Analyzing Two Competing Epidemic Processes - The Power of Influential Nodes -

Colin Cooper[1], Robert Elsässer[2], Adrian Ogierman[3], and Tomasz Radzik[1]

[1] Department of Informatics
King's College London
Strand, London WC2R 2LS, UK
{colin.cooper,tomasz.radzik}@kcl.ac.uk

[2] Department of Computing Sciences
University of Salzburg
5020 Salzburg, Austria
robert.elsaesser@sbg.ac.at

[3] Institute for Computer Science
University of Paderborn
33102 Paderborn, Germany
adriano@upb.de

Abstract. Epidemic processes are widely used for the design and analysis of algorithms in distributed computing. On the other side, such processes can also model the spread of a malicious piece of information in a distributed environment. In this paper, we consider two competing epidemic type processes in a network modeled by a random graph with power law degree distribution. The first process $\mathcal{V}$ models the spread of a malicious piece of information, while the second one $\mathcal{P}$ starts with a certain delay w.r.t. $\mathcal{V}$, and distributes a message which immunizes (uninfected) vertices. A node infected by process $\mathcal{V}$ in some time step t spreads its malicious piece of information to $\gamma(\mathcal{V}) \geq \delta$ randomly chosen neighbors in step $t+1$, where δ is the minimum degree in the graph. Process $\mathcal{P}$ becomes activated on this node with some delay d, i.e., it spreads its immunizing message to all of its neighbors in step $t+d$. Depending on the model, the message of $\mathcal{P}$ may or may not be transmitted from an immunized node to its neighbors. If a node becomes immunized without being infected before, then it cannot be infected anymore and will also not take part in propagating the infection by $\mathcal{V}$.

We analyze the behavior of these processes in different propagation models and obtain two results. First we prove that if $\delta \geq \Omega(\log\log n)$, where n is the size of the network, then there will be in total at most $o(n)$ infected vertices, and within $\mathcal{O}(\log n)$ steps $n - o(n)$ vertices are immunized without becoming infected, with high probability. Our second and main result considers the case when $\delta = \mathcal{O}(1)$. We assume here that the immunizing message of $\mathcal{P}$ may be transmitted from an immunized node (not infected so far) to its neighbors, provided that the (infected) node v from which this immunizing message originates was infected at most $\mathcal{O}(\log \deg_v)$ time steps before. Then, within $O(\log\log n)$ steps $\mathcal{V}$ runs out, and almost all nodes are immunized without becoming infected at all. Our analyses extend and quantify the results of Funk et al. (PNAS 106(16), 2009 or Physical Review E. 81: 036118, 2010) on random networks with power law degree distribution.

Keywords: Epidemic algorithms, random graphs, power law distribution

1 Introduction

Epidemic processes build the core of many efficient distributed algorithms such as randomized broadcasting, gossiping, leader election, or network management. On the other side, such processes are also used to model the spread of infectious diseases or malicious pieces of information in networks. While in the first case the objective is to cover all nodes as fast as possible, the goal in the second case is to stop the spreading process before it affects considerable portions of the network.

In this paper we assume that both types of these processes occur at the same time in a network. The first process we call P_V tries to infect the network with a disease or a malicious piece of information, i.e., in each step any infected node spreads the "disease" to a randomly chosen subset of its neighbors (see next section for a formal definition). The second process, called P_P, starts with a certain delay at each infected node, and warns all (uninfected) neighbors of this node about process P_V. If a node receives the message disseminated by P_P, then it becomes immunized, and cannot be infected anymore. Depending on the model, the message of process P_P may or may not be transmitted from an uninfected (but immunized) node to its neighbors. The question is whether (under certain assumptions) P_P can stop process P_V before it manages to infect large portions of the network.

We consider these processes in random networks with power law degree distribution. First, we assume that the smallest degree in the network is $\Omega(\log \log n)$, and the message of P_P is transmitted from each infected node v to all of its (uninfected) neighbors d steps after v becomes infected, where d is some constant. Then, we consider the ase when the smallest degree is larger than a given constant, and the message of P_P can be transmitted from an uninfected informed (i.e., prevented) node to all of its neighbors, provided that certain conditions hold (see next section for the details). Our results provide evidence for the power of high degree nodes, which play a crucial role in stopping the malicious process at the right time.

1.1 Motivation

We consider two different society models which we call *skeptical* and *credulous* society, respectively. We start with the skeptical society model. In this case, if a malicious piece of information (or a disease) is spread, from an individual to another, then a person who discovers that it is infected by the malicious process warns its neighbors [11]. However, only the direct neighbors of an infected individual trust the source of the warning, and therefore the preventive message will not propagate further. In the second model, the individuals are more credulous and - depending on the source of the message - they might propagate a warning further in the network, even if none of their neighbors is directly affected. As before, if one discovers the malicious process, then he or she starts to warn others in the network. In addition, individuals informed by a preventive message spread the warning further to their neighbors if they trust the source of the message. That is, the originator of the message is influential enough and the message is not too old. Here, the influence of a person is determined by the number of its neighbors in the graph that models the social network [13].

In the first case, the influence of a node does not play any significant role. In the second case, the influence of a node determines for how long a preventive message is propagated in the network. This behavior of the process tries to mimic the fact that messages initiated by people with many acquaintances and a high standing in the society will more likely be propagated to distant nodes in the network. On the other side, if someone is not well known to others, then it is more likely that a message sent out by this person will rapidly disappear.

The goal of this paper is to compare the society models described above, and to determine under which conditions one can guarantee that most of the nodes are protected if a malicious piece of information circulates in the network.

1.2 Related Work

There is a significant amount of work considering epidemiological type processes in various models and on different networks. The simplest mathematical model for disease spreading is the so called SIR scheme (see e.g. [15, 23]). The population consists of three categories: susceptible (S), i.e., the individuals which are not infected yet but can become infected, infective (I), i.e., the individuals which are infected and can transmit the disease to others, and recovered (R), i.e., all individuals which recovered and are immune. In most papers, the spread of epidemics is modeled by a differential equation, which assumes that any susceptible individual has uniform probability β to get the disease from any infective individual. Furthermore, any infected person recovers at some stochastically constant rate γ.

This fully mixed model can easily be generalized to a network. It is known that such a process can be described by bond percolation on the underlying graph [14, 22]. Callaway et al. [7] considered this model on graph classes constructed by the so called configuration model (i.e., a random graph with a given degree distribution) [20]. The SIR model was also studied in some other scenarios, including various kinds of correlations between the rates of infection or the infectivity times, in graphs containing different types of vertices [22], or in networks with correlations between the degrees of the vertices [21]. According to [23], for certain graphs with a power law degree distribution, there is no constant threshold for the epidemic outbreak as long as the power law exponent is less than 3. This is, in fact, the case in most real world networks, e.g. [10, 1, 3, 24]. If the graph is embedded into a low dimensional space, or has high transitivity, then there might exist a non-zero threshold for certain types of correlations between vertices. However, none of the papers above considered the impact of a warning process on the behavior of an epidemic (such as the one analyzed in [11]).

There are also plenty of papers analyzing the influence of certain countermeasures on the outbreak of an epidemic (e.g. [9, 12]). We only mention here the work of Borgs et al. [6]. This paper focused on how to distribute antidote to control epidemics. The authors analyzed a variant of the contact process in the *susceptible-infected-susceptible (SIS) model* on a finite graph in which the cure rate is allowed to vary among the vertices. That is, the rate ρ_v at which an infected node v becomes healthy is proportional to the amount of antidote it receives, given a fixed amount of antidote $R = \sum_{x \in V} \rho_x$ for the whole network.

However, in reality the spreading process is influenced by many factors. One of these factors is the awareness of the affected individuals about the epidemic. In [11] Funk et al. concentrated on this type of (natural) countermeasure, and analyzed the spread of awareness on epidemic outbreaks. That is, in addition to the disease the information about the epidemic is also spread in the network and has its own dynamic. In [11] the authors described the two spreading scenarios (awareness vs. disease) by the following model. Each individual has a level of awareness which depends on the number of hops the information has passed before arriving to this individual. This was combined with the traditional SIR model. It has been shown that in a well mixed population, the spread of awareness can result in a slower outbreak, however, it will not change the threshold of the epidemic. Nevertheless, if the spread of awareness about a disease is considered in certain (geometric) networks as a local effect in the proximity of an outbreak, then this process can completely stop the epidemic. Moreover, the impact of spreading awareness is even intensified if the social network of infections and informations overlap.

In [19] Liu et al. analyzed the influence of two parameters - the decay rate of the disease and the range a person can be infected in - on the spread of an epidemic in the Haizhu district of Guangzhou. The results imply the importance of both parameters. Especially the results of the distance parameter, which is influenced by peoples movement, imply significant impact on the disease spreading.

Concerning mobility in a 2D field, Valler et al. [25] examined the epidemic threshold for a mobile ad-hoc network. They showed that if the connections between devices are defined by a sequence of matrices $A_1, \dots A_T$, then for $\lambda_S < 1$ no epidemic outbreak occurs, with high probability, where λ_S is the first eigenvalue of $\Pi_{i=1}^{T}(1-\delta)I + \beta A_i$. Here, β and δ are the virus transmission probability and the virus death probability, respectively. They also approximated the epidemic threshold for different mobility models in a predefined 2D area.

In addition to the papers described above, plenty of empirical work is known [4, 17, 18, 2, 5]. Two of the most popular approaches to simulate the movement of individuals are the so called agent-based and structured meta-population-based models, respectively (cf. [2, 16]). To combine the advantages of both systems, hybrid environments were implemented (e.g. [5]). Many of the models used in these simulations integrate different locations like schools, theaters, etc... An agent may or may not become infected depending on his own movements and the ones made by other agents. Such agents are usually defined very precisely, including e.g., race, gender, educational level, age, etc. [17, 18]. Thus, this may result in a huge amount of very detailed data, which requires immense computing power. In our paper, however, we do not focus on empirical analyses at all.

1.3 Our Results

In this section we briefly describe our results. More details can be found in Section 2 and 3, respectively.

First, we show that, under the assumption of a constant delay l and a minimum degree of $\Omega(\log\log n)$, the infection disseminated by P_V in the *skeptical society model* will be eliminated in $\mathcal{O}(\log n)$ time steps, with high probability. Furthermore, with the same probability the total number of infected individuals will not exceed $o(n)$. Our result implies that an unbounded but relatively small degree is enough to prevent almost the whole population to be infected in this model.

Second, we show that $\mathcal{O}(\log\log n)$ time steps are sufficient to eliminate a malicious piece of information in the *credulous society model*. The preventive messages of process P_P immunize at least $n - o(n)$ nodes, with high probability, in the given time frame. We can interpret this as follows. Due to the social structure, messages from individuals with significant influence are treated with a high priority. Thus, they initiate a tree structure in which each level represents the number of hops a message traverses from the source of the message to the corresponding nodes. Now, even if the individuals that are far away from the origin of a certain preventive message ignore the warnings, the huge amount of messages which are generated leads to a very fast embankment of the malicious process.

