**Saint Petersburg State University**

**Department of mathematical game theory and statistical decisions**

Taynitskiy Vladislav Aleksandrovich

Master's thesis

# Optimal Program of Resistance to Propagation of Malwares in Computer Networks

Specialization 01.04.02

Applied Mathematics and Informatics

Master's Program Game Theory and Operations Research

Research advisor,

C. Sc.(PhD), Associate Professor

Gubar E.A.

Saint Petersburg

2016

# Contents

# 1 Abstract

The security risks incurred by the spread of malware in computer and wireless networks can be reduced by the immunization of nodes, using security and antivirus patches. Malware, which captures personal and corporate confidential data, induces different damages, including costs generated by the necessity to compensate disclosure of private information, loss of money and social damage caused by loss of reputation. The distribution of security and antivirus patches in a network enables the control of the proliferation of malicious software and decreases possible losses. We formulate an optimal control problem for the case whether two different types of malware can circulate at the same time in a computer network and present an optimal strategy of resistance to malwares.

## 2 Introduction

Spreading of information in computer and wireless networks has become faster and a greater number of people use their network access for different activities, i.e. to retrieving financial information, to managing their banking accounts to purchase goods online etc. At the same time, variety applications of networks increase the probability of security threats. Malware is able to gain illegal access to confidential data such as bank accounts, credit cards, email and social networks passwords, collect sensitive private information and disrupt computer functionalities. The malware attack may lead to direct and indirect losses such as the cost of repairing software and hardware and the recovery of compromised servers. For example, well-known successful attacks of computer viruses Kido in 2009, MyDoom and ILOVEYOU affect millions of computers worldwide, with approximate damage, 9.1 billion, 38 and 15 billion, respectively.

The self-propagation and replication of computer viruses are similar to those processes of biological viruses [7, 8]. In biology, a virus is an infectious agent that replicates only inside the living cells of other organisms. Once a virus invades the host cell, it copies itself, infects the host and leaves it. Host cells are not always killed by the action of the virus. In fact, there are viruses that leave the infected cell alive, but use it as a continuous media to produce generations of viruses. In a similar fashion, a computer virus generates new copies of itself, inject itself into the code of other programs, the system memory, and distributes its copies into a variety of communication channels. Such virus behavior can be captured by an epidemic process which is described by the system of nonlinear differential equations as well as in classical Susceptible-Infected-Recovered (SIR) model [11].

Each node in the network may be considered as susceptible while it is not invaded by the replica of the malware transmitted from the infected nodes. The node becomes recovered after the application of antivirus or security patches. We assume that the protection software can effectively protect the nodes, and they cannot be reinfected once antivirus software is installed. The adoption of antivirus software provides a mechanism to control the propagation of malware, and hence protects the network. However, the challenge with a wide adoption of the software is the tradeoff between resource utilization and security risks. The scanning and monitoring process of antivirus software consumes computational resources. When the security risks are low, it is more desirable to reduce the rate of scanning and monitoring to achieve a better usage of computer systems.

Some viruses can generate epidemic process in population periodically, for example epidemics of influenza in urban population can reach its peak two times during one epidemic period. Similarly, many examples of several waves of spreading the identical malicious software in computer and wireless networks are well-known. In paper [21], authors have analyzed that according to a surviving probability different homogeneous groups of viruses can preserve for long time period and provoke new attacks. As the examples of repeated virus epidemics it can be considered the attacks of Code Red, Code Red II and Conficker between 2001 and 2013 which caused the damages of more than 200000 of computers worldwide [14]. Due to these reasons it is possible to use a series of impulse control actions which can be applied in certain time moments or adhere to time interval.

For this reasons, we will consider different models describing the behaviour of control schemes of antivirus spreading. We establish an optimal control framework to characterize the fundamental characteristics for network security. We use epidemic dynamics to define the virus propagation, and a finite-horizon cost function to capture the tradeoffs between antivirus adoption and the impact of virus propagation. We aim to find an optimal control strategy of antivirus protection that minimizes resource consumption. In conclusion we will compare impulse and continuous types of spreading of protection software and make conclusion about the advantages of both types of antiviruses.

