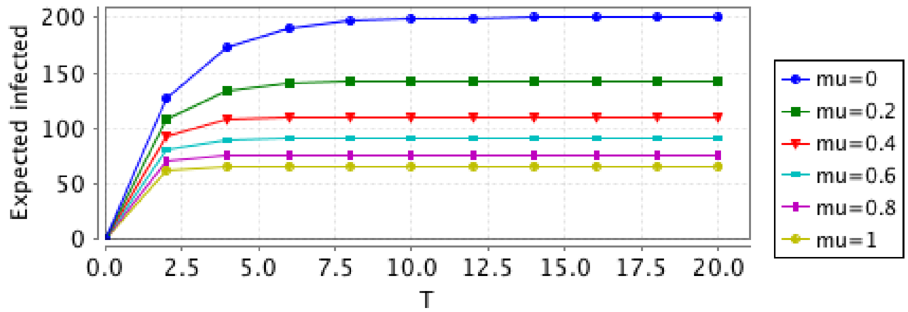

Fig. 4. The expected number of infected agents at time T in model 1 for $\lambda = 0.5$ and $m_a = 200$

agent properties contribute to the global effect (and ultimately in how actions by a mediating platform in a network service may contribute).

As an example of an agent feature that might influence contagion we consider how one idea may be associated with an "anti-idea" which might either cause an idea to be abandoned (analogous to recovering from infection in traditional model) or might cause other behaviour (e.g., greater adherence to the original idea, modifications to network structure and so on). Taking this example, which to the best of our knowledge has not been considered in social network analysis, is motivated by the insight from psychology that "once formed impressions are remarkably perseverant" [22]. In this case we use the current informational state of the agent to inform both how likely it is to adopt an opinion. Once adopted it will broadcast the opinion to its network.

It is natural, in such a case, to consider the transition system of our model in terms of the transitions of the individual agents. PRISM provides support for constructing a DTMC from a specification of a transition relation on individual modules within a system were the state s of the system is the product of the states, s_{m_i}, of each module, $s = (s_{m_1}, \ldots, s_{m_n})$. This support uses a labelling on transitions within modules which synchronizes across all modules. Each module specifies the probability that the module will transition to some new state when a particular labelled transition, say l, occurs within the system. From this PRISM can calculate the probability distribution for the next state of the whole system given transition l. PRISM then assigns an equal probability that any transition that can occur will occur to derive the transition probability matrix over all possible transitions[3].

In order to take an agent view of a social network, we will model each agent as a PRISM module and use the notation $s^{a_i} \xrightarrow{l} p_1 : s_1^{a_i} \wedge \ldots \wedge p_n : s_n^{a_i}$ to indicate that agent, a_i in state, s^{a_i} transitions to state $s_j^{a_i}$ with probability p_j where $\sum_{i=1}^{n} p_i = 1$ when the transition labelled l occurs.

[3] This is detailed in `http://www.prismmodelchecker.org/doc/semantics.pdf`.

Model 2 *We will use a DTMC to model a network of agents. Each agent, a_i, in the network is a PRISM module and can be in one of three states. Either the agent agrees with some idea ϕ (written as state $s_{\phi}^{a_i}$) or it disagrees with the idea ($s_{\neg\phi}^{a_i}$) or it is indifferent to ϕ (written as $s_{\perp}^{a_i}$)). If there are n agents in the network, there are 3^n states in S.*

An agent will broadcast a message in favour of ϕ (respectively $\neg\phi$) to all of its connections if it agrees with ϕ. We treat this as a transition labelled $a_i\_says_{\phi}$. On receiving a message in favour of ϕ (respectively $\neg\phi$) there is a probability of λ that the agent will adopt the idea ϕ (abandoning the idea $\neg\phi$ if necessary).