As mentioned in the introduction, in the credulous model individuals informed by the preventive message may transmit these messages further to their neighbors if certain conditions are fulfilled. On the other side, infected nodes in the network initiate process P_P with a certain delay after they are affected by P_V. However, if a preventive message arrives at an infected node, then this preventive message is **not transmitted further** in the network. Therefore, many infected nodes will block preventive messages. This might lead to a situation, in which most preventive messages are stopped after a few steps, even if the source of the message is a high degree node. The main challenge in our analysis was to show that this will not happen with high probability, i.e., most preventive messages will not be blocked. To perform the analysis, we split these *competing parts* into two major partitions: notification trees with depth at most $\mathcal{O}(\log\log n)$ on one side, and all other notification trees on the other side. Since the nodes of the first type may disturb the spreading process performed by the second kind of nodes, whenever the same informed node v belongs to two trees of two different types, v simply acts as a blocking node. That is, v does not forward the message belonging to the notification tree with the root of higher degree.

2 Model and Basic Definitions

In this section we present some basic concepts, models, and notation needed for the understanding of the upcoming formal analysis. We start by describing the network (also called environment), which is an undirected graph used by the dissemination processes. Afterwards, said dissemination processes are defined precisely.

Network Model We model the network used by the message dissemination process as an undirected random graph $G = (V, E)$ with n nodes and m edges. The nodes follow a power law degree distribution. Let $\deg_v$ denote the degree of a node $v \in V$. Then v has degree $\deg \geq \delta$ with probability proportional to $\deg^{-\alpha}$, where the constant $2 < \alpha < 3$ is the so called power law exponent and $\delta \in \mathcal{O}(\log\log n)$ denotes the minimum degree in G. Consequently, $m = {}^1\!/_2 \sum_{\forall v \in V} \deg_v = \Theta(n \cdot \delta)$. Further we assume that the maximum degree is proportional to $n^{1/\alpha}$.

Edges in G are randomly created using the so called *Configuration Model*. Note that the application of the configuration model may generate self-loops, disconnected components, and double edges. However, since $\delta \in \Omega(1)$ large enough and $2 < \alpha < 3$, G will be connected with high probability with at most a negligible amount of self loops and double edges. Additionally, all nodes have access to a global clock and work synchronously.

The nodes represent entities with states which are set accordingly to the dissemination processes. We model said states utilizing three categories/sets: uninformed/susceptible ($\mathcal{U}$), i.e., the nodes which are not notified yet in any way, infective ($\mathcal{I}$), i.e., the nodes which did not recover so far, and recovered ($\mathcal{R}$), i.e., all nodes $v \in \mathcal{I}$ which recovered and all nodes $v' \in \mathcal{U}$ which were already notified exclusively by the dissemination process. Hereby a node $v \in \mathcal{I}$ recovers (or heals) automatically after a specific delay d. The corresponding sets to a specific round i are denoted by $\mathcal{U}(i), \mathcal{I}(i)$ and $\mathcal{R}(i)$, respectively. If it is clear from the context, we simply use $\mathcal{U}, \mathcal{I}$ and $\mathcal{R}$.

For the analysis we need some additional structure based on the previously mentioned states. Let G_k be defined as the set (also called group) containing all nodes with degree k. Likewise let $G^{\mathcal{U}}_{k,i} = \{u \mid u \in \mathcal{U}(i), \deg_u = k\}$ be the group of all yet uninformed nodes in round i with degree exactly k.

Dissemination Processes The nodes in the network represent entities the message dissemination processes try to notify. We consider two main types of dissemination processes: 1.) harmful processes, i.e., dissemination processes aiming to inflict damage to the nodes, and 2.) remedying processes, i.e., dissemination processes counteracting the effects of the previously mentioned harmful processes. Our model includes three competing dissemination processes. A harmful process $\mathcal{V}$ with the goal to infect as many nodes with a malicious piece of information as possible, and two remedying processes Φ and Λ anxious to stop $\mathcal{V}$. Note, only one remedying process, depending on the model, is used to combat the harmful one (cf. below).

Process $\mathcal{V}$ The harmful dissemination process in our model is called $\mathcal{V}$. Its goal is to infect as many nodes of the network as possible and therefore inflict damage and/or costs. In the following we often use the term *notify* and *infect* with respect to $\mathcal{V}$ as synonyms. In short, in each round i process $\mathcal{V}$ proceeds as follows. Each newly infected node, i.e., each node that got infected in round $i-1$, *tries to infect* $\gamma(\mathcal{V})$ many of his neighbors in round i. Recall, only thus far uninfected nodes may be infected. Now we are ready to describe $\mathcal{V}$ more formally.

Let $u \in V$ be the starting point for $\mathcal{V}$, i.e., the first node infected by $\mathcal{V}$. In the first round $\mathcal{V}$ is only able to notify (a limited amount of) neighbors of u, i.e., $\mathcal{V}$ may notify at most $\gamma(\mathcal{V})$ different neighbors $\{v_1, ..., v_{\gamma(\mathcal{V})}\} \subseteq \Gamma(u)$, where $0 < \gamma(\mathcal{V}) < \delta$ and $\Gamma(u) = \{v \in V \mid \{v, u\} \in E\}$, of u. Let $\mathcal{V}_i = \mathcal{I}_i \setminus \mathcal{I}_{i-1}$ be the set of nodes $\{v_1, ...v_j\}$ of G which got notified by process $\mathcal{V}$ in round $i-1$. Then, each node $v \in \mathcal{V}_i$ chooses $\gamma(\mathcal{V})$ many of his neighbors $\{v_1, ..., v_{\gamma(\mathcal{V})}\} \subseteq \Gamma(v)$ uniformly at random and notifies them.

Consequently, $\mathcal{V}_0 = \{u\}$ while $\mathcal{V}_1$ is a randomly chosen subset of $\Gamma(u)$ of size $\gamma(\mathcal{V})$. In general, $\mathcal{V}_{i+1} = \mathcal{U}(i) \cap \bigcup_{u \in \mathcal{V}_i} \xi(u)$, where $\xi(u)$ is a randomly chosen subset of $\Gamma(u)$ of size $\gamma(\mathcal{V})$.

Process Φ Process Φ is one of two remedying processes we use to counteract $\mathcal{V}$. By definition, its goal is to combat $\mathcal{V}$ by reactively curing all infected nodes and proactively immunizing possible contact points of $\mathcal{V}$. In the following we often use the term *notify* and *immunize* with respect to Φ as synonyms. In short, process Φ is *activated* for each infected node after some delay of d rounds. W.l.o.g. let v be such a node and let i be the round Φ is activated on v. Then v notifies all its neighbors about $\mathcal{V}$ in round i, i.e., in the same round. All v's yet uninformed neighbors become immunized and v gets cured itself. However all v's already (and still) infected neighbors remain infected.

Let $\Phi_i = \mathcal{V}_{i-d}$ be the set of nodes $\{v_1, ...v_j\}$ of G where Φ is activated on in round i. It follows that process Φ is activated (also called triggered) on any node $v \in \mathcal{I}(i-1)$ in round i if and only if $v \in \mathcal{I}(i-d) \setminus \mathcal{I}(i-d-1)$. Further, once Φ was activated on a node $v \in \Phi_i$, v notifies all its neighbors and remains silent for the rest of the dissemination process. More formally:

$$\begin{aligned} \forall i \geq d \ : \ & \Phi_i = \mathcal{V}_{i-d}, \\ & \mathcal{I}(i) = \bigcup_{i-d<j\leq i} \mathcal{V}_j, \\ & \mathcal{U}(i) = V \setminus \left(\mathcal{I}(i) \cup \bigcup_{j<i} \Gamma(\Phi_j) \right), \\ & \mathcal{R}(i) = V \setminus (\mathcal{U}(i) \cup \mathcal{I}(i)) \end{aligned}$$

Process Λ Process Λ is the second remedying process in our model. By definition, its goal is identical to the one of Φ. However its behavior is more complex. In the following we often use the term *notify* and *immunize* with respect to Λ as synonyms.

Before we are able to describe Λ, we need to define the term *notification tree*. In short, a notification tree is a tree-like structure, starting with the root node, representing an *who-infected-whom-relationship* of the contained nodes. More formally, a *notification tree* $\Upsilon_{v,i}$ in G with root v in round i and depth j is a tree within G with the following property. Each (previously uninformed) node in $\Upsilon_{v,i} \setminus \{v\}$ was notified by process Λ and Λ only in round i', with $i - j \leq i' \leq i$, by another node of $\Upsilon_{v,i}$. W.l.o.g. let u and u' be the root nodes of two distinct notification trees in G with $\deg_u < \deg_{u'}$. Then if a node v' gets notified by two nodes $w \in \Upsilon_{u,k}$ and $w' \in \Upsilon_{u',k}$, then v' is assigned to the notification tree $\Upsilon_{u,k}$. That is, v' is always assigned to the notification tree with the minimum degree of the root among all competitors.

Let $\Lambda_i = \mathcal{V}_{i-d} \bigcup_{\forall \Upsilon_{v,i}} (\Upsilon_{v,i} \setminus \Upsilon_{v,i-1})$ be the set of nodes that may participate in Λ in round i. That is, Λ_i contains all nodes where Λ is activated on in round i as well as all uninformed nodes notified by Λ (for the first time) in round $i-1$. More formally, we define $\Upsilon_{v,i}$ recursively as

$$\Upsilon_{v,i} = \begin{cases} \{\emptyset\} & i < t(v) \\ \{v\} & i = t(v) \\ \Upsilon_{v,i-1} \cup [\Gamma(\Upsilon_{v,i-1} \cap \Lambda_{i-1}) \cap \mathcal{U}(i-1)] & t(v) \leq i \leq t(v) + \log(\deg_v) \\ \Upsilon_{v,i-1} & i > t(v) + \log(\deg_v) \end{cases}$$

where $t(v)$ is the round v was initially infected by $\mathcal{V}$ and $\Gamma(V') = \bigcup_{v' \in V'} \Gamma(v')$ for some subset $V' \subseteq V$. Note that infected or already immunized nodes are not considered as part of this notification tree. Further note that the number of rounds a notification tree with root v may participate in Λ is limited by $\log(\deg_v)$.

Now the description of Λ is quite intuitive. Similar to process Φ, Λ is activated on any node $v \in \mathcal{I}(i-1)$ in round i if and only if $v \in \mathcal{I}(i-d) \setminus \mathcal{I}(i-d-1)$. Then each newly activated node $v \in \mathcal{I}(i)$ forms a new notification tree $\Upsilon(v,i)$. This notification tree will then try to notify uninformed nodes for at most $\log(\deg_v)$ many rounds. Note that only newly immunized nodes, or the ones Λ is activated on for the first time, participate actively in Λ. That is, similar to process Φ, said nodes notify all their neighbors for one time only and remain silent afterwards.

Models During the analysis we consider two fixed combinations of the harmful and the remedying processes. We allow only one remedying process to combat $\mathcal{V}$ at once. The combination of $\mathcal{V}$ and Φ on the network G is called *Model 1* and the one of $\mathcal{V}$ and Λ *Model 2*.