Our work is related to previous studies in this area, including [3, 5, 9, 23]. In current study, the network can be attacked by a heterogeneous source of malware which captures the fact of coexistence of different types of exploits and vulnerability of the existing computing systems. The challenge for modeling heterogeneous malware spreading is multi-fold. We consider a network attacked by two types of malware [9], where each node can be infected either separately by each type of malware or by both types simultaneously. First, we model the dynamics of propagation of the malware in the network in case if both types coexist in one host node. Second, we formulate an optimal control problem and show the structure of the optimal strategies, which provides the minimum of the aggregated system costs depends on the properties of value-functions.

The work is organized as follows. Sections 3 and 4 present mathematical models of epidemics and formulates the optimal control problem in case of continuous control. In subsections 3.3 and 4.3, using Pontryagin's maximum principle, we define the structure optimal control policies and proof main results. Subsection 4.4 focuses on the stability analysis of the uncontrolled system. Numerical examples will be presented in subsection 3.4 and 4.5. Section 5 describes the mathematical model of spreading of viruses as impulse control problem. Structure of impulse control problem and numerical simulation are presented in subsections 5.3 and 5.4. Section 6 concludes the work and main results.

# 3 SIR model with continuous control 1

## 3.1 Mathematical model

In this section, we study a network of $N$ nodes, where two types of malicious software spread with different speeds. Malicious software propagates very fast, hence Susceptible-Infected-Recovery (SIR) model needs to be adapted to describe the epidemics of viruses in computer networks. All nodes in the network are divided into three groups: *Susceptible* ($S$), *Infected* ($I$) and *Recovered* ($R$) [11]. *Susceptible* is a group of nodes which have not contact with infected nodes yet and may be invaded by any forms of malicious software. *Infected* nodes are already attacked by the virus and *Recovered* is a group of dispatched nodes. Since two types of malware circulate in the network, the Infected consist of the subgroup of nodes infected by the first form of malware $V_1$, the subgroup of nodes infected by the second form $V_2$ and the group infected by both forms of viruses. We model the epidemic process as a system of nonlinear differential equations, where $n_S$, $n_{V_1}$, $n_{V_2}$, $n_R$ correspond to the number of susceptible, infected by different forms of malware $V_1$, $V_2$ and recovered nodes, respectively. The variable $n_{V_{12}}$ is a number of nodes simultaneously infected by both viruses. The total number of nodes in the network during the entire process remains constant and equal to $N$, $n_S + n_{V_1} + n_{V_2} + n_{V_{12}} + n_R = N$.

Let $S(t) = \frac{n_S(t)}{N}$, $I_1(t) = \frac{n_{V_1}(t)}{N}$, $I_2(t) = \frac{n_{V_2}(t)}{N}$, $I_{12} = \frac{n_{V_{12}}(t)}{N}$, $R(t) = \frac{n_R(t)}{N}$ as a fraction of the *Susceptible*, the *Infected* by virus $V_1$, $V_2$, both viruses together and the *Recovered* nodes, respectively. At the beginning of the epidemics $t = 0$, the majority of the nodes belong to the Susceptible group, and a small fraction of nodes is infected by different types of malware. Hence initial states are given by $0 < S(0) = S^0$, $0 < I_1(0) = I_1^0$,

$0 < I_2(0) = I_2^0$, $0 < I_{12}(0) = I_{12}^0$, $R(0) = R^0 = 1 - S^0 - I_1^0 - I_2^0 - I_{12}^0$.

A susceptible node becomes infected whenever it accepts or installs malicious software, spread by the infected nodes. Malware spreads in the network at the rate of $\beta_1$ and $\beta_2$ for $V_1$ and $V_2$, respectively. In particular, if the node is infected by the malware $V_1$, then with probability $\varepsilon\beta_2$ it can be infected by the second malware $V_2$, and vice versa, if the node is infected with a virus $V_2$ , then with probability $\varepsilon\beta_1$ it may be infected by $V_1$. Here, the variable $\varepsilon \in [0, 1]$ is a probability that a node infected by one type virus will be infected by another type virus. If a susceptible node contacts with a node infected by both viruses $V_1$ and $V_2$, then with probability $\beta_i$ it may be invaded by only one form of malware $V_i$, $i = 1, 2$.