Figure 5 shows the transition system for agent a_i where $cn(i,j)$ means i is

$$s_{\phi}^{a_i} \xrightarrow{a_i\_says_{\phi}} \quad 1 : s_{\phi}^{a_i} \tag{3}$$
$$s_{\phi}^{a_i} \xrightarrow{a_j\_says_{\phi}} \quad 1 : s_{\phi}^{a_i} \quad \text{if } cn(i,j) \tag{4}$$
$$s_{\phi}^{a_i} \xrightarrow{a_j\_says_{\neg\phi}} \quad \lambda : s_{\neg\phi}^{a_i} \wedge (1-\lambda) : s_{\phi}^{a_i} \quad \text{if } cn(i,j) \tag{5}$$
$$s_{\neg\phi}^{a_i} \xrightarrow{a_i\_says_{\neg\phi}} \quad 1 : s_{\neg\phi}^{a_i} \tag{6}$$
$$s_{\neg\phi}^{a_i} \xrightarrow{a_j\_says_{\neg\phi}} \quad 1 : s_{\neg\phi}^{a_i} \quad \text{if } cn(i,j) \tag{7}$$
$$s_{\neg\phi}^{a_i} \xrightarrow{a_j\_says_{\phi}} \quad \lambda : s_{\phi}^{a_i} \wedge (1-\lambda) : s_{\neg\phi}^{a_i} \quad \text{if } cn(i,j) \tag{8}$$
$$s_{\perp}^{a_i} \xrightarrow{a_j\_says_{\phi}} \quad \lambda : s_{\phi}^{a_i} \wedge (1-\lambda) : s_{\perp}^{a_i} \quad \text{if } cn(i,j) \tag{9}$$
$$s_{\perp}^{a_i} \xrightarrow{a_j\_says_{\neg\phi}} \quad \lambda : s_{\neg\phi}^{a_i} \wedge (1-\lambda) : s_{\perp}^{a_i} \quad \text{if } cn(i,j) \tag{10}$$

Fig. 5. Transition System for agent a_i in model 2

*connected to j in the network. Where a transition isn't specified (*i.e.*, for all the agents i is not connected to) then PRISM assumes a_i's state is unchanged by that transition (after all a_i is unaware of what a_j is saying). Where a transition is specified for only some of a_i's states (*e.g.*, $a_i\_says_{\phi}$ is specified only for state $s_{\phi}^{a_i}$) then that transition can only occur when a_i is in one of those states (a_i can not broadcast ϕ unless it agrees with ϕ).*

To start with we considered a fully connected network (FCN) of 10 agents. We seeded the network with one agent believing ϕ and one agent believing $\neg\phi$ and set the probability of infection, λ to 0.5. We created a reward function $\rho(s) = \sum_{i=0}^{n} : s^{a_i} = s_{\phi}^{a_i}$ (*i.e.,* the reward for a state s is the number of agents who believe ϕ in that state). Figure 6 shows that this network as quickly converges to a state where the expectation is that half the agents believe ϕ – the expected reward is 5.

We are not very interested in FCNs. Research in the information diffusion under the SIS model from early on has shown that the structure of the network has a big influence on the effectiveness of the contagion [26]. Mathematical anal-

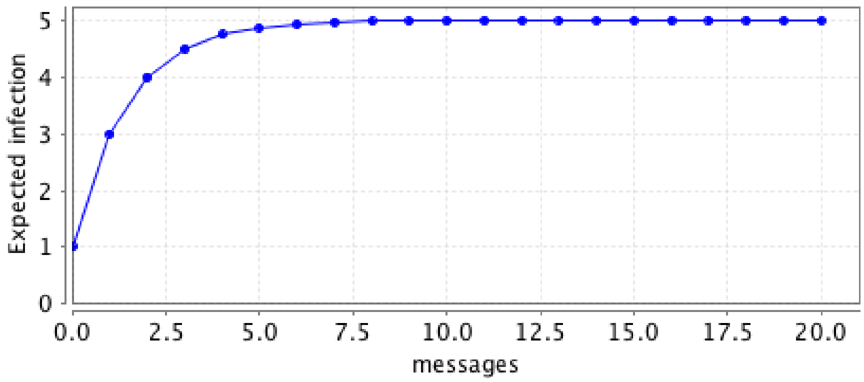

Fig. 6. Expected number of infected agents per message sent for model 2 on a FCN

ysis shows that the diffusion likelihood increases with the number of connections [19].