3 Analysis

In this section, we present a formal analysis for Model 1 and Model 2 separately. As we mentioned already, we make use of the Principle of Deferred Decisions and the Configuration Model in the analysis. Let throughout this section $s(V') = \sum_{v' \in V'} s(v')$ be defined as the amount of stubs for some subset $V' \subseteq V$. Further let $s_f(V') = \sum_{v' \in V'} s_f(v')$ be defined as the amount of free stubs for some subset $V' \subseteq V$.

Model 1 In this section, we prove the following theorem:

Theorem 1. *Let G be a network with minimum degree $\delta = \Theta(\log \log n)$. Further let Model 1 be in place with delay $d > 1$ being a constant. Then independent of the choice for $\gamma(\mathcal{V}) \in \Theta(1)$, $\mathcal{V}$ cannot infect more than $o(n)$ many nodes, with probability $(1 - o(1))$.*

This theorem states that it suffices to immunize only the direct neighborhood of a previously infected entity to stop a malicious message dissemination process, provided the minimum degree is large enough. Although the minimum degree is unbounded it indeed is very small compared to the maximum degree.

Before we present the formal proof, we want to clarify the main idea on a higher level. We take advantage of the unbounded minimum degree. The proof consists of two main parts. In the first part we show that if no more than $\mathcal{O}(n)$ many stubs, up to a step j, were already used by Φ and $\mathcal{V}$ combined, then $\mathcal{V}$ is able to continue its message dissemination quite comfortably. That is, $\mathcal{V}$ will keep growing exponentially. Afterwards, we show that Φ manages to immunize $n - o(n)$ many nodes in step $j+1$ and conclude that therefore no more than $o(n)$ many nodes can be infected by $\mathcal{V}$ overall.

Proof. [Theorem 1]

This proof consists of two main parts. In the first part we focus on the ability of $\mathcal{V}$ and Φ to continue their message dissemination provided that the overall amount of stubs used up by Φ and $\mathcal{V}$ combined is bounded by $\mathcal{O}(n)$. Note, the overall amount of stubs in the network is $\Theta(n \cdot \log \log n)$ since $\delta = \Theta(\log \log n)$. In the second part then Φ is able to immunize $n(1-o(1))$ within one round leaving $\mathcal{V}$ without any chance to inflict more damage.

To be more specific we show the following properties:

1. $\mathcal{V}$ grows at least by a constant with probability $1-o(1)$ throughout Part 1,
2. each node v provides at least $3/4 \deg_v$ many free stubs overall after being notified by $\mathcal{V}$ or Φ throughout Part 1, and
3. in the first round of Part 2, Φ immunizes $n-o(n)$ many nodes.

Combining the above properties we conclude that $\mathcal{V}$ cannot notify more than $o(n)$ many nodes overall with probability $1-o(1)$.

Part 1 Note that the network contains n nodes and $\delta = \Theta(\log\log n)$. Consequently $S_f(G) = \Theta(n \cdot \log\log n)$ at the very beginning. Let round j be the last round of Part 1. Assuming $|s(\Phi_i \cup \mathcal{V}_i)| = \mathcal{O}(n)$, for $1 \le i \le j$, it follows that G provides at least $|s(\mathcal{U}(i))| = S_f(G) - |s(\Phi_i \cup \mathcal{V}_i)| = \Omega(n \cdot \log\log n)$ many free stubs. Recall, the maximum degree is bounded by $n^{1/\alpha}$. Thus the probability for an arbitrary node $v \in \mathcal{V}_i \cup \Phi_i$ to be paired with another node $w \in \mathcal{U}(i)$ at least $\gamma(\mathcal{V})/c$ many times, where $c < \gamma(\mathcal{V})$ is some suitable constant, is upper bounded by

$$\binom{\gamma(\mathcal{V})}{\gamma(\mathcal{V})/c}\left(\frac{\deg_w}{\Omega(n \cdot \log\log n)}\right)^{\frac{\gamma(\mathcal{V})}{c}} \le \left(\frac{ce\gamma(\mathcal{V})}{\gamma(\mathcal{V})}\right)^{\frac{\gamma(\mathcal{V})}{c}}\left(\frac{\deg_w}{\Omega(n \cdot \log\log n)}\right)^{\frac{\gamma(\mathcal{V})}{c}} = o(1/n^{\Omega(1)}). \quad (1)$$

Further, the probability for a node in $\Phi_i \cup \mathcal{V}_i$ to be paired with an arbitrary node of $\mathcal{I}(i)$ is bounded by

$$\Omega\left(\frac{1}{n \cdot \log\log n}\right) = \frac{|s(\Phi_i \cup \mathcal{V}_i)|}{\Omega(n \cdot \log\log n)} = \mathcal{O}\left(\frac{1}{\log\log n}\right). \quad (2)$$

Consequently at most a $o(1)$-fraction of all pairings of nodes in $\Phi_i \cup \mathcal{V}_i$ will be made with nodes in $\mathcal{I}(i)$. Utilizing the union bound for the lower bound of Equation 2 and martingale techniques for the upper one yields the same result with high probability. Similarly, the probability for an arbitrary node $v \in \mathcal{V}_i \cup \Phi_i$ in round i to be paired with a node of group $G^{\mathcal{U}}_{k,i}$ is proportional to $|s(G^{\mathcal{U}}_{k,i})| \cdot k / \Omega(n \cdot \log\log n)$.

On the other hand, at least three quarters of all stubs of each node $w \in \mathcal{U}(i)$ remain free, i.e., are not paired with a node $w' \in \Phi_i \cup \mathcal{V}_i$, in round i with probability at least

$$1 - \binom{|s(\Phi_i \cup \mathcal{V}_i)|}{\deg_w/4}\left(\frac{\deg_w}{\Omega(n \cdot \log\log n)}\right)^{\deg_w/4} > 1 - \frac{1}{\Theta(\log^{\log\log\log n} n)}. \quad (3)$$

Standard martingale techniques imply that at most a $o(1)$-fraction of all nodes $v \in G^{\mathcal{U}}_{k,i}$ does not satisfy $|s_f(v)| \ge 3/4 \cdot \deg_v$ with high probability. Equation 3 implies that, as long as $|\mathcal{V}_i| \neq \emptyset$, $\Phi_{i+d} \ge \mathcal{V}_i \cdot 3/4 \log\log n$ for all groups $G^{\mathcal{U}}_{k,i}$.

It remains to show that $\mathcal{V}_{i+1} \ge c \cdot \mathcal{V}_i$, for some constant c. To proof this we consider $G^{\mathcal{U}}_{\delta,i}$ only. The probability for a node $w \in G^{\mathcal{U}}_{\delta,i}$ to be paired with at least two different nodes of $\mathcal{V}_i$ is at most

$$\binom{\mathcal{O}(|\mathcal{V}_i| \cdot \gamma(\mathcal{V}))}{2}\left(\frac{\deg_w}{\Theta(n \log\log n)}\right)^2 \le \frac{1}{\Omega\left((\log\log n)^2\right)}. \quad (4)$$

Note, at least a constant fraction of all nodes is contained in $G^{\mathcal{U}}_{\delta,i}$. Further note that Part 1 ends with round j where $|s(\Phi_{j+1})| = \Theta(n)$. Since $|s(G^{\mathcal{U}}_{k,i})| \cdot k / \Omega(n \cdot \log\log n)$ for all considered rounds i, $|\mathcal{V}_j| \le n/\delta$. Thus $|G^{\mathcal{U}}_{\delta,j}| \ge |G_{\delta,j}| \cdot (1-o(1))$. Therefore, with $\gamma(\mathcal{V})$ being chosen large enough (although constant), standard Chernoff bounds imply our claim. Note that at least a constant fraction of all pairing attempts will be made with nodes of group $G^{\mathcal{U}}_{\delta,i}$.

Since $|s(\Phi_i \cup \mathcal{V}_i)| = \mathcal{O}(n)$ implies $|\mathcal{V}_i| \le n/\delta$, the arguments thus far imply $|s(\Phi_{j+1})| = \Theta(n)$ and $|s(\mathcal{V}_{j+1})| = \mathcal{O}(n/\delta)$ after $j = \mathcal{O}(\log n - \log\log\log n)$ rounds.

Part 2 It remains to show that Φ notifies $n - o(n)$ many nodes in step $j+1$, i.e., $|\Phi_{j+2}| = n - o(n)$. First note that the previous arguments imply $|s(\Phi_{j+1})| \geq cn$, for some constant c, within the first $j = \mathcal{O}(\log n - \log\log\log n)$ rounds.

Since $|s(\mathcal{V}_{j+1})| = \mathcal{O}(n/\delta)$ and, by definition, $\Phi_{j+1} = \mathcal{V}_{j-d+1}$, we obtain $|\mathcal{U}(j+1)| \geq n - o(n)$. It follows that the expected number of pairings for an arbitrary node $v \in \mathcal{U}(j+1)$ with nodes of Φ_{j+1} is lower bounded by

$$cn \cdot \frac{\deg_v}{\Theta(n \cdot \log\log n)} = \frac{c \deg_v}{\Theta(\log\log n)}. \tag{5}$$

For what follows next we need two crucial insights. First, c can be (almost arbitrarily) manipulated by increasing/decreasing j. Secondly, expectation of the number of pairings between a node $v \in \mathcal{U}(j+1)$ and $v' \in \Phi_{j+1}$ is strictly greater than 1 with a proper choice of c (cf. Equation 5). To obtain an appropriate bound for the number of distinct nodes paired with the stubs of $s(\Phi_{j+1})$, we model the pairing attempts of Φ as a vertex exposure martingale.

Let the nodes $v_i \in \mathcal{U}(j+1)$ be ordered ascending to the nodes degree, i.e., $\deg_{v_i} \leq \deg_{v_{i+1}}$ for $1 \leq i \leq t'$ whereas $t' = |\mathcal{U}(j+1)|$. Further let X_i, $1 \leq i \leq t'$, be the random variable denoting how many stubs of $s(\Phi_{j+1})$ are paired with node $v_i \in \mathcal{U}(j+1)$. Note, these random variables are not independent. For example, the first i' nodes may be paired with all stubs of $s(\Phi_{j+1})$ and hence leave no stubs left for the remaining $t' - i'$ nodes. However, this does not concern us since the bound we use does not require independence of the random variables. Finally, let f be a real value function with the properties $f(X_1, ..., X_i) = \sum_{1 \leq i' \leq i} X_{i'}$ and $f = \sum_{1 \leq i' \leq t'} X_{i'}$. Certainly such a function exists. Further it must hold $|\mathbb{E}[f \mid X_1, ..., X_i] - \mathbb{E}[f \mid X_1, ..., X_{i-1}]| \leq c_k \leq \deg_{v_i}$.

