Virus diffusion through mobile phones
Complex Systems Project

Ivan Heibi

University of Bologna, Bologna, Italy,
ivan.heibi@studio.unibo.it

Abstract. The purpose of this project is to study the spreading patterns of mobile phone viruses. In order to do that we will observe the mobile phone users and try to understand the virus propagation through two mechanism: Bluetooth and MMS. For MMS we will introduce two different strategies of propagation, the first and most common strategy will have a Topological behavior, while the second one locates a vulnerable phone by generating a random number to be contacted, this strategy is called Scanning. In addition we will define the *User-Awareness* a new factor to try portray a more heedful behavior from the users, and see the consequences it brings down to the virus propagation for both Bluetooth and MMS.

1 Introduction: Mobile Phone Viruses

The development of smart phones, capable of sharing programs and data with each other, turned mobile phones into very vulnerable nodes for viruses . Indeed, since 2004 hundreds of smart phone viruses have been identified, having reached a state of sophistication in a few years that took computer viruses about two decades to achieve, this phenomena tells a lot about the high spreading speed of viruses in relation to what we used to see on computer viruses, which makes the security demand more necessary. A very important aspect to keep an eye on it is the mobile operating system's *Market-Share*, which indicates that a cell phone virus can operate only the phones with the operating system it was designed for. There are two main communication methods to transmit virus through mobile phones :

Viruses carried by multimedia messages (MMS):
MMS viruses follow a long-range spreading pattern which is independent from the infected phones physical location, through this type of mechanism an infected phone can directly reach the phones numbers of his phone book, this type of communication is *Topological* because it uses a call graph topology to find vulnerable phones. Another strategy for a MMS virus propagation is called *Scanning*, this method generates randomly the phone numbers of it's victims, so the infection may move from one side of the network to another in a very short amount of time, it's also important to notice that the phone numbers generated can be ineffective.

Bluetooth (BT):
Virus can infect all phones found within BT range (phones within a distance from 10 to 30m) from the infected phone. As physical proximity is essential for a BT connection, the transmission of a BT virus is determined by the owners location and mobility patterns. Therefor BT viruses follow a spreading pattern similar to the influenza diffusion.

2 Model description

To build the basic model environment we followed some previous studies [1]. So the background will include a set of users that use cell phones and a number of initial members infected with a virus. The model will explore the propagation of a singular virus (all viruses in the system belong to the same type), although the virus will have the possibility to evolve over time into a new and better version. The virus will operate only on phones with an OS which it is build for, so the Market-Share value (m) for the set of users in the system will be also assigned to test it's effect. As already mentioned before, the two main mechanism for mobile virus spreading are MMS and BT, so our model will still focus on these two, trying to rebuild a possible scenario and study relevant interesting phenomenons. To easily understand the model we will look at each spreading mechanism individually in order to explain the way it operates and what kind of

factors can characterize it's actions. Besides that some new factors were added to test and analyze new interesting aspects, the main one is an attempt to characterize the **User-Awareness(UA)** and it's consequences to the model, UA will gain different meanings and interpretations depending on the virus diffusion mode.

2.1 User-Awareness (UA)

Users are increasingly placed in a position where they must handle information security matters that they did not handle in days gone past. These new distributed systems and smart phones force users to play security roles that they had not previously had to play. User awareness incorporates advising on the responsibility for maintaining security through good security practices, and through increased awareness of the issues and risks that may affect them individually. This aspect can be improved through a school education and by highlighting it's importance.

2.2 Bluetooth

We will assume that Bluetooth port is always on, the infection attempt starts when an infected node is close enough to a susceptible node, all the nodes move randomly and in free directions. Infected nodes will have the opportunity to recover, the recovery proceeding will turn infected nodes to susceptible status again, so they still at risk for a possible future infection, so in some ways a recovery is a temporal effect, it can represent the detection and removal of the virus, this will not give the node a full resistance to the virus because we are taking in consideration the fact that viruses can evolve and overthrow past weaknesses, users will try to recover with a given time frequency. The infection process will take place only if the attacked node accepts the send request from the infected one, this act symbolize the UA level, so users more conscious of these type of risks will behave more responsibly and take in consideration the possible negative repercussions, as a result a more observant community will be generated. In Bluetooth virus propagation a user with a high UA will not always accept a sending request from the infected user. We consider this model as a kind of SIS-Model, such that it allowed an infection and the future recovery of the node, so it switches his status from **S**usceptible to **I**nfected and to **S**usceptible again with the recovery process . Our model will also check the consequences of the Market-Share using this propagation type.