We want to have a "higher detail" insight into the impact a particular graph has on the spread of information. We generated a random network that satisfies the criteria for modelling a social network as a random graph as outlined in [23]: the maximal degree of separation is low, the probability of an edge between two agents is higher if they have mutual neighbours, and the network has a skewed degree distribution. This network contained 10 agent nodes, some with a minimum of 2 connections within the network and one with 8 connections. We initially studied the spread of ideas within this network with an λ of 0.5 and ϕ and $\neg\phi$ inserted in poorly connected agents (i.e., agents with only two connections within the network), well-connected agents (i.e., agents with six connections) and when the agent with idea ϕ had 8 connections while the agent with idea $\neg\phi$ had only 2 connections. The results are shown in Figure 7.

As it can be seen in the case where the initial agents have similar numbers of connections, the expected number of infected agents converges to 5 (converging more rapidly in the case where the initial agents have more connections). However in the case where the agent initially wishing to disseminate ϕ has more connections than the agent wishing to disseminate $\neg\phi$ then the number of agents believing ϕ converges to just under 6 – showing that the **initial advantage had a long term effect**. This result came as a surprise to us and, as far as we are aware, is not one that has been studied in the context of diffusion in the literature. We generated a further 9 networks (for a total of 10) and observed the same effect in all of them. However we were unable to investigate whether the same effect held for larger network sizes using PRISM.

We also investigated the probability that all agents in the network would become infected with the idea ϕ – i.e., that its opposite idea was completely eradicated from the network. In the case of the unbalanced starting position the

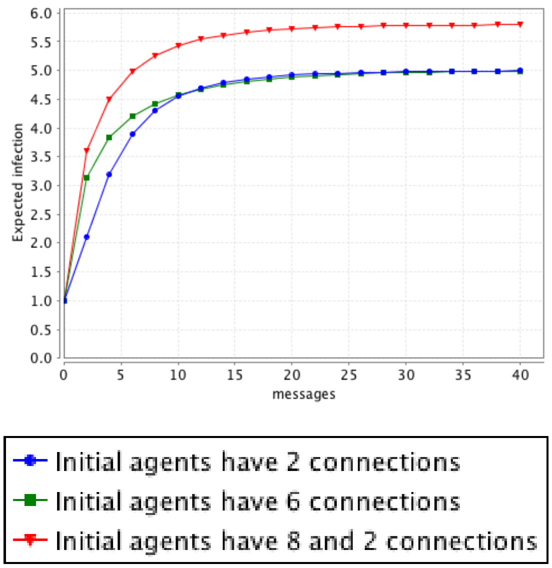

Fig. 7. Expected number of infected agents per message sent for model 2 on a randomly generated network

probability was 0.58 while in the balanced starting conditions the probability was 0.5.

5 Model-Checking Threshold Influence Models

We now consider threshold influence models. In these models it is not the receipt of a message bearing the idea ϕ that causes an agent to adopt the idea, but the perception that most of the agent's connections agree with ϕ.

We start with a FCN, as before.

Model 3 *We assume that an agent, a_i, can be in one of two states, $s_\phi^{a_i}$ (the agent publicly supports ϕ) or $s_\perp^{a_i}$ (the agent does not publicly support ϕ). All agents that publicly support ϕ broadcast the fact to all their neighbours, but we don't represent this as a transition in the network. Instead we have a joint transition, think, on all agents where they update a decision on whether or not they believe (or at least publicly support) ϕ themselves. Here the probability that they will adopt ϕ is proportional to the number of their connections who publicly support ϕ. n^{c_i} is the number of connections a_i has in the network and $n_\phi^{c_i}$ is the number of their connections who are broadcasting messages in support of ϕ so the chance of an agent deciding to espouse ϕ is $\lambda.\frac{n_\phi^{c_i}}{n^{c_i}}$ for some λ.*

This gives us the following transition system for agent a_i:

$$s_{\perp}^{a_i} \xrightarrow{think} \lambda.\frac{n_{\phi}^{c}}{n^{c}} : s_{\phi}^{a_i} \wedge (1 - \lambda.\frac{n_{\phi}^{c}}{n^{c}}) : s_{\perp}^{a_i} \tag{11}$$

$$s_{\phi}^{a_i} \xrightarrow{think} 1 : s_{\phi}^{a_i} \tag{12}$$

Figure 8 shows the expected number of agents who support ϕ, in a fully connected network of 10 agents, after T think transitions, given an influence probability of $\lambda = 0.5$. As can be seen this network converges to a state where we expect all 10 agents to support ϕ after 17 transitions.

