Analysis of the awareness impact on disease spreading in multi-layer networks

Pedro Santos
83550
Instituto Superior Técnico
Lisbon, Portugal

João Neves
83405
Instituto Superior Técnico
Lisbon, Portugal

Łukasz Niewiński
95018
Warsaw University of Technology
Warsaw, Poland

ABSTRACT

The presented work studies the interplay between the disease and awareness spreading in a multi-layer network. The epidemic process is not only regulated by a physical network (encoding real contacts between people) but also, indirectly, by a virtual network (encoding the disease's awareness spreading). Empirical results using scale free networks showed that the awareness spreading process can greatly contribute to slow down the spread of the disease. Moreover, the awareness spreading process can regulate the stationary prevalence of the disease and affect the epidemic threshold.

1 INTRODUCTION

The study of epidemic spreading processes is of extreme importance for human societies as it may allow to find effective control strategies to eradicate or mitigate diseases. The transmission of infectious diseases usually occurs by physical, direct interaction with infected individuals (exchange of biological elements).

Humans interact with each other under very different contexts. For example, an individual physically contacts with other people such as family members, friends and co-workers in their daily life. However, at the same time, the same individual can exchange information with other people without requiring the need of physical contact. Nowadays, mobile devices, mass media and online social networks greatly contribute to increase the amount of exchanged information.

Individuals tend to take extra "safety" measures when they know they are contacting someone infected, or even just because they are aware that the danger/chance of becoming infected is increasing in the community they live in. Therefore, it may be plausible to assume that the epidemic spreading is a result of the interplay between two different dynamical processes: the disease spreading process itself, that occurs at the layer of physical contacts, and the information/awareness spreading about a certain epidemic, that may occur not only at the physical level but also via non-physical interactions.

Multi-layer networks [1] are a generalization of the "traditional" network models that can be used to describe social interactions between individuals at different levels and contexts. Granell et al., 2013 [2] studied the interrelation between the disease and awareness processes in the spreading of epidemics, using multi-layer networks. The work is mainly focused on the study of the critical point for the onset of the epidemics, using analytical mathematical tools.

The present study closely follows the multi-layer network setting proposed in [2] but, instead, focus its attention to the study of the interrelation between both processes using computer simulations.

In the context of this work, the epidemic spreading process is modeled by an interplay between two different networks: the *physical/disease* network encodes the real physical contacts between individuals, therefore regulating the spread of the disease. On the other hand, the *virtual/awareness* network represents the non-physical interactions between the nodes that allow individuals to communicate information between each other. Both networks are composed of the same nodes but can have different topologies. With respect to the physical network, nodes can be in two states: susceptible (S) or infected (I). In the virtual layer, nodes can be in the unaware (U) or aware (A) states. Individuals on the unaware state do not have information about how to prevent infection, however, nodes on the aware state can reduce their risk of being infected (for example by avoiding direct contact with other people).

In the physical layer, infected nodes propagate the disease to their neighbours with a given probability β and infected nodes eventually recover with probability Δ. In the virtual layer, aware nodes make their neighbours aware of the disease with probability μ and aware nodes eventually forget the existence of the disease with probability Ω. Finally, aware nodes decay by a factor of λ their probability of getting infected.

It is assumed that infected nodes are always aware of the disease and, thus, each time a node gets infected it immediately gets aware of the disease. Therefore an individual can be in one of the states: susceptible and unaware <S,U>, susceptible and aware <S,A>, infected and aware <I,A> (note that the state <I,U> is spurious).

In summary, the joint epidemic and awareness spreading processes are parameterised by the following variables:

- Initially infected population size (γ): the number of initially infected nodes (i.e. the initial number of <I,A> nodes).
- Initially aware population (φ): the number of initially aware and susceptible nodes (i.e. the initial number of <S,A> nodes).
- Disease spread rate (β): probability of transitioning from the susceptible (S) state to the infected (I) state if node has at least one infected neighbour.
- Recovery rate (Δ): probability of transitioning from the infected state (I) to the susceptible (S) state.
- Awareness spread rate (μ): probability of transitioning from the unaware (U) state to the aware (A) state if node has at least one aware neighbour.
- Forgetness rate (Ω): probability of transitioning from the aware (A) state to the unaware (U) state.

- Infection decay (λ): decay in the probability of getting infected for nodes that are aware of the disease (e.g. if $\lambda = 0.5$ then the probability of getting infected decreases 50%; if $\lambda = 1.0$ aware nodes do not get infected; if $\lambda = 0.0$ then no decay).