2.3 MMS

For MMS virus propagation the nodes mobility are irrelevant, therefore nodes positions could be randomly generated and will not have any significant meaning. We divided the diffusion of virus through this mechanism in two different strategies as described in [2].

– Topological : the virus spreads by targeting individuals in the address books of the infected phones. The address book length can be changed. This method ensure that almost all the targets are active and real phones, although they can appertain to a different OS.
– Scanning : Randomly selecting the victims by randomly generating phone numbers to be attacked. Since previous works normally use a topological approach, ignoring the possibility that a virus can scan random phone numbers this method will introduce new interesting results. Generating random numbers can also aim ineffective users, and we still must consider the fact that some users can own a different OS.

These strategies can have different consequences depending on the Market-Share value, as [2] shows that the topological and scanning behaviors of MMS viruses can be more damaging in high and low market share cases, respectively. For our model we will consider the chance of virus diffusion throw MMS as one parameter that amass all those different aspects such the Market-Share, it's value will approximate the virus propagation strength depending on the situation taken in consideration.
Each infected node can perform a maximum number of attacks, this upper-bound will help the non detection of the virus: sending a big number of mms from same node makes it more suspicious [2]. If an infected node is conscious that a virus attacked him through a specific mms message, he can alerts all the nodes in his address book, this act will make them resistant to the virus, this operation represents the UA. Each infected node needs at least a day to auto detect himself, and the chance of detection is related to the UA value. Since noticed nodes can gain an immunity to the virus this should be considered as a SIR-Model (Susceptible-Infected-Recovered).

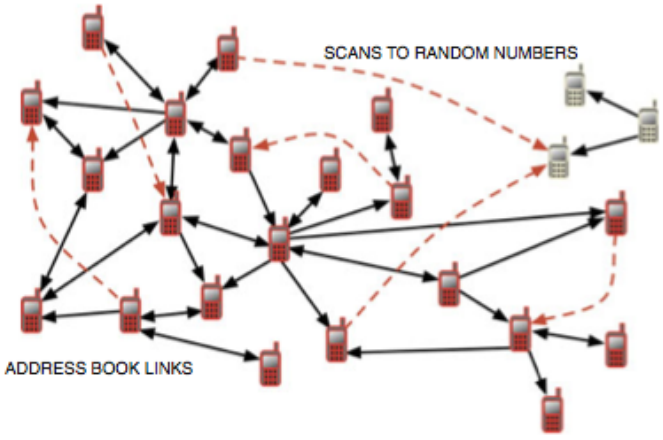

Fig. 1. Scanning and Topological behavior. Source [2].

3 NetLogo Model

NetLogo is an agent-based programming language it's a good environment for the exploration of complex systems, which makes it really useful for our study. The NetLogo agents will represent the mobile phones which can appear in green, red or grey color, making them in susceptible, infected or resistant status respectively. The total number of nodes and the initial number of infected phones can be initially chosen through two sliders . After the initial data are initialized a method of virus diffusion must be chosen (MMS or BT) throw a chooser input, as a result of this decision all the relative parameters to the selected mechanism. Each kind of method request different type of parameters. The last general parameter value to set with a slider is the UA percentage (it will gain different meanings for the diffusion method selected). For Bluetooth NetLogo ensure the possibility to decide the Market-Share and the recovery chance, the recovery process can be carried in a specific time frequency.
On the other hand for MMS we need to decide first which type of diffusion we would like to use (Topological or Scanning) with a chooser input, and the virus expansion probability, the virus diffusion will be made after each tick, which can represent the 2min needed for a typical MMS virus to copy itself on a new handset [1] .In addition we will set the maximum number of attacks the virus can accomplish and the address book size for all the phones throw two input boxes. Figure 2 summarize all the parameters introduced in NetLogo.

nodes-number	Integer
initial-infected-nodes	0 – nodes-number
virus-diffusion	MMS / BT
user-awareness	0 – 100%

(a)

market-share	0 – 100%
bt-check-frequency	Integers
recovery-chance	0 – 10%

(b)

mms-type	Topologcal / Scanning
mail-list-size	0 – nodes-number
max-attacks	Integer
mms-spread-chance	0 – 1%

(c)

Fig. 2. NetLogo model parameters. (a) General, (b) BT, (c) MMS

3.1 Output Interface

A black window will show the evolution of the world, in case we are using a BT virus diffusion the nodes will move randomly and change color depending on their status, otherwise for MMS, nodes will only change color without any mobility since such behave is irrelevant in MMS.
Beside this, it's possible to view the development of the model by monitoring the number of susceptible, infected and resistant nodes in a dynamical chart. Figure 3 shows the output interface.