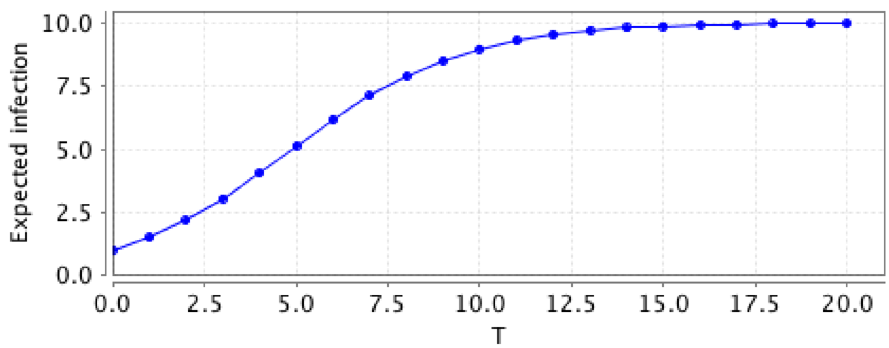

Fig. 8. Expected number of agents expressing ϕ after T transitions in model 3 on a FCN

Figure 9 shows the spread of ϕ on our more realistic network in the case where ϕ is first adopted by an agent with only 2 connections and in the case when ϕ is first is first adopted by an agent with 8 connections.

As in infection models, we can also add $\neg\phi$ into our influence model – with the chance that an agent expresses the opinion ϕ or $\neg\phi$ depending upon their perception of how many of their connections believe ϕ or $\neg\phi$.

Model 4 *We extend our transition system from model 3 with an agent state, $s_{\neg\phi}^{a_i}$ (the agent is expressing support for $\neg\phi$) and the variable $n_{\neg\phi}^{c_i}$ (the number of a_i's connections expressing support for $\neg\phi$). Therefore an agent's transitions become:*

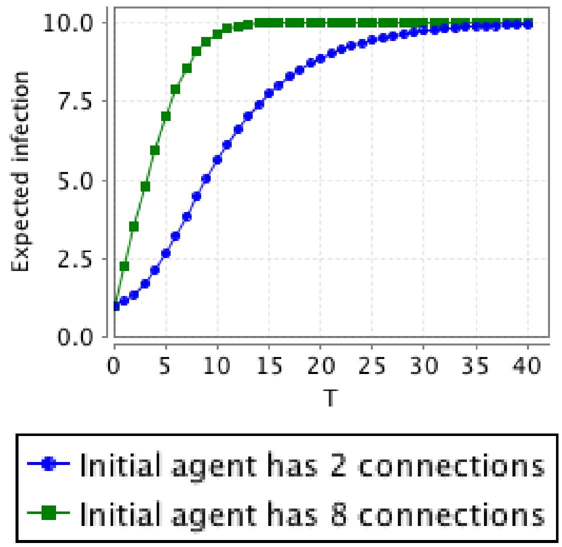

Fig. 9. Expected number of agents expressing ϕ after T transitions in model 3 on a randomly generated network

$$
\begin{aligned}
s_{\perp}^{a_i} \xrightarrow{think} \lambda.\frac{n_{\phi}^{c_i}}{n^{c_i}} : s_{\phi}^{a_i} \wedge \lambda.\frac{n_{\neg\phi}^{c_i}}{n^{c_i}} : s\neg\phi^{a_i} \wedge \\
(1 - \lambda.\frac{n_{\phi}^{c_i} + n_{\neg\phi}^{c_i}}{n^{c_i}}) : s_{\perp}^{a_i}
\end{aligned} \tag{13}
$$

$$
s_{\phi}^{a_i} \xrightarrow{think} \lambda.\frac{n_{\neg\phi}^{c_i}}{n^{c_i}} : s_{\neg\phi}^{a_i} \wedge (1 - \lambda.\frac{n_{\neg\phi}^{c_i}}{n^{c_i}}) : s_{\phi}^{a_i} \tag{14}
$$

$$
s_{\neg\phi}^{a_i} \xrightarrow{think} \lambda.\frac{n_{\phi}^{c_i}}{n^{c_i}} : s_{\phi}^{a_i} \wedge (1 - \lambda.\frac{n_{\phi}^{c_i}}{n^{c_i}}) : s_{\neg\phi}^{a_i} \tag{15}
$$

As before we start with a fully connected model the results of which are shown in Figure 10. Here, instead of all agent eventually expressing ϕ, we expected to reach a state where half the agents express ϕ (and by extension half are expressing $\neg\phi$).