1.1 Objectives

The main focus of this work is to empirically study, via computer simulations, how the awareness process can affect the spread of a disease. More specifically, the present study is focused on the following two major objectives:

- **Understand how the awareness process affects the spread velocity of a disease**: to accomplish this, the epidemic is modeled and studied using a SI model (section 3.1).
- **Understand how the awareness process affects the stationary prevalence distribution of a disease**: to accomplish this, the epidemic is modeled and studied using a SIS model (section 3.2).

1.2 Outline

The following section (section 2) describes the setup of the experiments: the networks' properties and generation procedure, as well as the simulation process. Section 3 presents the results of the experiments. Finally, section 4 discusses the main conclusions of this study.

2 EXPERIMENTS SETUP

2.1 Network description

In order to simulate the spread of an epidemic, 10 scale-free networks were generated using the Barabási–Albert model. All 10 networks have 10,000 nodes and 19,994 edges.

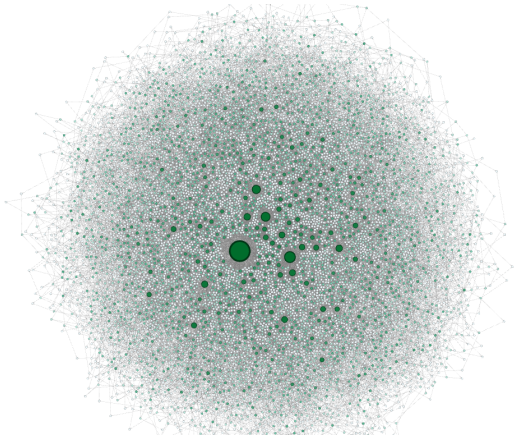

Figure 1: One of the 10 generated networks. Green, large nodes, are hubs with high degree. Small white nodes have low degree.

At each iteration of the generation process, a new node is created and attached with two edges to already existing nodes. This ensures all 10 networks have an average degree of 4. The generation process applies preferential attachment, meaning that on the insertion of a node, nodes with larger degree are preferred as a connection as opposed to nodes with smaller degree. The networks are static, as in, during simulation, their nodes and edges do not change.

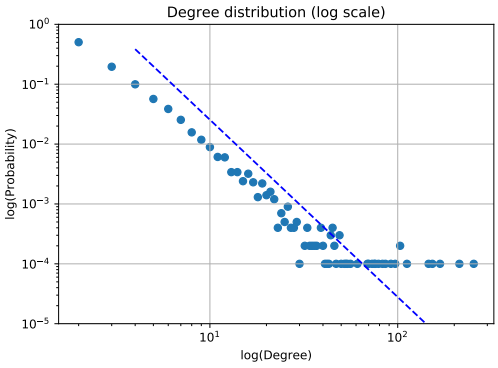

Figure 2: The degree distribution for the generated networks shows a power law distribution. The average γ coefficient is equal to 2.9.

The degree distribution and average power law γ coefficient is pictured in figure 2. Additionally, the largest hub in this network has a degree of 360. There are a few smaller hubs, and most (50.32%) of the nodes have a degree of 2 due to the preferential attachment.

2.2 Simulation process

All results were obtained using Monte Carlo simulations. Therefore, all plots except the 2D grid-plots (figures 11, 14 and 16), are computed by averaging the results of 2500 simulations: 250 simulations for each of the 10 networks. Due to limited computational power, the 2D grid-plots are computed using 125 simulations: 25 simulations for each of the 5 networks. For the presented results, the same network topology was used for both the physical and virtual layers (for the impact of the networks' topology on the simulations please check section 3.3).

3 EXPERIMENTS

This section presents the experimental results obtained using the SI and SIS multi-layer models. The simulation of both epidemic models was implemented from scratch in C++. All plots were produced with Python using the Matplotlib library [3]. The developed code can be found here:

https://github.com/PPSantos/network-epidemics.

3.1 SI model

This section assesses how the awareness process can influence the spreading speed of the epidemic disease. In order to accomplish this, both the awareness and disease processes are modeled using an SI model (infected nodes never recover from the disease). On the physical layer the disease spread rate β governs the speed of the epidemic spreading, while on the virtual layer, the awareness spread rate μ controls the speed of the information spread. For this section parameters Δ and Ω parameters are discarded.

While just considering the epidemic process, as time goes to infinity, the percentage of infected individuals will converge to 100%. However, when considering both processes, experimental results show that the awareness process can greatly slow down the epidemic spread. Moreover, for the special case of $\lambda = 1.0$, i.e. when the aware nodes become invulnerable to the disease, the awareness process can actually prevent a part of the population of becoming infected.