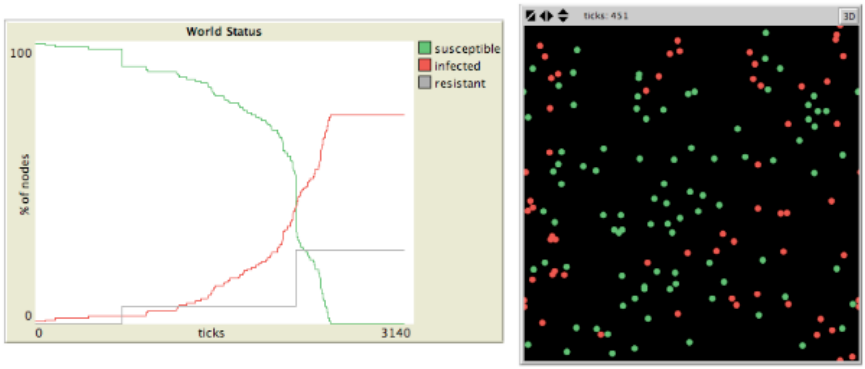

Fig. 3. Output interface in NetLogo

4 Experimental results

As already mentioned UA is the major focus of our study, so we would like to answer the question: How much is the impact of UA in the virus diffusion ?. To answer this question we take each method of virus diffusion separately and evaluate it singularly.
The total number of nodes used for the experiments were 150, and the simulation starts with only 1 infected node. For MMS mode we have two types of diffusion (Scanning/Topological), for Scanning we fixed the maximum number of attacks to 3 (2% of the nodes) which means infected nodes can perform 3 random attacks, one attack every tick, the spread chance was set to 0,2% this approximation is based on other studies made in [2]. While keeping these parameters constant we varied the User Awareness value five different times setting it's value to 5%, 25%, 45%, 65% and 90%. Figure 4 shows how User Awareness can positively effect the infection obstruction, all these values are average values for 20 simulations. As we can see for example, with a 45% of Awareness level we can almost guarantee 27.46% of nodes resistant to the virus and the infected nodes

percentage decreases to 72.55. It's important remind that infected nodes checking time is equal to a day equivalent to 720 ticks in NetLogo (1Tick = 2min), so in order to have effective results the virus diffusion must be slow enough to ensure a bigger immunity to it, the time chart in Figure 4 shows that the average time needed to spread the virus throw a Scanning strategy without UA is equal to 2828.1 ticks, so this will ensure at least 3-4 checking.

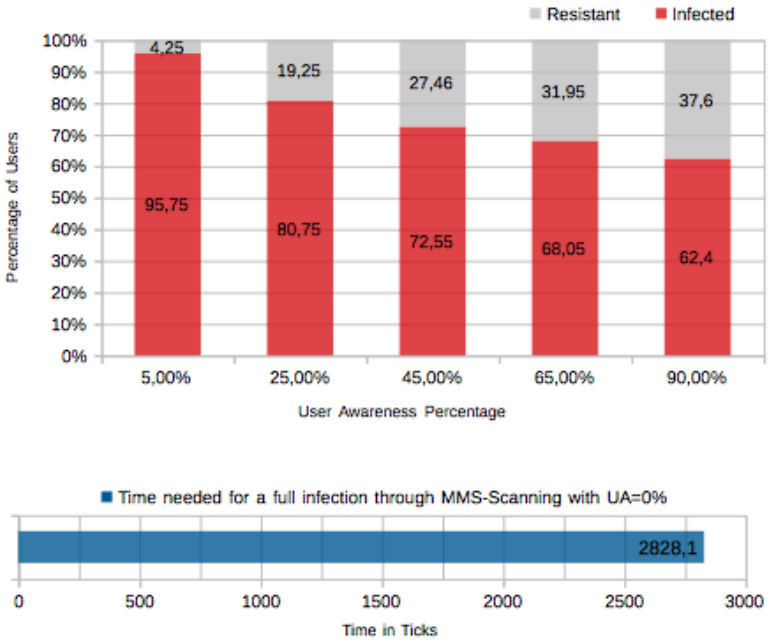

Fig. 4. The effect of UA on MMS-Scanning virus diffusion

Another important factor is the maximum number of attacks an infected node can do, so we tried to increased it to 6 (4% of the nodes) and left all the other parameters values as before. As Figure 5 shows, the percentage of infected nodes is slightly higher but still with good results.