Now we turn to our randomly generated model and examine the effect on the expected spread of ϕ, given the connectivity of the initial agents expressing ϕ and $\neg\phi$. The results are shown in Figure 11. In the infection models (Figure 7) the models converged to a state where half the agents adopted ϕ when both ϕ and $\neg\phi$ had similar starting states while it converged to a state where roughly 60% of the agents adopted ϕ when ϕ had an advantage over $\neg\phi$ at the start. In the case of influence models we see that the advantage conveyed by a better initial state is larger than it is in infection models, with the network converging to a state where we expect over 8 agents to be expressing ϕ.

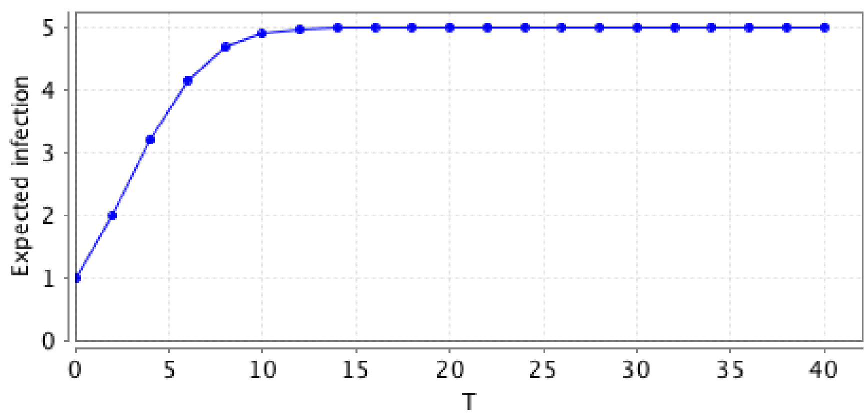

Fig. 10. Expected number of agents expressing ϕ after T transitions for model 4 on a FCN

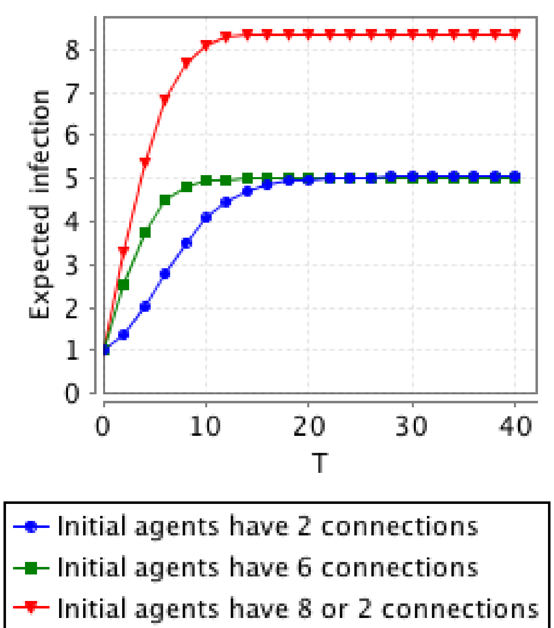

Fig. 11. Expected number of agents expressing ϕ after T transitions for model 4 on a randomly generated network

6 Analysing Larger Networks

Clearly networks of 10 agents are inadequate models of behaviour over large social networks. Unfortunately PRISM proved incapable of analysing larger models. In some cases PRISM couldn't even construct a larger model rendering even its simulation capabilities out of reach. Figure 12 shows the time taken to build a model of a FCN and to use PRISM model-checking to find the probability that the entire network would be fully infected with ϕ after 10 time steps for both model 2 and model 4.