Usually, a majority of nodes are protected by permanent antivirus software which is effective against known viruses. Then, we can consider a recovery rate $\gamma$, which show the probability that susceptible nodes are recovered by permanent antivirus software. However, periodically, the epidemics of new computer viruses appear, and the permanent antivirus software is often inefficient against new or unknown malicious software. In this case, special patches can be applied to protect the network. We define $u_i(t)$, $i = 1, \dots, 4$ as a control parameter which corresponds to the application of special antivirus patches. Parameter $\sigma_1$, $\sigma_2$ and $\sigma_3$ may be interpreted as self-recovery rate in biological system. We may discuss about the rate of self-recovery if a node is repaired without any external resources or employees.

We modeled the epidemics of two forms of malware using a system

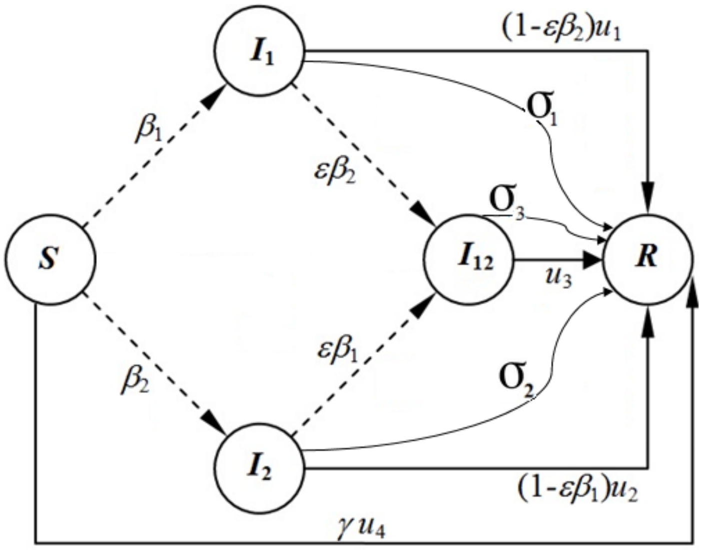


Figure 1: The scheme of the propagation of two forms of malwares.

of nonlinear differential equations:

$$
\begin{aligned}
&\dot{S} = -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) - \gamma S u_4, \\
&\dot{I}_1 = \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_1 I_1 - (1 - \varepsilon\beta_2)u_1 I_1, \\
&\dot{I}_2 = \beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2 - (1 - \varepsilon\beta_1)u_2 I_2, \\
&\dot{I}_{12} = \epsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_3 I_{12} - u_3 I_{12}, \\
&\dot{R} = \gamma S u_4 + (1 - \varepsilon\beta_2)u_1 I_1 + (1 - \varepsilon\beta_1)u_2 I_2 + u_3 I_{12} + \sigma_1 I_1 + \\
&\qquad + \sigma_2 I_2 + \sigma_3 I_{12}, \\
&S + I_1 + I_2 + I_{12} + R = 1.
\end{aligned}
\tag{1}
$$

Hence initial states are

$$
\begin{aligned}
&S(0) = S^0,\ I_1(0) = I_1^0,\ I_2(0) = I_2^0,\ I_{12}(0) = I_{12}^0, \\
&R(0) = R^0 = 1 - S(0) - I_1(0) - I_2(0) - I_{12}(0).
\end{aligned}
\tag{2}
$$

### 3.2 Objective function

The objective of the system designer is to minimize the aggregated cost on time interval $[0, T]$. At any given $t$, the overall system costs include infected costs $f_1(I_1(t))$, $f_2(I_2(t))$, $f_3(I_{12}(t))$ and protection costs $h_i(u_i(t))$, $i = 1, \ldots, 4$. Infected costs are includes losses caused by infected nodes; protection costs are generated by the consumption of resources for the application of antivirus or stationary security patches. Functions $f_i$ are non-decreasing and twice-differentiable, such as $f_i(0) = 0$, $f_i(I_i) > 0$ for $I_i > 0$. Functions $h_i(u_i(t))$ are twice differentiable and increasing in $u_i(t)$ such as $h_i(0) = 0$, $h_i(u_i) > 0$, when $u_i > 0$, $u_i \in [0, 1]$ for $i = 1, \ldots, 4$.