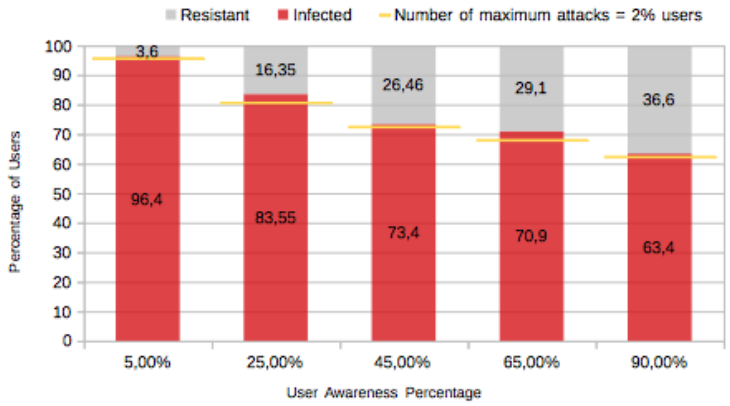

Fig. 5. The effect of UA on MMS-Scanning virus diffusion with a higher number of maximum attacks

The other diffusion strategy is the Topological one, we can notice that the diffusion speed is highly related to the address book size of the users as we can see in the second chart in Figure 6, so as expected if users maintain in their phones more contacts the virus spread speed will increase, for example if users address book contain 5 (3,33% of the nodes) the time needed for a total diffusion will be almost 700 ticks, which will impede a virus check for the infected users (720 ticks), On the other hand it's important to keep in mind that few contacts will make some users unreachable for the virus, because non of them is present in an address book of any user, so their status will remain susceptible. we can see this relation in the first chart in Figure 6.

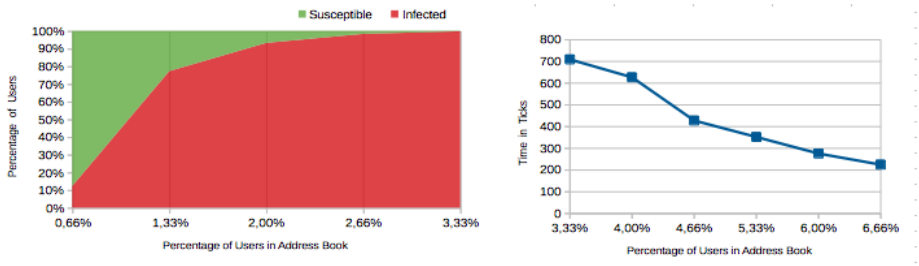

Fig. 6. Percentage of susceptible/infected nodes and time needed for a full infection with different address book sizes

That said, in order to study the User Awareness in this case, we set the percentage of phone numbers in the address book of each user to 2% and the same

number of maximum attack of the Scanning tests which is 3. The virus diffusion probability assigned is three time higher 0.6%: which is an approximation to indicate that a topological infection process is more powerful than a Scanning approach with a lower number of maximum attacks, as already studied in [2]. Figure 7 shows the results obtained with 3 different UA values, the graphics shows also a part of users that remained in susceptible status, which means they weren't included in any address book of any infected phone. The percentage of resistant users is getting bigger as we increase the Users Awareness value as expected, despite the fact that the results were way lower than the results obtained with the scanning experiments, the main reason is the virus spreading speed, with a Topological strategy the spreading process is fast enough to limit the propagation of an immunity to it.