3.1.1 Infection spread rate (β)

The standard SI model (without the awareness spreading process) was briefly studied. In order to properly evaluate the awareness impact, the following results will be used as a baseline for all future comparisons. The only parameter regulating the standard SI model is β. As seen in figures 3, 4 and 5, higher β values impose a faster spreading of the epidemic. As β increases, the disease propagation gets faster and reaches its highest spread rate earlier (figure 4).

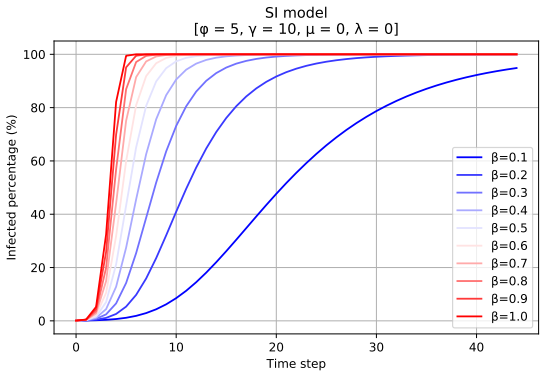

Figure 3: The impact of the infection rate (β) on the epidemic's spreading speed. Higher β values impose a faster spread.

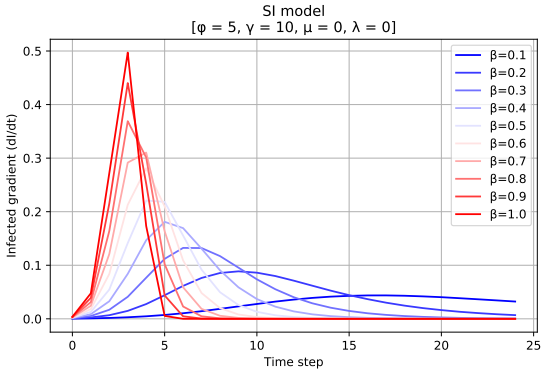

Figure 4: The impact of the infection rate (β) on the epidemic's spreading speed. Higher β values not only impose a higher gradient overall, but also push the timestep at which the highest gradient is recorded towards zero.

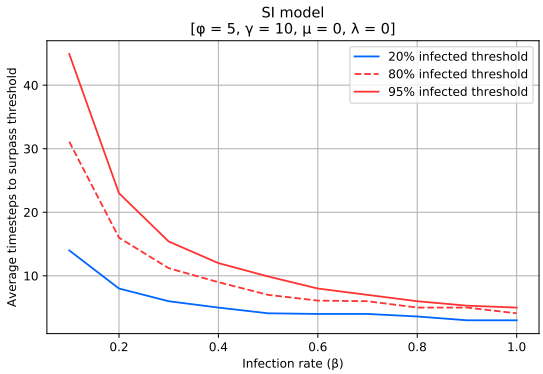

Figure 5: The impact of the infection rate (β) on the epidemic's spreading speed. As shown, the average time steps to exceed the respective thresholds are lower as β increases.

3.1.2 Awareness spread rate (μ)

The awareness spread rate μ controls the speed of the information exchange on the virtual layer. As seen in the following figures as μ gets higher, the epidemic spreading slows down. This happens because the information will spread faster on the virtual layer, therefore increasing the amount of individuals in the state <S,A> before the arrival of the epidemic at them. Since $\lambda < 1.0$ those individuals will resist to the disease for a number of time steps but will also become infected at some point in the future.

In the figure 6, for $\lambda = 0.6$ and independently of the value of μ, at time step 20 all the population is already infected. However, as seen in figure 7, higher μ values greatly contribute to smooth the growth in infected population throughout time steps.

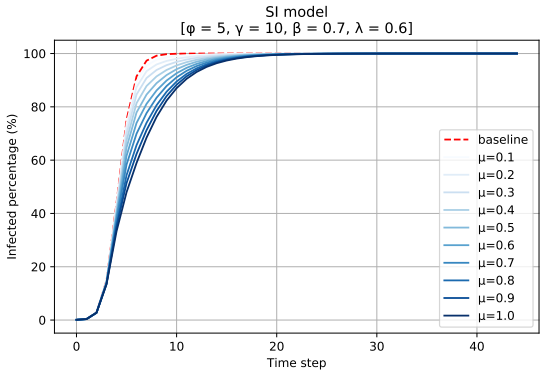

Figure 6: The impact of the awareness spread rate on the epidemic's spreading speed. Higher μ values impose a slower spread of the epidemic. The baseline curve corresponds to the equivalent SI process without the awareness spreading process.

Other interesting thing to notice is that μ barely affects the percentage of infected individuals during the first 5 time steps.

Since initial <I,A> and <S,A> nodes are randomly selected, happens that it takes on average 5 time steps for the infection to reach the <S,A> individuals. In other words, the infection can freely spread, on average, for the first 5 time steps without reaching susceptible and aware (<S,A>) nodes.