Aggregated system costs are defined as the functional:

$$\begin{aligned} J(u_1, u_2, u_3, u_4) = \int_0^T (f_1(I_1(t)) + f_2(I_2(t)) + f_3(I_3(t)) \\ + h_1(u_1(t)) + h_2(u_2(t)) + h_3(u_3(t)) + h_4(u_4(t)))dt. \end{aligned} \tag{3}$$

The optimal control problem is to minimize the functional in a time interval $[0, T]$, i.e.,

$$\min_{u_1, u_2, u_3, u_4} J(u_1, u_2, u_3, u_4). \tag{4}$$

## 3.3 Structure of the optimal control

We will use Pontryagin's maximum principle [16] to solve the problem described above. There exist continuous and piecewise continuously differentiable co-state functions $\lambda_i$ that at every point $t \in [0; T]$ where $u_i(t)$, $i = 1, \ldots, 4$ are continuous, satisfy (8) and (9).

$$(u_1, \ldots, u_4) \in \arg\max_{\underline{u_1}, \ldots, \underline{u_4}} H = (\lambda, S, I_1, I_2, I_{12}, R, \underline{u_1}, \ldots, \underline{u_4}). \tag{5}$$

Here, $\underline{u_1}, \underline{u_2}, \underline{u_3}$ and $\underline{u_4}$ are feasible controls.

Define functions $\psi_i$, $i = 1, \dots, 4$ as follows:

$$\begin{aligned} &\psi_1 = (1-\varepsilon\beta_2)(\lambda_R - \lambda_{I_1})I_1, \quad &&\psi_2 = (1-\varepsilon\beta_1)(\lambda_R - \lambda_{I_2})I_2, \\ &\psi_3 = (\lambda_R - \lambda_{I_{12}})I_{12}, &&\psi_4 = \gamma S(\lambda_R - \lambda_S). \end{aligned} \tag{6}$$

Hamiltonian $H$ of the system (1):

$$\begin{aligned} H = &-(f_1(I_1) + f_2(I_2) + f_3(I_{12}) + h_1(u_1) + h_2(u_2) + \\ &h_3(u_3) + h_4(u_4)) + \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) + \\ &\beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) + \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) + \\ &\varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) + (\lambda_R - \lambda_{I_1})\sigma_1 I_1 + \\ &(\lambda_R - \lambda_{I_2})\sigma_2 I_2 + (\lambda_R - \lambda_{I_{12}})\sigma_3 I_{12} + \\ &(1-\varepsilon\beta_2)u_1 I_1(\lambda_R - \lambda_{I_1}) + (1-\varepsilon\beta_1)u_2 I_2(\lambda_R - \lambda_{I_2}) + \\ &u_3(\lambda_R - \lambda_{I_{12}})I_{12} + (\lambda_R - \lambda_S)\gamma S u_4. \end{aligned} \tag{7}$$

Adjoint system is

$$\begin{aligned} \frac{d\lambda_S}{dt} &= \beta_1(\lambda_S - \lambda_{I_1})(I_1 + I_{12}) + \beta_2(\lambda_S - \lambda_{I_2})(I_2 + I_{12}) - \\ &\quad \gamma u_4(\lambda_R - \lambda_S); \\ \frac{d\lambda_{I_1}}{dt} &= \frac{df_1(I_1)}{dI_1} + \beta_1 S(\lambda_S - \lambda_{I_1}) + \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) + \\ &\quad \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + (\lambda_{I_1} - \lambda_R)\sigma_1 I_1 + u_1(1-\varepsilon\beta_2)(\lambda_{I_1} - \lambda_R); \\ \frac{d\lambda_{I_2}}{dt} &= \frac{df_2(I_2)}{dI_2} + \beta_2 S(\lambda_S - \lambda_{I_2}) + \varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}}) + \\ &\quad \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + (\lambda_{I_2} - \lambda_R)\sigma_2 I_2 + u_2(1-\varepsilon\beta_1)(\lambda_{I_2} - \lambda_R); \\ \frac{d\lambda_{I_{12}}}{dt} &= \frac{df_3(I_{12})}{dI_{12}} + \beta_1 S(\lambda_S - \lambda_{I_1}) + \beta_2 S(\lambda_S - \lambda_{I_2}) + \\ &\quad (\lambda_{I_{12}} - \lambda_R)\sigma_3 I_{12} + \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \\ &\quad u_3(\lambda_{I_{12}} - \lambda_R); \\ \frac{d\lambda_R}{dt} &= 0. \end{aligned} \tag{8}$$