In other words, the nodes in $\mathcal{U}(j+1)$ are exposed one by one revealing their pairings with the nodes in Φ_{j+1}. It follows by Equation 5 that with a proper choice of c, the expectation of the overall number of different choices is at least $c''n$. Here, c'' is a constant arbitrarily close to 1 depending on the value of c. Consequently we obtain the claim by using the so called *Method of Bounded Martingale Differences* [8]:

$$Pr\left[f \leq \mathbb{E}[f] - (1-c'')n\right] \leq exp\left(-\frac{(1-c'')^2 n^2}{2\sum_{k=1}^{t'} c_k^2}\right) \overset{(Eq.6)}{\leq} exp\left(-n^{\Omega(1)}\right),$$

with

$$\begin{aligned} 2\sum_{k=1}^{t'} c_k^2 &< \mathcal{O}(1) \cdot \sum_{k=2+\log\log n}^{n^{1/\alpha}+\log\log n} \frac{n}{(k - \log\log n)^{\alpha}} (k)^2 \\ &\leq \mathcal{O}(n) \cdot \sum_{k=1}^{n^{1/\alpha}} \frac{(k + \log\log n)^2}{k^{\alpha}} \\ &\leq \mathcal{O}(n) \cdot \frac{2n^{2/\alpha}}{3-\alpha} \frac{1}{n^{(\alpha-1)/\alpha}} \\ &\leq \mathcal{O}(n^{1+3-\alpha/\alpha}). \end{aligned} \tag{6}$$

□

Model 2 As we have seen in the section before, in the first model a $o(n)$ number of individuals become infected, if the average degree in the network is increasing in n. In the second model, $\mathcal{V}$ competes with Λ instead of Φ. Recall, Λ is triggered by $\mathcal{V}$ on every infected node v after a specific delay d. Then these nodes establish a notification tree each (cf. Section 2). Therefore the dissemination of Λ may be much faster than the one of Φ. On the other hand, the minimum degree δ a constant in this model. Since the nodes of the network G follow a power law degree distribution this implies that only $o(n)$ many nodes will have a degree of at least $\log\log n$, which was the minimum degree in Model 1. The question is whether Λ is able to achieve an embankment of $\mathcal{V}$ regardless these circumstances.

The main statement of this section is as follows:

Theorem 2. *Let G be a network with minimum degree $\delta = \Theta(1)$. Further let Model 2 be in place. Then there is a $\gamma(\mathcal{V}) = \Theta(1)$ and delay $d = \Theta(1)$, such that Λ can immunize at least $n - o(n)$ many nodes within $\mathcal{O}(\log\log n)$ many rounds with probability $1 - o(1)$.*

We prove the statement in several stages. In the first stage, i.e. the first $\mathcal{O}(\log\log n)$ rounds, we show that Λ grows by a double exponential rate. This implies that Λ immunizes $n/\log^{\mathcal{O}(1)} n$ many nodes during the first stage with high probability. Note, the growth of $\mathcal{V}$ on the other hand is naturally limited by $\gamma(\mathcal{V})$. Hence $\mathcal{V}$ cannot infect more than a polylogarithmic number of nodes overall during the first stage. These statements then enable us to perform a more in depth analysis specifically focusing on the notification trees. Recall, each notification tree established by a node v has depth at most $\log(\deg_v)$. Clearly, Λ benefits the most from such nodes if they have a large degree (at least polylogarithmic). Hence, these notification trees $\Upsilon_{v,i}$ could potentially immunize other nodes throughout the whole timespan of Theorem 2. However, due to our model, there might be conflicts between notification trees with a large root degree Λ^+ and the ones with a low one Λ^-. We examine these effects during stage two of our analysis for each $G^{\mathcal{U}}_{k,i}$ separately. We conclude by obtaining a significant domination of Λ^+ over Λ^-.

To increase readability, we first prove some auxiliary results. By definition only nodes in Λ, i.e., immunized or recently activated ones, or $\mathcal{V}$, i.e., infective ones, may initiate transmissions trying to immunize or infect. Further, recall that we make use of the Principle of Deferred Decisions and the Configuration Model in the analysis. Consequently, most of the pairings occur either in Λ or $\mathcal{V}$ themselves, or between nodes in $\Lambda/\mathcal{V}$ and $\mathcal{U}$. Note, at this point we omit pairings between nodes in $\mathcal{V}$ and Λ and refer to them when needed. The following definitions simplify the presentation and, if not otherwise stated, hold throughout the entire section.

Definition (ϖ_i). *Let $i > 0$ be an arbitrary round. Further, let $\xi(v)$ be a randomly chosen subset of $s(v)$ of size $\gamma(\mathcal{V})$ and let $\xi'(v) := \{s \in \xi(v) \mid s \in s_f(v)\}$ be the subset of free stubs among the ones of $\xi(v)$. Then*

$$\varpi_i = \sum_{v\in\Lambda_i} |s_f(v)| + \sum_{v\in\mathcal{V}_i} |\xi'(v)|$$

is the amount of stubs of nodes in $\Lambda_i \cup \mathcal{V}_i$ determined to be paired in round i. We say that ϖ_i is constraint on Λ_i and $\mathcal{V}_i$ if $\varpi_i = \sum_{v\in\Lambda_i} |s_f(v)|$ and $\varpi_i = \sum_{v\in\mathcal{V}_i} |\xi'(v)|$, respectively.

Definition (k_{max}). *k_{max} defines the largest node degree with respect to a specific round i, such that the expected amount of newly notified nodes in $G^{\mathcal{U}}_{k_{max},i}$ is at least $\log n$.*

The following lemma basically states that the more stubs may be used for pairing, the more likely it is that nodes with high degree are chosen as pairing partners. This statement is intuitively comprehensive since a higher number of tries implies a higher chance for success.

Lemma 1. *Let $\varpi_i \in [\log^{1+\Omega(1)} n, n^{\frac{\alpha-1}{\alpha}} \cdot \log^{-2} n]$ and let $|s(\mathcal{U}(i))| \geq \Theta(n) - o(n)$. Then*

$$k_{max} = \Theta\left[\left(\frac{\varpi_i}{\log n}\right)^{1/\alpha-1}\right].$$

Proof. By definition it holds $|s(\mathcal{U}(i))| = |s_f(\mathcal{U}(i))| \geq \Theta(n) - o(n)$. Hence the probability for a node $v \in \mathcal{U}(i)$ with degree $\deg_v$ to be paired with at least one stub of ϖ_i is given by

$$1 - \left(1 - \frac{\deg_v}{\Theta(n)}\right)^{\varpi_i} = 1 - e^{-\frac{\deg_v \varpi_i}{\Theta(n)}} = \frac{\varpi_i \deg_v}{\Theta(n)}.$$

Since $|s(\mathcal{U}(i))| \geq \Theta(n) - o(n)$ it follows that $|\mathcal{U}(i)| = \Theta(n)$. Thus

$$\log n = c \cdot \frac{\Theta(|\mathcal{U}(i)|)}{k^{\alpha}_{max}} \cdot \frac{\varpi_i k_{max}}{\Theta(|s(\mathcal{U}(i))|)} \Leftrightarrow k_{max} = \Theta\left[\left(\frac{\varpi_i}{\log n}\right)\right]^{1/\alpha-1},$$

where c is a normalizing constant. □

The next lemma provides crucial insight about the number of rounds needed for process Λ to immunize $n/\log^{\mathcal{O}(1)} n$ many nodes. It states that, up to a specific round j, the uninformed nodes in $\mathcal{U}(j)$ still follow a power law distribution (with a negligible error rate). Thus, for all rounds $i \leq j$, we can treat the network in an uniform way. Using this property we conclude that the growth of $s(\Lambda)$ in Stage 1 is double exponential, which implies that only $\mathcal{O}(\log\log n)$ many rounds are required to immunize $n/\log^{\mathcal{O}(1)} n$ many nodes.

Lemma 2. *Let i^* be the round just before process Λ has informed $\frac{n}{\log^{\mathcal{O}(1)} n}$ many nodes in total. Further, for each $i \leq i^*$, let ϖ_i be constrained on Λ_i. Then for any $i \leq i^*$*

$$\varpi_{i+1} \geq \begin{cases} \Theta\left(\frac{\varpi_i^{\frac{2}{\alpha-1}}}{\log^{\frac{3-\alpha}{\alpha-1}} n}\right), & \text{if } \varpi_i \leq \Theta\left(n^{\alpha-1/2\alpha}\log n\right) \\ \Theta\left(\varpi_i \cdot \frac{n^{\frac{(3-\alpha)(\alpha-1)}{2\alpha^2}}}{\log^{1-(3-\alpha)/\alpha} n}\right), & \text{otherwise} \end{cases}$$

with high probability, where $\delta \leq k \leq n^{1/\alpha}$.

Proof. Let $x_{k,i,j}$ be the event that node $v_j \in G^{\mathcal{U}}_{k,i-1}$ with degree k is informed in round i. Let further $X_{k,i} = \sum_{v_j \in G^{\mathcal{U}}_{k,i-1}} x_{k,i,j}$. Note, Lemma 1 implies that with an increasing ϖ_i also k_{max} will increase. Consequently, nodes in groups with larger degrees will become notified over time and henceforth be part of the dissemination process themselves.

Below we show that ϖ_i increases constantly up to round i^*. The upcoming analysis is two fold. For the first part we restrict the dissemination of Λ to groups with degree at most k_{max}. That is, we only consider groups $G^{\mathcal{U}}_{k,i-1}$ with an expected number of newly informed nodes $\mathbb{E}[X_{k,i}] = \Omega(\log n)$, where the constant hidden in Ω is large. Therefore we simply reset all notified nodes with degree above k_{max}. i.e., these nodes will remain uninformed. Then it is possible to show a double exponential growth of ϖ_i, up to some round i', where it holds $w_{i'} = \Theta\left(n^{\alpha-1/2\alpha}\log n\right)$. On the other hand, the second part of the analysis works as follows. Starting with round $i'+1$, we use a different restriction for Λ. To be more specific, we restrict the dissemination of Λ to groups with degree at most

$$k'_{max} = \left(n^{\frac{\alpha-1}{2\alpha}}\log n\right)^{1/\alpha}.$$

As before, all nodes with larger degree are reset and thus remain uninformed. In addition we allow only a $(1/\log n)$-fraction of all notified nodes to participate in Λ. The other ones are reset as well. In other words, only $\frac{\mathbb{E}[X_{k,i}]}{\log n}$ many nodes remain notified for each $k \leq k'_{max}$ and $i' < i \leq i^*$. Similar as before, we argue that ϖ_i grows by at least a polynomial amount n^{ϵ}, for some $\epsilon < 1$.

Note that both modifications of the dissemination behavior stochastically dominate process Λ due to the following reasons. First, there might be nodes with large degree which are no longer allowed to participate in Λ, i.e., slowing Λ down and thus giving $\mathcal{V}$ the opportunity to remain active. Secondly, as we will see in Lemma 3, the restrictions of this lemma only apply to Λ after a polylogarithmic amount of nodes was infected by $\mathcal{V}$, which implies that $\mathcal{V}$ remained unaffected thus far.

Part 1 First note, all nodes of a group with degree larger than k_{max} may become notified. However, these nodes will be reset anyways. By assumption, there are ϖ_i stubs ready to be paired with other nodes. We are especially interested in the pairings with nodes of $\mathcal{U}$. In short, we focus on the distribution of said pairings in all groups up to $G^{\mathcal{U}}_{k_{max},i'-1}$.