Finally, since $\lambda \neq 1.0$, based on the SI model assumptions and as time goes to infinity all nodes will become infected.

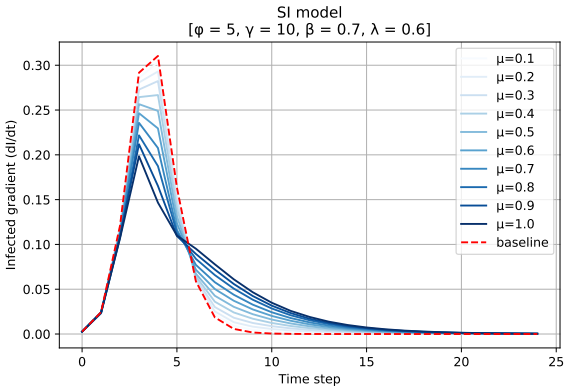

Figure 7: The impact of the awareness spread rate on the epidemic's spreading speed. Higher μ values impose a slower spread of the epidemic. The baseline curve corresponds to the equivalent SI process without the awareness spreading process. Higher μ values result in smoothing and extending the infection gradient.

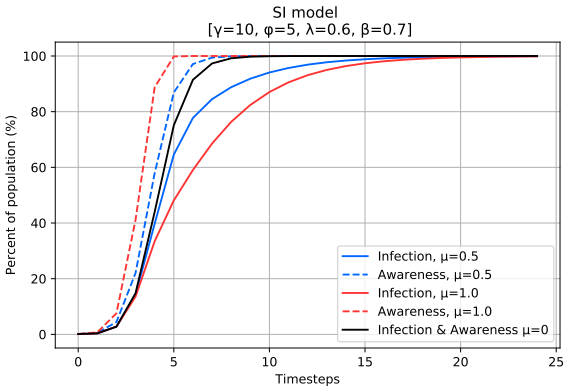

Figure 8: Infection and awareness curves for two different awareness spreading rates. Higher μ values impose a slower spread of the epidemic. Dashed lines represent the percentage of aware population per time step. Note that infected individuals are always aware, therefore, the gap between the curves of same color represents the susceptible but aware individuals <S,A>. Worth pointing is the fact that, for $\mu = 1$, the whole population gets aware within 5 time steps only.

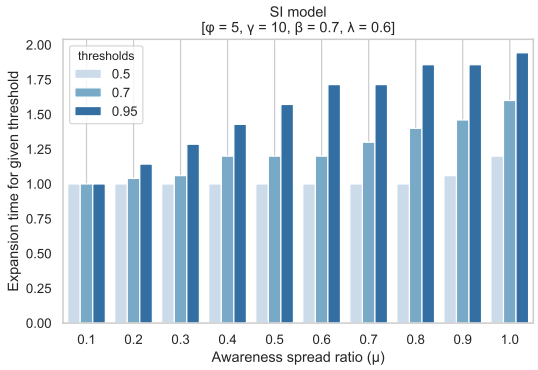

Figure 9: The impact of the μ parameter on the epidemic's spreading speed. Each bar encodes how many more times does that parameters setting take to reach the threshold in comparison to the baseline SI model (for example, a bar with an height of 2 means that, for that specific setting, the time to reach the given threshold was 2 times greater than the time to reach the same threshold while considering the baseline SI model). For the 0.7 and 0.95 thresholds, higher μ values impose a slower spread of the epidemic. However, μ barely affects the time needed to reach the 0.5 threshold. As seen on the chart, awareness significantly postpones reaching higher infection thresholds.

3.1.3 Infection decay (λ)

The infection decay λ controls the decrease in the probability of getting infected for the susceptible and aware individuals (<S,A>), in other words, the resistance of those individuals to the disease. For the case where $\lambda < 1.0$, as seen in figure 10, higher λ values greatly contribute to slow down the propagation of the epidemic.

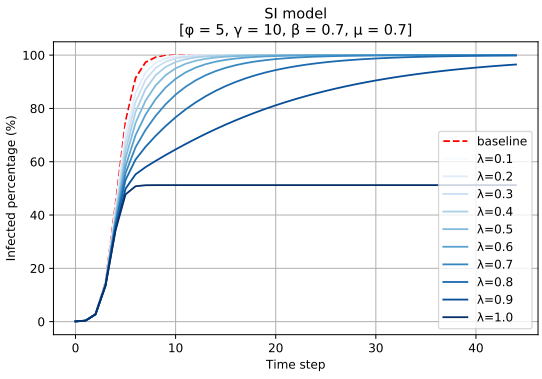

Figure 10: The impact of the λ parameter on the epidemic's spreading speed. For $\beta = 0.7$ and $\mu = 0.7$, the disease decay parameter significantly delays reaching infection thresholds above 50%. For the special case when $\lambda = 1$, the disease spread stops at 50%. The level at which the epidemic stops spreading depends on both β and μ.