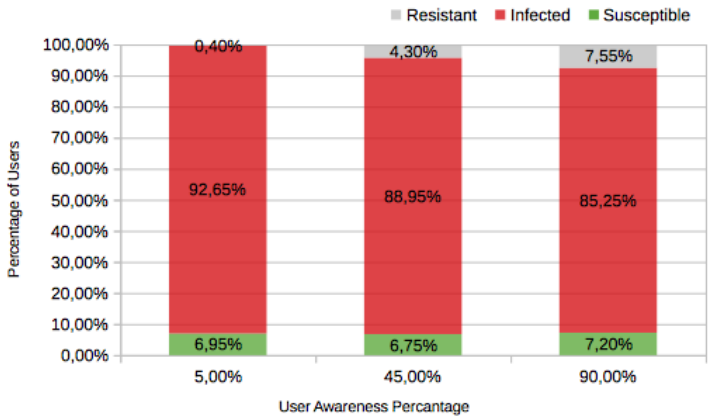

Fig. 7. The effect of UA on MMS-Topological virus diffusion

Now we can move on to the next diffusion strategy, the Bluetooth mechanism is strictly connected to the users movement as we already mentioned, but in our experiments we will not take in consideration the mobility patterns and suppose that users move randomly without any specific pattern. Another important factor is the Market Share (m), which indicates that a cell phone virus can infect only the phones with the operating system it was designed for, so for example a low value of m will make a virus propagation through Bluetooth difficult, because users with same OS will be harder to reach. To study the time needed for a full diffusion of a virus with different Market Share values we made some experiments varying the m value and writing down the time needed for each different m value. Figure 8 shows the results obtained.

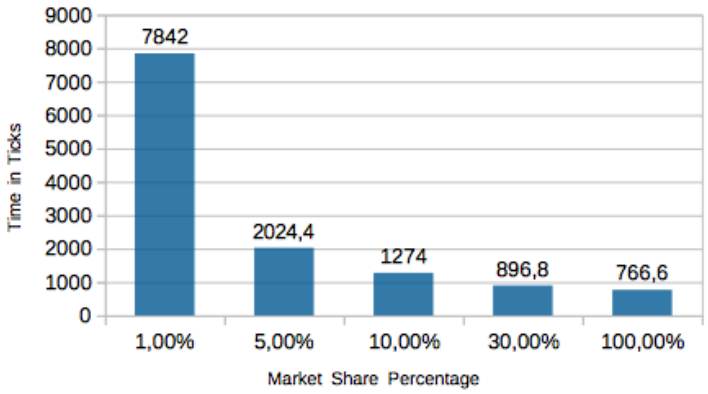

Fig. 8. Full virus propagation time with different Market-Share values

What we would like to know now is the effect of the User-Awareness on Bluetooth diffusion, so we fixed the Market Share value to 5% and varied the User Awareness level. As we can see in Figure 9 as the UA value get's higher the time needed for the virus to reach and infect all the users is getting higher, for example with a 65% User-Awareness the time requested has almost doubled to an average of 4296.2 ticks. So the virus is making a full diffusion anyway but with different time, because as we should remind, the UA in this case will make a node accept or decline a message temporarily, so a future request can still be made by the infected phones.

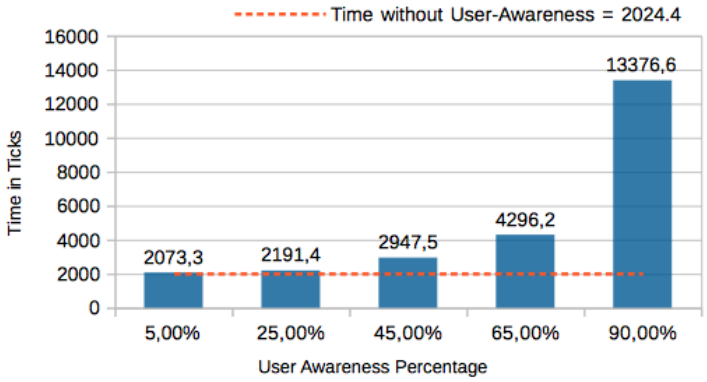

Fig. 9. Full virus propagation time with different UA values

So this factor is making the propagation slower, which should help the diffusion of opposing systems to the virus, in order to evaluate and quantify this contribution we set the recovery chance value to 0.5%, and each user will try to recover after a maximum time frequency of 20 ticks, if the recover process ends positively, the infected node will turn his status into susceptible again, and not into resistant status, as already mentioned we assume that the virus can evolve into a stronger version and infect again the same user. Figure 10 shows the fluctuations of infected and susceptible nodes as the time goes on, in the first graphic the UA value is null while for the next graphics we increased it's value. As expected the User Awareness is a good support for the progression of an antiviral process, so we can see that infected users fluctuations are getting closer to smaller percentage values, for instance with 90% UA value the percentage of susceptible phones is even higher than the infected one.