Thus, using an inductive argument, the expected number of newly generated stubs participating in the dissemination process Λ in step $i+1$ is given by

$$\begin{aligned}\mathbb{E}[\varpi_{i+1}] &\geq -\varpi_i + \sum_{k=\delta}^{k_{max}} \frac{n}{k^\alpha}(1-o(1)) \cdot \frac{\varpi_i k}{\Theta(n)} \cdot k \\ &\geq -\varpi_i + \sum_{k=\delta}^{k_{max}} \frac{1}{k^{\alpha-2}} \cdot \Theta(\varpi_i) \\ &\geq \Theta(\varpi_i) \cdot k_{max}^{3-\alpha} \\ &\geq \Theta\left(\frac{\varpi_i^{\frac{2}{\alpha-1}}}{\log^{\frac{3-\alpha}{\alpha-1}} n}\right).\end{aligned}$$

By definition the term ϖ_i in the inequalities represents the stubs in the ith round used for pairing. Consequently, they cannot be used in the next step anymore. The sum in the first inequality contains a $(1-o(1))$ term due to the induction itself. Utilizing standard Chernoff bounds we obtain the same lower bound with high probability. Note, with $\varpi_i = \Theta\left(n^{\alpha-1/2\alpha}\log n\right)$, we obtain $k_{max} = n^{1/(2\alpha)}$ and consequently $\mathbb{E}[X_{n^{1/2\alpha},i}] = \Theta(\log n)$ (cf. Lemma 1 with the above limits for $\deg_v$ and ϖ_i, which imply the validity of the lemma). Thus, ϖ_i grows by a double exponential rate until $\varpi_{i'} = \Theta\left(n^{\alpha-1/2\alpha}\log n\right)$.

Hence, for every group with degree $k \leq k_{max}$ in step $i \leq i'$ there are

$$c' \cdot \left(\frac{n}{k^\alpha} - \frac{\varpi_{i'}}{k^{\alpha-1}}\right) \geq c'' \cdot \left(\frac{n}{k^\alpha}\left(1 - \frac{\varpi_{i'}k_{max}}{n}\right)\right) = \mathcal{O}\left(\frac{n}{k^\alpha}(1-o(1)\right)$$

remaining uninformed nodes, with c', c'' being proper constants. Further with $2 < \alpha < 3$ it holds $\frac{\varpi_{i'}k_{max}}{n} \leq \frac{1}{n^{\Omega(1)}}$ and $|G^{\mathcal{U}}_{k,i'}| \geq |G^{\mathcal{U}}_{k,1}|(1-o(1))$, which concludes the first part.

Part 2 By definition we start in round $i > i'$. Hence it holds $\varpi_i \geq \Theta\left(n^{\alpha-1/2\alpha}\log n\right)$. Consequently, with $k'_{max} = \left(n^{\frac{\alpha-1}{2\alpha}}\log n\right)^{1/\alpha}$, it holds $\mathbb{E}[X_{k,i}] \geq k$ for every $k \leq k'_{max}$ as long as ϖ_i is non-decreasing. Recall, in this part we reset all notified nodes with degree larger than k'_{max}. Further we allow only a $(1/\log n)$-fraction of all notified nodes to participate in Λ. Thus we obtain

$$\begin{aligned}\mathbb{E}[\varpi_{i+1}] &\geq -\varpi_i + \sum_{k=\delta}^{k'_{max}} \frac{n}{k^\alpha}(1-o(1)) \cdot \frac{\mathbb{E}[X_{k,i}]}{\log n} \cdot k \\ &\geq -\varpi_i + \frac{n(1-o(1))}{\log n} \cdot \sum_{k=\delta}^{k'_{max}} \frac{1}{k^{\alpha-2}} \\ &\geq -\varpi_i + \varpi_i \cdot \Theta\left(\frac{n}{\varpi_i \log n} k'^{3-\alpha}_{max}\right) \\ &\geq \Theta\left(\varpi_i \cdot \frac{n^{\frac{(3-\alpha)(\alpha-1)}{2\alpha^2}}}{\log^{1-(3-\alpha)/\alpha} n}\right) \\ &= \Theta\left(\varpi_i \cdot \frac{n^{\Omega(1)}}{\log^{\Theta(1)} n}\right).\end{aligned}$$

Utilizing standard Chernoff bounds we obtain the same lower bound with high probability. Recall, in round i' all groups in $\mathcal{U}$ still followed a power law degree distribution. Further note that all groups up to $G^{\mathcal{U}}_{k'_{max},1}$ contain at least a polynomial amount of nodes. Hence, after removing at most a $(1/\log n)$-fraction of them it still holds $|G^{\mathcal{U}}_{k,i}| \geq |G^{\mathcal{U}}_{k,1}|(1-o(1))$, for all $k \leq k'_{max}$.

Observation 1 *An amount of* $n/\log^c n$ *many nodes can provide at most*

$$n \cdot \sum_{k=\log^{c/\alpha} n} k^{-(\alpha-1)} = \mathcal{O}\left(\frac{n}{\log^{\mathcal{O}(1)} n}\right)$$

many stubs in our model.

By Observation 1 we know that Λ must have immunized at least $\mathcal{O}\left(n/\log^{\mathcal{O}(1)} n\right)$ many nodes if $\varpi_{i^*+1} \geq \mathcal{O}\left(n/\log^{\mathcal{O}(1)} n\right)$. Since $\varpi_{i'} \geq \Theta\left(n^{\Omega(1)}\right)$ and $\varpi_{i+1} \geq \Theta\left(\varpi_i \cdot n^{\Omega(1)}\right)$, for all $i' < i \leq i^*$, it follows that only a constant number of additional rounds are necessary to obtain $\varpi_{i^*+1} \geq \mathcal{O}\left(n/\log^{c'} n\right)$. Note, we can obtain a suitable amount of immunized nodes $n/\log^c n$ for each constant c' by modifying the constant c properly. □

From now on we will mainly focus on notification trees. It is crucial to understand that Λ will contain many competing notification trees during the dissemination process. That is, there will be many different notification trees of various maximum depth competing for nodes in $\mathcal{U}$.

Definition (Type A notification tree). *Notification trees with a maximum depth of* $\mathcal{O}(\log\log n)$.

Definition (Type B notification tree). *All notification trees not of Type A.*

The next lemma copes with this phenomenon by splitting these competing notification trees into two major partitions: notification trees with depth at most $\mathcal{O}(\log\log n)$ (Type A), and all other notification trees (Type B). The intention behind this partitioning is as follows. Type B notification trees will be potentially active, i.e., containing at least one node $v \in \Lambda_i$ for each considered i, over the whole dissemination process since their maximum depth is large enough. On the other hand Type A notification trees may reach their maximum depth way sooner. Consequently conflicts[4] of Type A and Type B notification may result in *lost nodes* for the Type B notification trees.

In other words, whenever an yet uninformed node u is simultaneously notified by a node v of a Type A and another node v' of a Type B notification tree, v is *blocking* the dissemination of v' since, by definition, u will be integrated into vs notification tree. Note, these kind of problems also occur between Type A notification trees. However, these intern conflicts do not provide any problems whatsoever for Type B notification trees. On the contrary, they even help them since potential conflicts between Type A and Type B notification trees are mitigated this way.

Before we can prove the following lemma, we need two more definitions.

Definition (Λ_i^-). *Let j be the round, where the following properties hold for the first time:*

1. $|\mathcal{U}(j)| = n - \mathcal{O}(\log^c n)$, *and*
2. $|\Lambda_j| \geq \Theta(\log^c n)$,

for some constant c.

Then the set

$$\Lambda_i^- := \{v \in \Lambda_i \mid v \in \bigcup_{\forall w \in \Lambda_j : \deg_w \leq \log^q n} \Upsilon_{w,i}\},$$

with q being some proper constant and $i \geq j$, contains all nodes notified by a notification tree (Type A) created in round j with maximum depth $\mathcal{O}(\log\log n)$.

Definition (Λ_i^+). *Let j be the round, where the following properties hold for the first time:*

1. $|\mathcal{U}(j)| = n - \mathcal{O}(\log^c n)$, *and*
2. $|\Lambda_j| \geq \Theta(\log^c n)$,

[4] A conflict of two notification trees occurs whenever one node $u \in \mathcal{U}$ is simultaneously notified by nodes of both notification trees.

for some constant c.

Then the set

$$\Lambda_i^+ := \{v \in \Lambda_i \mid v \in \bigcup_{\forall w \in \Lambda_j : \deg_w > \log^q n} \Upsilon_{w,i}\},$$

with q being some proper constant and $i \geq j$, contains all nodes notified by a notification tree (Type B) created in round j with depth at least $\Omega(\log\log n)$.

Lemma 3. *Let i and j be some fixed rounds, where the following properties hold for the first time:*

$$|\mathcal{U}(i)| = n - \mathcal{O}(\log^c n), |\Lambda_i| \geq \Theta(\log^c n), \text{ and}$$
$$|\mathcal{U}(j)| = n - \mathcal{O}\left(\frac{n}{\log^{c'} n}\right), |\Lambda_j| \geq \Theta\left(\frac{n}{\log^{c'} n}\right)$$

for some properly chosen constants c and c'.

Then it holds

1. $|\Lambda_i^+| - |\Lambda_i^-| = |\Lambda_i^+|(1 - o(1))$,
2. $|\Lambda_j^+| = o(n), |\Lambda_j^-| = o(|\Lambda_j^+|)$,
3. $i = \mathcal{O}(\log\log n), j - i = \mathcal{O}(\log\log n)$.

Proof. To increase readability we split the proof into two main parts, each corresponding to one of the above claims. Our third claim then follows directly from both individual parts.

Property 1 For the first part we slightly alter the dissemination of Λ. Until round i, we simply reset all nodes of Λ as soon as they become activated, thus $\Lambda_{i-1} = \emptyset$. Note that this slightly modified behavior stochastically dominates the original one. However, starting with round i, we use the behavioral modifications of Λ as introduced in Lemma 2. Note, the first nodes contained in Λ_i will be the infected ones in round $i - d$.

According to Lemma 2, we have at least a double exponential growth of Λ till round i and $|s(\mathcal{U}(i))| = \Theta(n)$ since, by assumption, $|\mathcal{U}(i)| = n - \mathcal{O}(\log^c n)$. Further, by definition, $\Lambda_{i'} = \emptyset$ for all $i' < i$. Additionally Lemma 2 implies that $|\mathcal{V}_{i'}| \geq q \cdot |\mathcal{V}_{i'-1}|$, for all $i' < i$ and some constant $q > 1$. Thus $|\mathcal{V}_i| = \Theta(\log^c n)$ after at most $\mathcal{O}(\log\log n)$ many rounds and, by definition, $|\Lambda_i| = \Theta(|\mathcal{V}_{i-d}|) = \Theta(\log^c n)$. Consequently, the amount of stubs overall used by either process is bounded by $|s(\mathcal{V}_i \cup \Lambda_i)| = \mathcal{O}(\log^c n)$ in round i.