For the special case where $\lambda = 1.0$, the number of infected individuals will converge to a stationary distribution (different than 100 % of infected nodes), as seen in figure 10. This distribution depends on the values of both β and μ, as seen in figure 11.

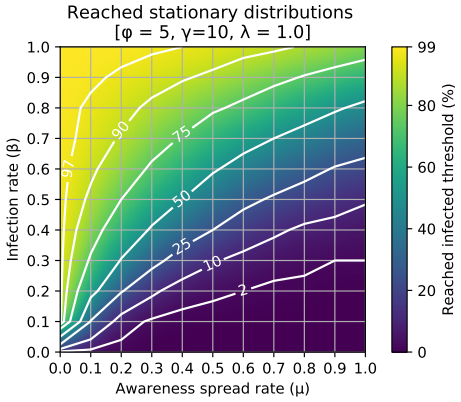

Figure 11: Stationary distributions for different μ and β values. The 100% threshold of infected population is never reached. For the presented SI model and $\lambda = 1$ (nodes once aware of the disease are immune to getting infected), the awareness and disease race to reach a higher number of nodes. As the heat-map presents, for low β and high μ, the majority of the nodes becomes aware before getting sick. On the other hand, for low β and high λ most of the nodes become sick before getting aware.

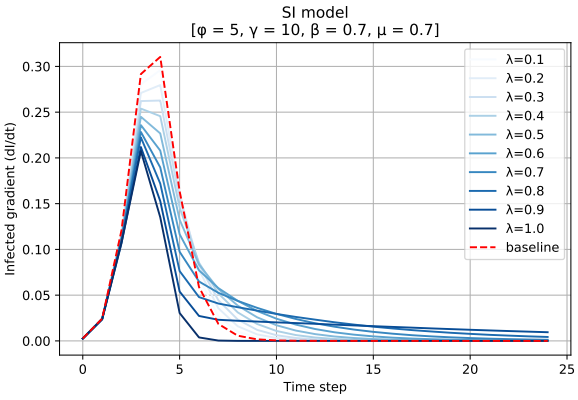

Figure 12: The impact of the infection decay parameter on the epidemic's spreading speed. For high λ values, the disease spread slows downs as it is infecting aware nodes with a lower probability of getting infected. Therefore, as a natural consequence it takes more time steps to reach 100% of infected individuals. For $\lambda = 1$, the spread stops earlier so the gradient reaches a value 0 first.

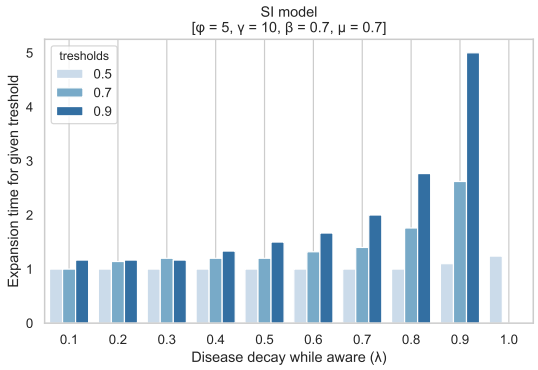

Figure 13: The impact of the λ parameter on the epidemic's spreading speed. Each bar encodes how many more times does that parameters setting takes to reach the threshold in comparison to the baseline SI model (for example, a bar with an height of 2 means that, for that specific setting, the time to reach the given threshold was 2 times greater than the time to reach the same threshold while considering the baseline SI model). For the presented parameters and $\lambda = 0.9$, it takes 5 times longer to reach the 90% threshold. Once again, it is possible to notice that, for $\lambda = 1$, high thresholds are not reached. Low λ values do not postpone infecting population significantly.

3.1.4 Multivariate μ and λ analysis

Although both μ and λ parameters are intrinsically correlated and dependent on one another, λ has a bigger impact on stopping the epidemic spread, as seen in figure 14.