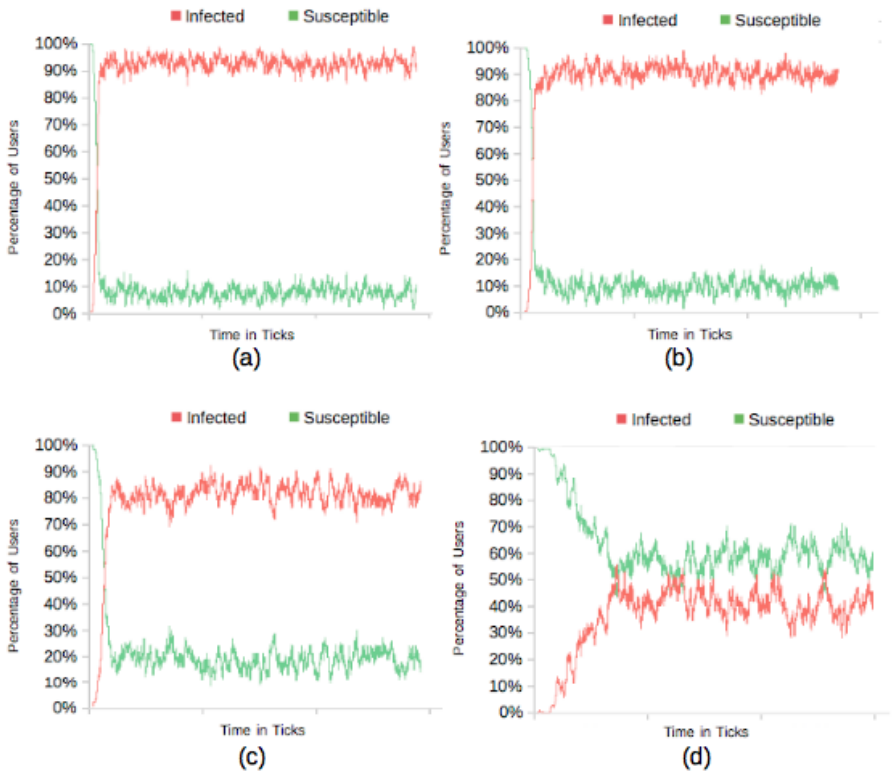

Fig. 10. System status with different UA values. (a) UA=0, (b) UA=25%, (c) UA=65%, (d) UA=90%

5 Discussion and conclusion

The main aim of this paper was to characterize the User-Awareness behavior, so the purpose of the experiments we made was to show it's effect on a virus propagation scenario.
Our tests showed us that UA can act and have different consequences on virus diffusion :

- Delay the virus diffusion : the tests made for the Bluetooth mechanism showed that slowing down the virus propagation will help the spread of an antiviral systems, the benefits in terms of time we can gain from even only a 45% of UA can really turn to be a big factor.
- Give total immunity to the virus: we made the MMS tests assuming that the UA will give the users a resistance to the virus, this behave can be really powerful, so in this case we demand the users to be more careful when opening strange messages and on the other hand to act singularly as an informer for their friends when they detect it.

So the UA is some kind of conscious conduct a phone user can obtain throw an education on the new technologies each day we are facing, nowadays we still don't see that much attention on this topic, although the large increase of internet and connection technologies. As always operating with a powerful and very useful tool needs also a much higher concern and knowledge about it's risks. In this case we can face really dangerous risks for our personal data and hardware devices.
We thing that achieving a 30% of users that act with awareness regard this kind of problems isn't so far from our reality, we just need to focus and dedicate some time to educate users on the way they should act when dealing with new technologies.

References

1. P. Wang, M. Gonzalez, C. A. Hidalgo, A.-L. Barabasi,: Understanding the spreading patterns of mobile phone viruses. Science 324, 1071-1076 (2009)
2. P. Wang, M. Gonzalez, R. Menezes, A-L Barabasi: Understanding the spread of malicious mobile-phone programs and their damage potential. 20 June 2013
3. X-P. Han, Z-D. Zhao, T. Hadzibeganovic, B-H. Wang: Epidemic spreading on hierarchical geographical networks with mobile agents. September 2013