With the above arguments in mind, we are now able to calculate the locations of these nodes, i.e., the groups to which the nodes in Λ_i belong. Due to the above modification of Λ until round i, these groups then determine whether a node of Λ_i is the root of a Type A or Type B notification tree.

Note, as long as the amount of nodes in $\Lambda \cup \mathcal{V}$ is sufficiently small, the stubs of $s_f(\Lambda)$ will be paired with nodes of $\mathcal{U}$ with high probability. In other words, almost all of the stubs of Λ will be paired with uninformed nodes without overlapping with each other. Since $|\Lambda_i| = \Theta(\log^c n)$, Lemma 1 states that the groups with degree up to $\Theta(\log^{c-1/\alpha-1} n)$ generate at least a polylogarithmic amount of newly notified nodes with high probability. Further note that at this point the nodes in $\mathcal{U}$ still follow a power law distribution (cf. Lemma 2).

Therefore we obtain

$$|\Lambda_i^-| = \Theta(\log^c n) \sum_{d=\delta}^{\log^q n} 1/d^{\alpha-2} = \Theta\left(\log^{c+q(3-\alpha)} n\right) = \Theta\left(\log^{c_2} n\right) \tag{7}$$

$$|\Lambda_i^+| = \Theta(\log^c n) \sum_{d=\log^q n+1}^{\log^{c-1/\alpha-1} n} 1/d^{\alpha-2} = \Theta\left(\log^{c+\frac{c-1}{\alpha-1}(3-\alpha)} n\right) = \Theta\left(\log^{c_1} n\right) \tag{8}$$

with high probability, where $q \ll c$, $c_1 = c + \frac{c-1}{\alpha-1}(3-\alpha)$, and $c_2 = c + q(3-\alpha)$. Note, group $G^{\mathcal{U}}_{\Theta(\log^{c-1/\alpha-1} n),i}$ contains $n/\log^{\mathcal{O}(1)} n$ many nodes. Since $|\Lambda_i| = \Theta(\log^c n)$, using the Union Bound it follows that each newly notified node of said group is paired with no more than one stub from $s(\Lambda)$ with high probability. Hence, with a proper choice of q and c, we obtain $|\Lambda_i^+| - |\Lambda_i^-| = |\Lambda_i^+|(1 - o(1))$.

Property 2 Before we start to prove the second property, it is important to note that all modifications made to Λ in Lemma 2 are still in place here throughout the corresponding round. Therefore it always holds $|G^{\mathcal{U}}_{k,i'}| \geq |G^{\mathcal{U}}_{k,1}|(1-o(1))$, for all considered k and i'.

Note $|\Lambda^+_i| - |\Lambda^-_i| = |\Lambda^+_i|(1-o(1))$, i.e., Λ^+_i dominates Λ^-_i in round i. Then it follows directly by an inductive argument and Equation 10 that the intersections in all considered groups $G^{\mathcal{U}}_{k,i'}$ between $\Lambda^+_{i'}$ and $\Lambda^-_{i'}$, for each round $i < i' \leq j$, are limited by $o(|\Lambda^+_{i'}|)$, provided $|\Lambda^+_{i'}| - |\Lambda^-_{i'}| = |\Lambda^+_{i'}|(1-o(1))$. The latter property holds true, as we will argue below. Therefore we can analyze $\Lambda^+_{i'}$ and $\Lambda^-_{i'}$ separately.

Due to the initial gap between c_1 and c_2 (cf. Equation 7 and 8), the expected amount of pairings for the last group of $\Lambda^+_{i'}$ we consider is logarithmic wheres only $1/\log^{\mathcal{O}(1)} n$ for $\Lambda^-_{i'}$. We will show that this holds true for the entire span of Equation 10. Obviously, whenever $\Lambda^+_{i'}$ is larger than $\Lambda^-_{i'}$, also the growth rate of $\Lambda^+_{i'}$ will be larger than the one of $\Lambda^-_{i'}$. To upper bound the growth of $\Lambda^-_{i'}$, we set the amount of newly notified nodes by $\Lambda^-_{i'}$ to $\log n$ whenever the expectation of said amount is less than $\log n$. However, since an expectation of newly notified nodes of $1/\log n$ implies that pairings within such groups are very unlikely, we do not consider groups with an expectation below this value. Note that the expectation of newly notified nodes drops below $\log n$ for groups with degree larger than $(\varpi_{i'}/\log n)^{1/(\alpha-1)}$ and is below $(1/\log n)$ for groups with degree larger than $(\varpi_{i'} \log n)^{1/(\alpha-1)}$.

Consequently we can upper bound the growth of $\Lambda^-_{i'}$ by

$$\begin{aligned}
\varpi_{i'+1} &\leq \Theta(\varpi_{i'}) \sum_{k=\delta}^{(\varpi_{i'}/\log n)^{1/\alpha-1}} 1/k^{\alpha-2} + \Theta(\log n) \sum_{k=(\varpi_{i'}/\log n)^{1/\alpha-1}}^{(\varpi_{i'} \log n)^{1/\alpha-1}} 1/k^{\alpha-2} \\
&\leq \Theta(\varpi_{i'}) \left(\varpi_{i'}/\log n\right)^{3-\alpha/\alpha-1} + \Theta(\log n)(\varpi_{i'} \log n)^{3-\alpha/\alpha-1} \\
&= \mathcal{O}\left(\frac{\varpi_{i'}^{2/\alpha-1}}{\log^{3-\alpha/\alpha-1} n} \cdot (1 + 1/\log n)\right),
\end{aligned} \tag{9}$$

whenever $\varpi_{i'}$ is large enough. Recall, we start with an arbitrary polylogarithmic amount. Further it holds $(1+1/\log n)^{\mathcal{O}(\log\log n)} = (1+o(1))$. Therefore, we are able to use the same growth prediction formula for both, $\Lambda^+_{i'}$ and $\Lambda^-_{i'}$ (cf. below). Note, bounding the growth of $\varpi_{i'+1}$ utilizing Chernoff or Martingale techniques adds an additional constant or $1 \pm o(1)$ term to the expectation value. However, since the amount of rounds to be considered is bounded by $\mathcal{O}(\log\log n)$ (cf. below), it follows that the final amount of newly generated nodes by Λ^+_j and Λ^-_j is only affected by a multiplicative value of at most $1/\log^{\mathcal{O}(1)} n$ and $\log^{\mathcal{O}(1)} n$, respectively. As we will show below, this does not affect our claim.

According to Lemma 2,

$$\varpi_{i'+1} \geq \begin{cases} \Theta\left(\dfrac{\varpi_{i'}^{\frac{2}{\alpha-1}}}{\log^{\frac{3-\alpha}{\alpha-1}} n}\right), & \text{if } \varpi_{i'} \leq \Theta\left(n^{\alpha-1/2\alpha} \log n\right) \\ \Theta\left(\varpi_{i'} \cdot \dfrac{n^{\frac{(3-\alpha)(\alpha-1)}{2\alpha^2}}}{\log^{1-(3-\alpha)/\alpha} n}\right), & \text{otherwise} \end{cases}$$

with high probability. We consider these cases separately.

As long as Case 1 holds, we have $\varpi_{i'+1} \geq \Theta\left(\varpi_{i'}^{2/\alpha-1} \log^{-(3-\alpha)/\alpha-1} n\right)$. On the other hand, if $\varpi_{i'} = \Theta\left(n^{\alpha-1/2\alpha} \log n\right)$, then Case 2 holds (cf. Lemma 2). Obviously, since we start with a polylogarithmic amount of stubs $\varpi_1 = \log^{c''} n$ (where the constant c'' is set accordingly to the value of $\Lambda^+_{i'}$ and $\Lambda^-_{i'}$, respectively), we can rewrite the formula for $\varpi_{i'+1}$. It holds $\varpi_{i+j} = \log^x n$, where the exponent x in Case 1 satisfies the following:

$$\begin{aligned}
&\text{step } i+1: \ \frac{2c''}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\\
&\text{step } i+2: \ \left(\frac{2c''}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\right)\frac{2}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\\
&\text{step } i+3: \ \left(\left(\frac{2c''}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\right)\frac{2}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\right)\frac{2}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\\
&\text{step } i+4: \ \left(\frac{2^3c'' - 2^2\cdot 3 + 2^2\alpha}{(\alpha-1)^3} - \frac{2^1\cdot 3 - 2^1\alpha}{(\alpha-1)^2} - \frac{2^0\cdot 3 - 2^0\alpha}{(\alpha-1)^1}\right)\frac{2}{\alpha-1} - \frac{3-\alpha}{\alpha-1}\\
&\qquad = \frac{2^4c'' - 2^3\cdot 3 + 2^3\alpha}{(\alpha-1)^4} - \frac{2^2\cdot 3 - 2^2\alpha}{(\alpha-1)^3} - \frac{2^1\cdot 3 - 2^1\alpha}{(\alpha-1)^2} - \frac{2^0\cdot 3 - 2^0\alpha}{(\alpha-1)^1}\\
&\qquad \vdots\\
&i' \text{ steps done}: \ \frac{2^{i'}c''}{(\alpha-1)^{i'}} - \frac{3-\alpha}{\alpha-1}\sum_{j'=0}^{i'-1}\left(\frac{2}{\alpha-1}\right)^{j'} = (c''-1)\left(\frac{2}{\alpha-1}\right)^{i'} + 1 \qquad (10)
\end{aligned}$$

Recall, we start at step i with $\Theta(\log^{c_1} n)$ many stubs (cf. Equation 8) for Λ_i^+, i.e., for the notification trees which may send throughout at least $\Omega(\log\log n)$ many rounds. Further recall it holds $|\Lambda_i^+| - |\Lambda_i^-| = |\Lambda_i^+|(1-o(1))$, i.e., Λ_i^+ dominates Λ_i^- in round i.

Our calculation for the exponent (Equation (10)) implies that $\log_{\frac{2}{\alpha-1}} \log_{\log^{c_1-1} n} n^{\frac{\alpha-1}{2\alpha}}$ many rounds are sufficient to achieve the point in time in which Case 2 applies. Recall, Case 2 holds when $\varpi_{i'} \geq \Theta\left(n^{\alpha-1/2\alpha}\log n\right)$.

Thus with the right choice of c_1, c_2, i.e., $c_1 \gg c_2$, we obtain

$$\begin{aligned}
|\Lambda_{j'}^+| &= \mathcal{O}\left({}^1/_{\log^{\mathcal{O}(1)} n} \cdot n^{\frac{\alpha-1}{2\alpha}}\right), \text{ and}\\
|\Lambda_{j'}^-| &= \mathcal{O}\left(\log^{\mathcal{O}(1)} n \cdot n^{\frac{\alpha-1}{2\alpha}\cdot\left(\log_{\log^{c_2-1} n}\log^{c_1-1} n\right)^{-1}}\right) = o(|\Lambda_{j'}^+|),
\end{aligned}$$

for some round j'. Additionally this implies that $\Lambda_{i'}^+$ dominates $\Lambda_{i'}^-$ in all rounds $i' \leq j$ as well.