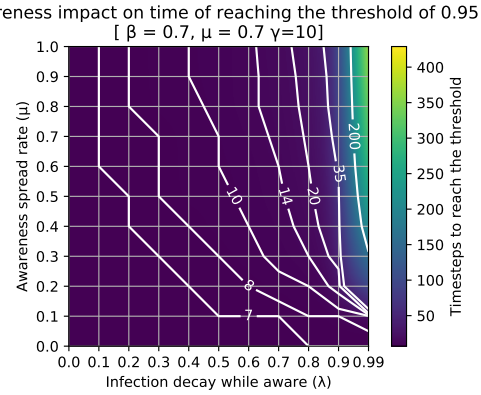

Figure 14: The impact of both the infection decay λ and the awareness spread μ on the time to reach a threshold of 95%. It is visible that the time steps to reach the threshold increase mostly with the increase of the infection decay parameter. As lambda approaches 1.0 the number of timesteps to reach the threshold approach infinity (as the threshold might never be reached). Check figure 11 for the special case where $\lambda = 1$.

For the given parameters setting, λ appears to be more important in the slow-down of the disease spreading in comparison to the awareness spreading rate μ. However, notice that the resistance to the disease only becomes noticeable if the nodes have the chance to get informed about the epidemic in the first place.

3.2 SIS model

This section assesses how the awareness spreading process can influence the stationary distribution of the epidemic disease. In order to accomplish this, both the awareness and disease processes are modeled using an SIS model. On the physical layer, the previously described SI model is now extended by allowing individuals to eventually recover from the infection with probability Δ. On the virtual layer, the previously described SI model is extended by allowing aware nodes to eventually forget about the existence of the disease with probability Ω.

While just considering the epidemic process, the stationary distribution of infected individuals is governed by both β and Δ. However, when considering both processes, experimental results show that the awareness process allows to decrease the stationary distribution of the epidemic. Moreover, the information spreading process can greatly contribute to increase the epidemic threshold.

3.2.1 Infection (β) and recovery (Δ) rates

The standard SIS model (without the awareness spreading process) was briefly studied. The stationary distribution of infected individuals is a result of the interaction between both β and Δ parameters. As expected, for a fixed β value, higher recovery rates contribute to reduce the observed stationary distribution of infected nodes, as seen in figure 15. Figure 16 displays the epidemic's stationary distribution as a function of both β and Δ parameters. Important to notice that, for almost all the phase space, the disease is on the endemic/active state. For all the combinations of tested parameters, the only combination with $\beta > 0$ that achieved an healthy state was ($\beta = 0.1, \Delta = 1.0$).

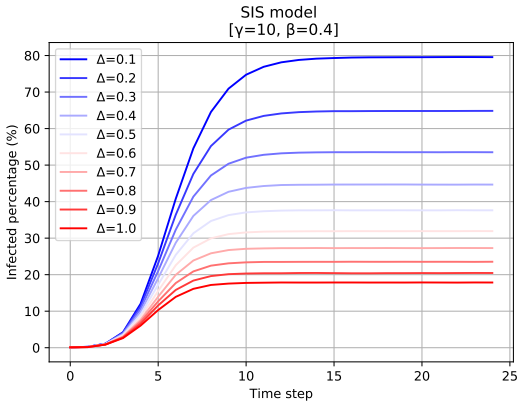

Figure 15: The impact of the recovery rate (Δ) on the epidemic's stationary distribution for a fixed disease spreading rate ($\beta = 0.4$). Higher Δ values impose a lower number of infected individuals as time goes to infinity.

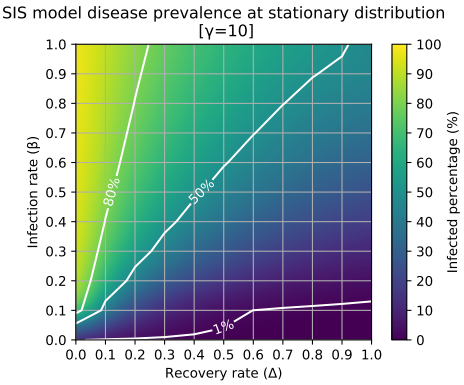

Figure 16: The impact of the infection rate (β) and the recovery rate (Δ) on the epidemic's stationary distribution. Higher β and Δ values contribute to, respectively, increase and decrease the stationary distribution of infected nodes. The plot was computed by discretizing each parameter into the $[0.0, 0.1, ..., 1.0]$ values.

3.2.2 Awareness spread (μ) and forgetness (Ω) rates

With respect to the awareness spread rate, while keeping all the other parameters fixed, as μ increases the epidemic's stationary distribution decreases, as seen in figure 17.