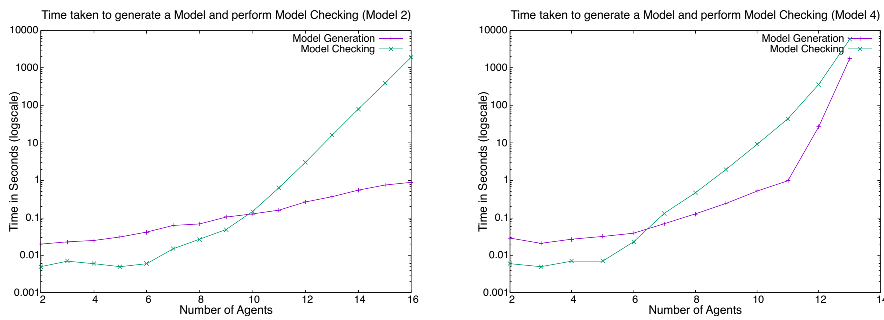

Fig. 12. Times taken to build a model and perform model checking for fully connected networks

The reasons for these problems are unclear[4]. While we did not expect to be able to model networks containing hundreds of agents in PRISM we had expected to model networks of more than ten. We have been invited to submit the social network models as challenge problems.

This does mean, however, that for the time being the use of model-checking as a tool for analysing information diffusion in social networks is limited although the same formalism can be used for simulation based analysis.

7 Related work

The influence of network structure on diffusion has been extensively studied in economics, see *e.g.*,[19] for an extensive literature list and [14] for a more general overview of the impact of social network structure on behaviour.

The methodology used to study network structure impact on diffusion throughout the literature is numerical analysis, simulation and experiments. Both micro and macro aspects of the network structure have been considered, but in both cases these aspects refer to statistical properties of the network. For example, a macro network aspect example is the degree distribution in the network, while a micro network aspect example is the average distance between two agents in the

[4] Ernst Moritz Hahn, private communication.

network and network component diameters. In nearly all diffusion models, the likelihood of adopting new information or behaviour increases with the increase of adjacent agents who have adopted it and a higher agent degree leads to higher contagion [19]. We also observe this here.

Bolzern et al's [3] approach is most similar to our own, using Markov chain models to capture network structure and to show how opinions among the agents in the network may vary among a fixed set of opinions (a generalisation of the idea of an idea, an "anti-idea", and indifference that we use here). However in their model the chance an agent will change its opinion does not depend upon its existing opinion, only upon the opinions of its neighbours. They use both formal analysis to generate results about the behaviour of the general system and monte-carlo simulation to analyse a specific system consisting of a fully connected network and two possible opinions.

Model-checking information in social networks has been studied from a theoretical perspective in [24] and [8]. Pardo and Schneider [24] consider the problem of verifying knowledge properties over social network models (SNM's) and shows that the model checking problem for epistemic properties over SNMs is decidable. Dennis et al [8] introduce a formal specification for SNM's and privacy properties that can be established to hold using model-checking using PRISM. Belardinelli and Grossi [2] present a model-checking algorithm and property specification logic for studying contagion-type models in open dynamic networks. This takes an agent view but does not explicitly consider the informational states of the agents. The proposed model-checking algorithm has not been implemented. Kouvaros and Lomuscio [15] use parameterised model-checking in the `MCMAS` system to study opinion formation protocols for swarm robotics. These protocols are similar to threshold models and involve agents in a swarm switching their opinion to the majority opinion of their neighbours. The interest in this work was primarily on answering whether the protocol guaranteed convergence to an opinion, not on analysing the behaviour of information diffusion itself and probabilistic aspects were not studied. Lastly, Zonghao et al [30] use PRISM to evaluate the efficacy of methods for controlling harmful network propagation using different protection strategies for individual nodes. Although Zonghao et al [30] are interested in security an protecting networks from e-viruses, the approach and methodology can be seen as related to ours in the case of information diffusion.

8 Discussion

While we have successfully made steps to account for internal informational states of agents in models of information diffusion. We have not successfully managed to use formal verification to analyse these models.

Clearly, there is no free lunch and, at least for now, there are technical limitations to the number of agents we can model. While we did not expect to model networks containing thousands of agents we had hoped that model-checking would provide a useful tool for exploration of networks of sufficiently large size to allow reasonable variation in network structure to be studied. This has not proved to be the case.