Further, a constant number of additional steps is sufficient to achieve the desired amount of $\frac{n}{\log^{c'} n}$ many notified nodes using the growth estimates of Case 2 (cf. Lemma 2). Similar arguments as above imply that $\Lambda_{i'}^-$ will not grow faster in the remaining rounds than $\Lambda_{i'}^+$. Thus, with a proper choice of c_1, c_2 and c' the lemma follows. □

Definition (ϖ_i^k). *Amount of stubs used by a dissemination process for the pairing process with nodes of $G_{k,i}^{\mathcal{U}}$.*

Let $i > 0$ be an arbitrary round. Then

$$\varpi_i^k = \varpi_i \cap G_{k,i}^{\mathcal{U}}.$$

We say that ϖ_i^k is constraint on Λ_i and $\mathcal{V}_i$ if ϖ_i is constraint on Λ_i and $\mathcal{V}_i$, respectively.

Definition (Section s_1). *Section s_1 contains all groups $G_{k,i}^{\mathcal{U}}$ with an expected amount of newly notified nodes by Λ^- of at least logarithmic size.*

Definition (Section s_2). *Section s_2 contains all groups with a degree larger than the ones of Section s_1 up to the groups with an expected amount of newly notified nodes by Λ^- of at most ${}^1/_{\log^{\Theta(1)} n}$.*

Definition (Section s_3). *Section s_3 contains all groups, up to degree at most k_{max} w.r.t. Λ^+, not contained in Section s_1 nor Section s_2.*

Finally we are ready to prove Theorem 2.

Proof. [Theorem 2]

To show our claim we need to analyze both processes $\mathcal{V}$ and Λ throughout the whole lifetime of $\mathcal{V}$. We prove our claim in two main phases. First, we show a double exponential growth of Λ during the first $\mathcal{O}(\log\log n)$ rounds. This implies that within the first $\mathcal{O}(\log\log n)$ rounds Λ immunizes $n/\log^{\mathcal{O}(1)} n$ many nodes with high probability. The growth of $\mathcal{V}$ on the other hand is naturally limited by $\gamma(\mathcal{V}) = \Theta(1)$. Therefore, $\mathcal{V}$ cannot infect more than a polylogarithmic amount of nodes overall within the first $\mathcal{O}(\log\log n)$ rounds. To complement this statement we show that $\mathcal{V}$ remains active within these first $\mathcal{O}(\log\log n)$ rounds, i.e., $\gamma(\mathcal{V}) \cdot |s(\mathcal{V}_i)| \geq |s(\mathcal{V}_{i+1})| \geq |s(\mathcal{V}_i)| \cdot (1 + \Omega(1))$ for every round i.

From this point forward we mainly concentrate on the behavior of Λ. Recall, each notification tree created by a node $v \in \Lambda$ has depth at most $\log(\deg_v)$. Therefore the highest benefit of such notification trees for Λ is provided by root nodes with at least a polylogarithmic degree. These notification trees $\Upsilon_{v,i}$ could potentially stay active, i.e., $\Lambda_i \cap \Upsilon_{v,i} \neq \emptyset$ for all considered i, throughout the whole timespan of this proof. We cope with this phenomenon by dividing process Λ into two subprocesses Λ^+ and Λ^-. Then we analyze the conflicts between said processes during the remaining dissemination process. To be more specific, we study the impact of Λ^+ and Λ^- in each $G^{\mathcal{U}}_{k,i}$ separately and show that three properties hold:

1. $\frac{|\Lambda_i^-|}{|\Lambda_i^+|} \leq \frac{(\log^{\Theta(1)} n)^i}{n^{\Omega(1)}}$, for every considered i,
2. the amount of newly notified nodes in $G^{\mathcal{U}}_{k,i}$ is $\varpi_i^k \cdot (1 - o(1))$, whenever $\varpi_i^k/|G^{\mathcal{U}}_{k,i}| = o(1)$, and
3. $|G^{\mathcal{U}}_{k,i+1}| \leq \frac{|G^{\mathcal{U}}_{k,i}|}{\Theta(1)}$, whenever $\varpi_i^k/|G^{\mathcal{U}}_{k,i}| \neq o(1)$.

<u>*Phase 1*</u> Let i' be the round where $|\mathcal{V}_{i'}| = \Theta(\log^c n)$ for the first time, for some constant c. Note, in the following we use the same modification as in Lemma 2 for Λ. Since P_V spreads over a constant number of connections at each infected node, it follows that within $\mathcal{O}(\log\log n)$ steps the disease cannot infect more than $\mathcal{O}(\log^c n)$ nodes. Further the lowest and the average degree in the considered network is constant. Recall, $\Lambda_i = \emptyset$ for all rounds $i < i'$, which in turn implies $|s(\mathcal{V}_{i'} \cup \Lambda_{i'})| = \Omega(\log^c n)$. Since $\gamma(\mathcal{V}) = \Theta(1)$ it follows $|s(\mathcal{V}_{i'} \cup \Lambda_{i'})| = \Theta(\log^c n)$. Therefore $\mathcal{V}$ will only infect uninformed nodes until round i' with probability at least

$$\left(1 - \frac{\log^{\mathcal{O}(1)} n}{\Theta(n)}\right)^{\log^{\mathcal{O}(1)} n \cdot \mathcal{O}(\log\log n)} \leq 1 - \frac{\log^{\mathcal{O}(1)} n}{\Theta(n)}.$$

Additionally, each node $v \in \mathcal{U}$ is notified at least two times by nodes in $\mathcal{V}$ with probability at most

$$\binom{|s(\mathcal{V}_{i'})|}{2}\left(\frac{\deg_v}{\Theta(n)}\right)^2 \leq \left(\frac{\Theta(\log^c n)}{n^{1-1/\alpha}}\right)^2 = o(n^{-1/2}),$$

since $\deg_v \leq n^{1/\alpha}$ for all $v \in V$. Utilizing the union bound, we conclude that $\mathcal{V}$ will remain active until round i' with high probability. At this point we apply Lemma 2 and Lemma 3 to obtain the rest of our claim for this part.

<u>*Phase 2*</u> From the previous phase we know that Λ managed to immunize $n/\log^{\mathcal{O}(1)} n$ nodes within the first $\mathcal{O}(\log\log n)$ rounds. To analyze the further behavior of process Λ we consider the situation in each $G^{\mathcal{U}}_{k,i}$ separately from now on.

The following case analysis divides all groups $G^{\mathcal{U}}_{k,i}$ into two major categories:

– Groups with the property $\varpi_i^k/|G^{\mathcal{U}}_{k,i}| = o(1)$ (Case 1), and
– all groups without said property (Case 2).

Note, since we use the same behavioral modification of Λ as we did in Lemma 2, the node degrees in $\mathcal{U}$ are still distributed accordingly to the power law distribution (up to a negligible factor) at the beginning of this phase. During the rest of this analysis this is not necessarily the case anymore. However, for all nodes $v \in \mathcal{U}(i)$ it holds $|s_f(v)| = \deg_v$. Thus each node $v \in \mathcal{U}$ will be notified with probability proportional to $|s_f(v)|$. We prove the rest of our claim by induction. To be more specific we explicitly show that

$$\frac{|\Lambda_i^-|}{|\Lambda_i^+|} \leq \frac{(\log^{\Theta(1)} n)^i}{n^{\Omega(1)}} \tag{11}$$

holds for each considered round i.

Case 1 In this case it holds $\varpi_i^k/|G_{k,i}^{\mathcal{U}}| = o(1)$ for every considered group k. Since we start with $|\Lambda_i| = \Theta\left(n/\log^c n\right)$, where c is a constant, we apply Lemma 3 and obtain the domination of Λ_i^+ over Λ_i^- for this round. Note, we only consider groups up to k_{max} in this case. Therefore it is not possible for Λ^+ to lose its domination over Λ^- for round $i+1$ due to the following reasons:

- Whenever the expected amount of newly notified nodes for both Λ^+ and Λ^- is at least logarithmic, said expected amount is tight (up to some constant factor).
- Due to the initial domination of Λ^+ over Λ^- it follows that the amount of newly notified nodes by Λ^+ in $G_{k_{max},i'}^{\mathcal{U}}$ is at least logarithmic whereas $o(\log^{-\Theta(1)} n)$ for Λ^-.

Hence Λ_i^- may gain at most a polylogarithmic factor over $\mathcal{O}(\log\log n)$ many steps.

Next we observe that the probability for a node in $G_{k,i}^{\mathcal{U}}$ to be contacted at least twice (either from one or both sets Λ_i^+ and Λ_i^-) is less than

$$\binom{w_i^k}{2}\left(\frac{k}{|G_{k,i}^{\mathcal{U}}|\cdot k}\right)^2 \leq \Theta(1)\left(\frac{w_i^k}{|G_{k,i}^{\mathcal{U}}|}\right)^2 = o(1). \tag{12}$$

This implies that the amount of immunized nodes will grow rapidly. Additionally, the domination of Λ_i^+ over Λ_i^-, Equation 12, and $\varpi_i^k/|G_{k,i}^{\mathcal{U}}| = o(1)$ imply

$$\frac{|\Lambda_{i+1}^-|}{|\Lambda_{i+1}^+|} \leq \frac{|\Lambda_i^-|}{|\Lambda_i^+|\cdot(1-o(1))} < \frac{(\log^2 n)^{i+1}}{n^{\Omega(1)}}.$$

Recall, we only consider groups with degree up to k_{max}. For all groups which have larger degree, the expected amount of pairings to nodes in Λ^- is less than $1/\log^{\Theta(1)} n$ (cf. Lemma 1 and Case 2). Furthermore, we have $\Theta(1) \leq k_{max} \leq n^{1/\alpha}$, otherwise Case 1 would not apply. Thus, standard Chernoff bounds imply the same results, up to constant factors, with high probability.

Case 2 At some point Case 1 does not hold for each group anymore. Then, Case 2 applies for these groups. Before we start the formal part, we first give a short outline for the rest of the proof.

At the beginning it holds $\varpi_i \leq o(n)$, for all rounds i. Thus, the Sections s_1, s_2 and s_3 are non-empty and well defined. During the procedure ϖ_i increases and the following problem arises: Section s_3 followed by Section s_2 may disappear. That is, at this point none of the remaining groups $G_{k,i}^{\mathcal{U}}$ are contained in said sections anymore. The cause is the growth of ϖ_i itself.

An increasing ϖ_i implies an increasing expected amount of newly informed nodes in all groups $G_{k,i}^{\mathcal{U}}$. With ϖ_i large enough, there is simply no possibility for an arbitrary group to have an expected amount of newly notified nodes of no more than $1/\log^{\Theta(1)} n$. Thus, there may be a point in time where only Section s_1 remains. Therefore an increasing ϖ_i shifts the transition points between these sections. However, we will show that throughout all steps Λ_i^+ will not lose its domination over Λ_i^-.