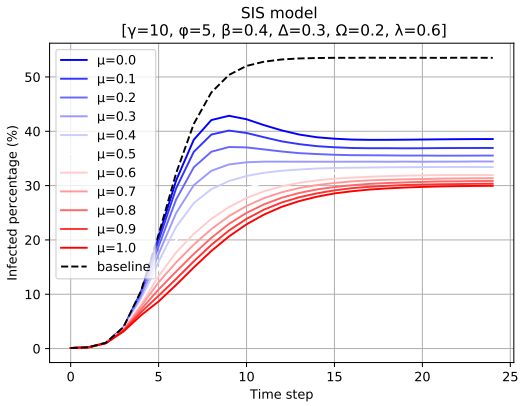

Figure 17: The impact of the awareness spread rate (μ) on the epidemic's stationary distribution for a fixed set of parameters. Higher μ values impose a lower number of infected individuals as time goes to infinity.

Interestingly, as seen in figure 17, the gap between the baseline and the line for $\mu = 0$ uncovers the fact that individuals, usually, do not forget that they were ill a short time after recovering from the disease and, therefore, during that time, they are extra careful to not get infected again. For the parameters setting used in the creation of figure 17, this happens because infected nodes (<I,A>) can recover from the disease (since $\Delta > 0$) and, therefore, become aware of the epidemic but not infected (<S,A>). Then, in turn, since $\lambda > 0$ these nodes will contribute to decrease the stationary distribution

in comparison to the equivalent SIS model that doesn't take into account the information spread.

The forgetness parameter Ω controls the rate at which aware individuals forget about the disease. As seen in figure 18, higher Ω values contribute to increase the stationary distribution of infected nodes. This result is intuitive because, for high Ω values, individuals more easily forget about the existence of the disease and therefore become more vulnerable to it.

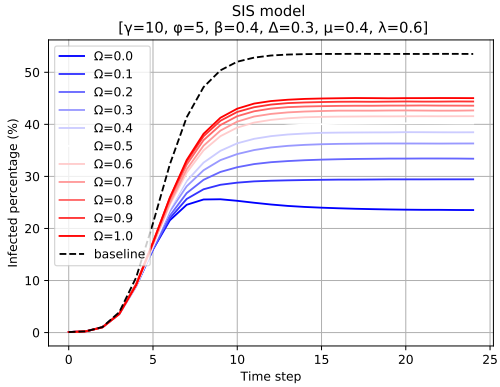

Figure 18: The impact of the forgetness rate (Ω) on the epidemic's stationary distribution for a fixed set of parameters. Higher Ω values impose a higher number of infected individuals as time goes to infinity.

3.3.3 Infection decay (λ)

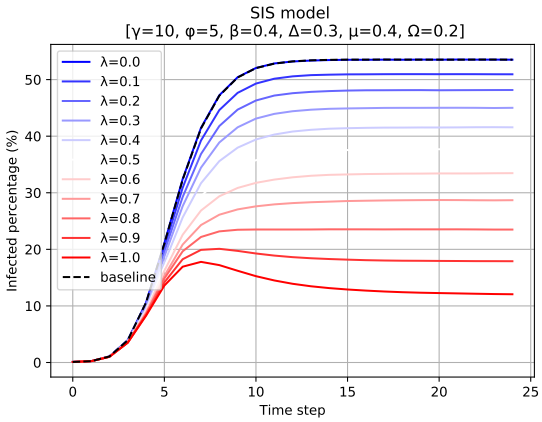

Figure 19: The impact of the infection decay parameter (λ) on the epidemic's stationary distribution for a fixed set of parameters. Higher λ values impose a lower number of infected individuals as time goes to infinity. Note that for $\lambda = 0$ the awareness process doesn't affect the stationary distribution.

Finally, figure 20 sums up the overall trend for the previously studied parameters.

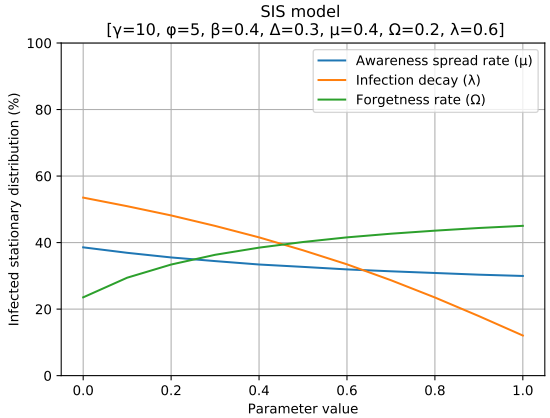

Figure 20: The impact of μ, λ and Ω on the stationary distribution. Each line was calculated by varying one parameter at a time, while keeping all the other fixed.

3.3.4 Epidemic threshold

As previously shown in section 3.2.1, under the standard SIS model, i.e. without considering the awareness spread, for almost all the phase space the epidemic is on the active state (low epidemic threshold). This is mostly due to the fact that the present work comprises the study of scale-free networks. Even in the best case scenario (when $\Delta = 1.0$), the epidemic threshold is ≈ 0.1, as it can be seen in figure 21 (blue dot).