There are two approaches to overcoming this problem both of which we intend to pursue. We intend to continue using Markov chain models to study information diffusion in social networks – in particular we wish to study networks where an agent's informational state, its decision to broadcast a particular opinion and the decision of the network itself to propagate a broadcast to particular other agents all interact. While we could not use PRISM to simulate on our models we hope to either adopt or build a suitable alternative tool that can be used in this way.

Secondly, it is common in model-checking to develop abstractions of the problem which allow systems of realistic size to be studied. Work on parameterised model checking in swarm scenarios is also a promising avenue of research and, indeed, this is the approach taken in [15].

9 Summary

We have developed an initial framework for modelling information diffusion in social networks which takes an agent-centred view that includes an account of the agent's informational state when considering changes in the network. This framework uses Markov chain models to represent the agents within the network and their relationships to each other. Unfortunately even comparatively simple models proved intractable for analysing models of interesting size in PRISM, a current state-of-the-art tool for probabilistic model-checking.

Open Data

The PRISMmodels, network graphs, output and timing data reported in this paper can all be found in the University of Liverpool Data Catalogue DOI: https://doi.org/10.17638/datacat.liverpool.ac.uk/909.

References

1. Bailey, N.: The Mathematical theory of Infectious Diseases and its applications. Charles Griffin and Company Ltd., London (1975)
2. Belardinelli, F., Grossi, D.: On the formal verification of diffusion phenomena in open dynamic agent networks. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems. pp. 237–245. AAMAS '15, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2015), http://dl.acm.org/citation.cfm?id=2772879.2772912
3. Bolzern, P., Colaneri, P., Nicolao, G.D.: Opinion influence and evolution in social networks: A markovian agents model. Automatica **100**, 219 – 230 (2019). https://doi.org/https://doi.org/10.1016/j.automatica.2018.11.023, http://www.sciencedirect.com/science/article/pii/S0005109818305557

4. Christoff, Z., Hansen, J.U.: A two-tiered formalization of social influence. In: Grossi, D., Roy, O., Huang, H. (eds.) Logic, Rationality, and Interaction: 4th International Workshop, LORI 2013, Hangzhou, China, October 9-12, 2013, Proceedings, pp. 68–81. Springer Berlin Heidelberg, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40948-6_6
5. Christoff, Z., Hansen, J.U.: A logic for diffusion in social networks. J. Applied Logic **13**(1), 48–77 (2015). https://doi.org/10.1016/j.jal.2014.11.011
6. Christoff, Z., Hansen, J.U., Proietti, C.: Reflecting on social influence in networks. Journal of Logic, Language and Information **25**(3), 299–333 (2016). https://doi.org/10.1007/s10849-016-9242-y
7. Clarke, E.M., Grumberg, O., Peled, D.: Model Checking. MIT Press (1999)
8. Dennis, L.A., Slavkovik, M., Fisher, M.: "how did they know?" - model-checking for analysis of information leakage in social networks. In: COIN-XII 2016 International Workshops, COIN@AAMAS, Singapore, Singapore, May 9, 2016, COIN@ECAI, The Hague, The Netherlands, August 30, 2016, Revised Selected Papers. pp. 42–59 (2016). https://doi.org/10.1007/978-3-319-66595-5_3
9. Duflot, M., Kwiatkowska, M., Norman, G., Parker, D., Peyronnet, S., Picaronny, C., Sproston, J.: Practical applications of probabilistic model checking to communication protocols. In: Gnesi, S., Margaria, T. (eds.) FMICS Handbook on Industrial Critical Systems, pp. 133–150. IEEE Computer Society Press (2010), to appear
10. Fisher, M., Dennis, L.A., Webster, M.: Verifying autonomous systems. ACM Communications **Accepted for publication** (2013)
11. Grandi, U., Lorini, E., Novaro, A., Perrussel, L.: Strategic disclosure of opinions on a social network. In: Proceedings of the 16th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2017) (2017)
12. Granovetter, M.: Threshold models of collective behavior. American Journal of Sociology **83**(6), 1420–1443 (1978). https://doi.org/10.1086/226707
13. Jackson, M.O.: Social and Economic Networks. Princeton University Press, Princeton, NJ, USA (2008)
14. Jackson, M.O., Rogers, B.W., Zenou, Y.: The economic consequences of social-network structure. Journal of Economic Literature **55**(1), 49–95 (March 2017). https://doi.org/10.1257/jel.20150694
15. Kouvaros, P., Lomuscio, A.: Formal verification of opinion formation in swarms. In: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. pp. 1200–1208. AAMAS '16, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2016), `http://dl.acm.org/citation.cfm?id=2936924.2937099`
16. Kwiatkowska, M., Norman, G., Parker, D.: Stochastic model checking. In: Bernardo, M., Hillston, J. (eds.) Formal Methods for the Design of Computer, Communication and Software Systems: Performance Evaluation (SFM'07). LNCS (Tutorial Volume), vol. 4486, pp. 220–270. Springer (2007)
17. Kwiatkowska, M., Norman, G., Parker, D.: PRISM 4.0: Verification of probabilistic real-time systems. In: Gopalakrishnan, G., Qadeer, S. (eds.) Proc. 23rd International Conference on Computer Aided Verification (CAV'11). LNCS, vol. 6806, pp. 585–591. Springer (2011)
18. Kwiatkowska, M., Thachuk, C.: Probabilistic model checking for biology. In: Software Safety and Security. NATO Science for Peace and Security Series - D: Information and Communication Security, IOS Press (2014), to appear
19. Lamberson, P.J.: Linking network structure and diffusion through stochastic dominance. In: Complex Adaptive Systems and the Threshold Effect, Papers from the 2009 AAAI Fall Symposium, Arlington, Virginia, USA, November 5-7, 2009 (2009)