According to Lemma 3 our claim holds for some round i' with $|\Lambda_{i'}| = \Theta\left(n/\log^{\mathcal{O}(1)} n\right)$. Recall, the expected amount of newly notified nodes in group $G_{k,i}^{\mathcal{U}}$, for all rounds i, is proportional to its amount of free stubs at any given point in time. This holds regardless of the degree distribution in $\mathcal{U}$. Now let us assume that our claim holds for round $j \geq i'$ and let the groups be sorted in ascending order according to their degree. The following subcases cover the corresponding sections, i.e., a subcase applies for all groups contained in the corresponding section.

Subcase s_1 Let X_j^+, X_j^- be the number of newly notified nodes by Λ_j^+ and Λ_j^-, respectively. Note, the expected amount of newly notified nodes by Λ_j^- in the group with the highest degree of Section s_1 is of logarithmic size whereas $\Omega\left(n^{\Omega(1)}/\log^{\Theta(j)-1} n\right)^5$ for Λ_j^+. Thus, for all groups in Section s_1, the lower and upper bounds for the expectation of $\mathbb{E}[X_j^+], \mathbb{E}[X_j^-]$ are tight up to (at most) a constant factor. Recall, by assumption it holds $|\Lambda_j^-| < \log^{\Theta(j)} n/n^{\Omega(1)} \cdot |\Lambda_j^+|$ at the beginning of step j.

[5] Due to our assumption $|\Lambda_j^+| \geq |\Lambda_j^-| \cdot n^{\Omega(1)}/\log^{\Theta(j)} n$.

Thus we obtain

$$\frac{|\Lambda_{j+1}^{+}|}{|\Lambda_{j+1}^{-}|} \geq \frac{c_1 \cdot \mathbb{E}[X_j^{+}]}{c_2 \cdot \mathbb{E}[X_j^{-}]} = \Theta(1) \cdot \frac{\Lambda_j^{+} \cdot p}{\Lambda_j^{-} \cdot p} \geq \Theta(1) \cdot \frac{n^{\Omega(1)}}{\log^{\Theta(j)} n}$$

with high probability, for some constants c_1, c_2 and the probability to notify an yet uninformed node p.

Furthermore, the intersection of notified nodes by Λ_j^{+} and Λ_j^{-} is still limited by $o(|\Lambda_j^{+}|)$, even if we assume that all notified nodes by Λ_j^{-} intersect with the ones notified by Λ_j^{+}.

Let v be an arbitrary node of $G_{k,j}^{\mathcal{U}}$ and let ϖ_j^k be constraint on Λ. Since we are in Case 2, $\varpi_j^k \geq \Omega\left(|G_{k,j}^{\mathcal{U}}|\right)$ holds by assumption. Thus the probability for v to be notified is at least

$$1 - \left(1 - \frac{k}{k|G_{k,j}^{\mathcal{U}}|}\right)^{\varpi_j^k} \geq 1 - e^{-\Omega(1)}.$$

It follows that at least a constant fraction of the nodes in $G_{k,j}^{\mathcal{U}}$ will be immunized with high probability.

Subcase s_3 By definition we obtain for the first group of Section s_3: the expected amount of newly notified nodes by Λ^{-} is less than $1/\log^{\Theta(1)} n$ whilst $\Omega\left(n^{\Omega(1)}/\log^{\Theta(j)} n\right)^5$ for Λ^{+}. Note, in this section Λ^{+} increases while Λ^{-} does not. This implies our claim. Consequently, intersections between nodes notified by Λ^{-} and Λ^{+} do not exist, with high probability. Further, by the same arguments as in Subcase s_1, it follows that at least a constant fraction of the nodes in $G_{k,j}^{\mathcal{U}}$ will be immunized with high probability.

Subcase s_2 Since the expected amount of newly notified nodes by Λ^{-} is less than logarithmic, this subcase is the most problematic one. However, we can bound the amount of newly notified nodes by Λ^{-} to $\mathcal{O}(\log n)$ in each group of Section s_2. Note, this bound holds with high probability using standard Chernoff bounds, for example. By definition we obtain that the number of newly notified nodes by Λ^{+} is $\Omega\left(n^{\Omega(1)}/\log^{\Theta(j)-1} n\right)$ for the first group of Section s_2 and $\Omega\left(\frac{n^{\Omega(1)}}{\log^{\Theta(j)} n}\right)$ for the first group of Section s_3. Similarly, the number of newly notified nodes by Λ^{-} is $\log n$ for the first group of Section s_2 and $1/\log^{\Theta(1)} n$ for the first group of Section s_3.

Hence it must hold

$$\frac{\log n \cdot \log^{\Theta(j)-1} n}{\Omega\left(n^{\Omega(1)}\right)} = \frac{\log^{\Theta(j)} n}{\Omega\left(n^{\Omega(1)}\right)} \leq \frac{|\Lambda^{-}|}{|\Lambda^{+}|} \leq \frac{\log^{\Theta(j)} n \cdot \log n}{\Omega\left(n^{\Omega(1)}\right)} = \frac{\log^{\Theta(j)+1} n}{\Omega\left(n^{\Omega(1)}\right)},$$

for all groups in Section s_2. Further, by the same arguments as in Subcase s_1, it follows that at least a constant fraction of the nodes in $G_{k,j}^{\mathcal{U}}$ will be immunized with high probability.

All arguments so far imply Equation 11. Further, the same arguments imply that within $\mathcal{O}(\log\log n)$ rounds process Λ will notify at least $n - o(n)$ many nodes with probability $1 - o(1)$. □

4 Conclusion

We presented two models, which describe a simple behaviour pattern of individuals in random power law graphs. Two different parameter settings were used for the analysis. In the first case we assumed that the minimum degree in the network is unbounded, and a malicious piece of information (also called infection) is spread nearly unhindered while a preventive message may only be sent to the neighbours of infected nodes. Our result implies that w.r.t. this model an infection/misinformation of persons can be stopped by establishing routines and subnetworks which guarantee a specific amount of connections for each node.

In the second case we assumed that the smallest degree in the network is bounded by some large constant. We observed that the spread of a malicious piece of information can efficiently be stopped if warnings are propagated for a sufficient number of hops, depending on the degree of the source of the preventive message. One can conclude that although there are only a few high degree nodes - compared to the size of the network - these have a huge impact on the behaviour of the processes we considered.

References

[1] Adamic, L.A., Huberman, B.A.: Power-law distribution of the world wide web. Science 287(5461), 2115 (2000)

[2] Ajelli, M., Goncalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J., Merler, S., Vespignani, A.: Comparing large-scale computational approaches to epidemic modeling: Agent-based versus structured metapopulation models. BMC Infectious Diseases 10(190) (2010)

[3] Amaral, L.A., Scala, A., Barthelemy, M., Stanley, H.E.: Classes of small-world networks. PNAS 97(21), 11149–11152 (2000)

[4] Balcan, D., Hu, H., Goncalves, B., Bajardi, P., Poletto, C., Ramasco, J.J., Paolotti, D., Perra, N., Tizzoni, M., den Broeck, W.V., Colizza, V., Vespignani, A.: Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Medicine 7, 45 (2009)

[5] Bobashev, G.V., Goedecke, D.M., Yu, F., Epstein, J.M.: A hybrid epidemic model: combining the advantages of agent-based and equation-based approaches. In: Proc. WSC '07. pp. 1532–1537. IEEE Press, Piscataway, NJ, USA (2007)

[6] Borgs, C., Chayes, J., Ganesh, A., Saberi, A.: How to distribute antidote to control epidemics. Random Struct. Algorithms 37, 204–222 (September 2010)

[7] Callaway, D.S., Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Network robustness and fragility: Percolation on random graphs. Physical Review Letters 85(25), 5468–5471 (2000)

[8] Dubhashi, D., Panconesi, A.: Concentration of Measure for the Analysis of Randomized Algorithms. Cambridge University Press, New York, NY, USA, 1st edn. (2009)

[9] Eubank, S., Guclu, H., Kumar, V., Marathe, M., Srinivasan, A., Toroczkai, Z., Wang, N.: Modelling disease outbreaks in realistic urban social networks. Nature 429(6988), 180–184 (2004)

[10] Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. In: SIGCOMM '99. pp. 251–262 (1999)

[11] Funk, S., Gilad, E., Watkins, C., Jansen, V.A.A.: The spread of awareness and its impact on epidemic outbreaks. PNAS 106(16), 6872–6877 (2009)

[12] Germann, T.C., Kadau, K., Longini, I.M., Macken, C.A.: Mitigation strategies for pandemic influenza in the United States. PNAS 103(15), 5935–5940 (2006)

[13] Gladwell, M.: The Tipping Point: How Little Things Can Make a Big Difference. London: Academic Press (2002)

[14] Grassberger, P.: On the critical behavior of the general epidemic process and dynamical percolation. Mathematical Biosciences 63(2), 157 – 172 (1983)

[15] Hethcote, H.W.: The mathematics of infectious diseases. SIAM Review 42(4), 599–653 (2000)

[16] Jaffry, S.W., Treur, J.: Agent-Based and Population-Based Simulation: A Comparative Case Study for Epidemics. In: Louca, L.S., Chrysanthou, Y., Oplatkova, Z., Al-Begain, K. (eds.) European Conference on Modelling and Simulation, ECMS'08. pp. 123–130 (2008)

[17] Lee, B.Y., Bedford, V.L., Roberts, M.S., Carley, K.M.: Virtual epidemic in a virtual city: simulating the spread of influenza in a us metropolitan area. Translational Research 151(6), 275 – 287 (2008)

[18] Lee, B.Y., Brown, S.T., Cooley, P.C., Zimmerman, R.K., Wheaton, W.D., Zimmer, S.M., Grefenstette, J.J., Assi, T.M., Furphy, T.J., Wagener, D.K., Burke, D.S.: A computer simulation of employee vaccination to mitigate an influenza epidemic. American Journal of Preventive Medicine 38(3), 247 – 257 (2010)

[19] Liu, T., Li, X., Ai, B., Fu, J., Zhang, X.: Multi-agent simulation of epidemic spatio-temporal transmission. In: Fourth International Conference on Natural Computation, ICNC '08. vol. 7, pp. 357–361 (2008)

[20] McKay, B., Wormald, N.: Uniform generation of random regular graphs of moderate degree. Journal of Algorithms 11, 52–67 (1990)

[21] Moreno, Y., Vázquez, A.: Disease spreading in structured scale-free networks. European Physical Journal B 31(2), 265–271 (2003)

[22] Newman, M.E.J.: Spread of epidemic disease on networks. Phys. Rev. E 66(1), 016128 (2002)

[23] Newman, M.E.J.: The structure and function of complex networks. SIAM Review 45(2), 167–256 (2003)

[24] Ripeanu, M., Foster, I., Iamnitchi, A.: Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system. IEEE Internet Computing Journal 6(1), 50–57 (2002)

[25] Valler, N., Prakash, B.A., Tong, H., Faloutsos, M., Faloutsos, C.: Epidemic spread in mobile ad hoc networks: Determining the tipping point. pp. 266–280. IFIP NETWORKING, Valencia (2011)