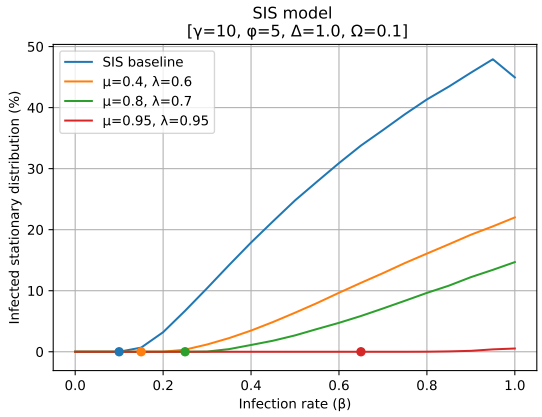

Figure 21: The stationary distributions (lines) and epidemic thresholds (points) for different settings of μ and λ parameters, with $\Delta = 1.0$.

However, interestingly, different results are obtained when the awareness process is also taken into account. As figure 21 illustrates, as the awareness process increasingly becomes more noticeable (by increasing the μ and λ parameters), not only the infected stationary distribution decreases independently of the β value, but also the epidemic threshold becomes larger.

For convenience, figure 21 shows the results obtained by setting $\Delta = 1$, but equivalent results are obtained for different recovery rates.

3.3 Topology impact

All previously presented results were obtained by using the same network topology for both the physical and virtual layers. However, further experiments were conducted in which two networks were picked from the set of initially generated networks, in such a way that different topologies are assigned to the physical and virtual layers. Experimental results showed that, for this new setup, all obtained results are equivalent to the previously presented results.

4 CONCLUSION

The presented study presents a comprehensive analysis of the awareness impact on disease spreading in a two-layer network. The study of the interplay between both the disease and awareness processes uncovers interesting findings about human behaviour, impossible to obtain while just considering the disease spreading process alone.

Firstly, on section 3.1, experimental results showed that the awareness process contributes to slow down the spread of the epidemic. Moreover, it is interesting to think again about the case when $\lambda \approx 1$, i.e. aware individuals are cautious/guarded or immune to the infection. The interplay between both the disease and awareness processes might help to explain why infectious diseases such as VIH/AIDS affect a much lower population than what predicted while considering the standard SI model, by formulating the following hypothesis:

Humans are good communicators (high μ)
+
Humans are consistently reminded about the disease (low Ω)
+
Humans take extra safety measures while contacting infected individuals and are afraid of getting infected ($\lambda \approx 1$)
+
(Humans have a finite life-time)
=
The stationary distribution of the disease is $<<$ 100%

Moreover, it is also interesting to show that the awareness process might clarify why VIH/AIDS affects a greater amount of population on developing countries in comparison to the developed countries: for developed countries μ and λ are higher and Ω lower in comparison to developing countries, therefore, contributing to reduce the stationary distribution of the disease. Also, as studied in section 3.1.4, being aware of the disease is not enough if extra safety measures are not taken (λ plays a huge role regulating the stationary distribution of the disease). This might also explain why the lack of resources in the developing countries might contribute to the spread of HIV, even in the case individuals are aware of the infectious disease. Concluding, the study presented in section 3.1.4 shows that, besides being important to aware individuals about the existence of the disease, it is crucial to properly educate them on how to prevent infection, as well as provide them the necessary conditions to fight the disease.

Secondly, on section 3.2, experimental results showed that the awareness process contributes to decrease the stationary distribution of the disease and increase the epidemic threshold. While just considering the standard SIS model (without the awareness spreading process) it was showed that it is hard for the disease to be on the healthy state. However, while taking the awareness process into consideration, the epidemic threshold can greatly increase, allowing the disease to be more easily kept in the healthy state and, therefore, under control.

Future work could comprise the study of how the network's properties, such as average-degree and assortativity, can affect both the disease and awareness spread. Minor tweaks to both the disease and awareness spreading processes could also be considered. For example, it could be interesting to study what would happen if individuals only realised they were infected (become aware of the disease) some time after being infected. Additional parameters could also be used to represent an increase in the speed of recovery only after individuals discovered they were infected - this could represent the beginning of a treatment process. Finally, it would be interesting to recompute some of the previously presented plots (especially the 2D grid-plots) with a finer granularity. Access to a computer-cluster would significantly reduce the computation time since the simulation process is embarrassingly parallel.

REFERENCES

[1] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, Jul 2014.

[2] Clara Granell, Sergio Gómez, and Alex Arenas. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Physical Review Letters*, 111(12), Sep 2013.

[3] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.