20. Li, M., Wang, X., Gao, K., Zhang, S.: A survey on information diffusion in online social networks: Models and methods. Information **8**(4) (2017). https://doi.org/10.3390/info8040118
21. Liu, F., Seligman, J., Girard, P.: Logical dynamics of belief change in the community. Synthese **191**(11), 2403–2431 (2014). https://doi.org/10.1007/s11229-014-0432-3
22. Mercier, H., Sperber, D.: The Enigma of Reason. Harvard University Press, London,UK (2017)
23. Newman, M.E.J., Watts, D.J., Strogatz, S.H.: Random graph models of social networks. Proceedings of the National Academy of Sciences **99**(suppl 1), 2566–2572 (2002). https://doi.org/10.1073/pnas.012582999
24. Pardo, R., Schneider, G.: Model checking social network models. In: Proceedings Eighth International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2017, Roma, Italy, 20-22 September 2017. pp. 238–252 (2017). https://doi.org/10.4204/EPTCS.256.17, `https://doi.org/10.4204/EPTCS.256.17`
25. Pedersen, T., Slavkovik, M.: Formal models of conflicting social influence. In: PRIMA 2017: Principles and Practice of Multi-Agent Systems - 20th International Conference, Nice, France, October 30 - November 3, 2017, Proceedings. pp. 349–365 (2017). https://doi.org/10.1007/978-3-319-69131-2_21
26. Rapoport, A.: Spread of information through a population with socio-structural bias: I. assumption of transitivity. The bulletin of mathematical biophysics **15**(4), 523–533 (Dec 1953). https://doi.org/10.1007/BF02476440, `https://doi.org/10.1007/BF02476440`
27. Schelling, T.: Dynamic models of segregation. Journal of Mathematical Sociology **1** (1971)
28. Smets, S., Velázquez-Quesada, F.: How to make friends: A logical approach to social group creation. In: Baltag, A., Seligman, J., Yamada, T. (eds.) Logic, Rationality, and Interaction - 6th International Workshop, LORI 2017 Sapporo, Japan, September 11-14, 2015, Proceedings. Forthcoming (2017)
29. Winship, C.: Threshold models of social influence. In: Bearman, P., Hedström, P. (eds.) The Oxford Handbook of Analytical Sociology. Oxford University Press (2011). https://doi.org/10.1093/oxfordhb/9780199215362.013.20
30. Zonghao, G., Ou, W., Peng, Y., Lansheng, H., Weiming, W.: Model checking probabilistic network propagation protection strategies. In: 2016 IEEE Trustcom/BigDataSE/ISPA. pp. 354–361 (Aug 2016). https://doi.org/10.1109/TrustCom.2016.0084