together with the transversality conditions

$$\lambda_S(T) = \lambda_{I_1}(T) = \lambda_{I_2}(T) = \lambda_{I_{12}}(T) = \lambda_R(T) = 0. \tag{9}$$

According to [3, 5, 16] we construct the optimal program to prevent the spreading of malicious software in a computer network.

**Proposition 1** *1) If $h_i$ is strictly convex function ($h_i''(u_i) > 0$) then there exist time moments $t_0, t_1 \in [0, T]$, $0 \leq t_0 \leq t_1 \leq T$ such that:*

$$u_i = \begin{cases} 0, & \text{if } \ t_1 < t < T; \\ h_i'^{-1}(\psi_i), & \text{if } \ t_0 < t \leq t_1; \\ 1, & \text{if } \ 0 < t < t_0. \end{cases} \tag{10}$$

*2) If $h_i$ is concave function ($h_i''(u_i) < 0$) then there exist time moment $t_1 \in [0, T]$ such that:*

$$u_i = \begin{cases} 0, & \text{if } \ t_1 < t < T; \\ 1, & \text{if } \ 0 < t < t_1. \end{cases} \tag{11}$$

*where functions $\psi_i$, $i = 1, \ldots, 4$ are follows:*

**Proof of the Proposition 1.** Rewrite Hamiltonian in terms of function $\psi_i$ and we obtain:

$$\begin{aligned} H = &-(f_1(I_1) + f_2(I_2) + f_3(I_{12})) + \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) + \\ &\beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) + \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) + \varepsilon\beta_1 I_2(I_1 + I_{12}) + \\ &(\lambda_R - \lambda_{I_1})\sigma_1 I_1 + (\lambda_R - \lambda_{I_2})\sigma_2 I_2 + (\lambda_R - \lambda_{I_{12}})\sigma_3 I_{12} + \\ &(-h_1(u_1) + u_1\psi_1) + (-h_2(u_2) + u_2\psi_2) + (-h_3(u_3) + u_3\psi_3) + \\ &(-h_4(u_4) + u_4\psi_4). \end{aligned} \tag{12}$$

According to the algorithm of maximum principle we consider derivatives:

$$\frac{\partial H}{\partial u_i} = -\dot{h}_i(u_i) + \psi_i = 0, \ i = 1, \ldots, 4. \tag{13}$$

As $h_i(u_i)$ are increasing functions and $I_i \geq 0$ and $S \geq 0$ then Hamiltonian reaches its maximum if $\psi_i = \dot{h}_i(u_i) \geq 0, \ i = 1, \ldots, 4$. It follows if and only if the following conditions are satisfied: $(\lambda_R(t) - \lambda_{I_1}(t)) \geq 0$, $(\lambda_R(t) - \lambda_{I_2}(t)) \geq 0$, $(\lambda_R(t) - \lambda_{I_{12}}(t)) \geq 0$ and $(\lambda_R(t) - \lambda_S(t)) \geq 0$. To complete the proof of proposition we consider auxiliary lemma.

**Lemma 1** *For all* $0 \leq t \leq T$, *we have* $(\lambda_R(t) - \lambda_{I_1}(t)) \geq 0$, $(\lambda_R(t) - \lambda_{I_2}(t)) \geq 0$, $(\lambda_R(t) - \lambda_{I_{12}}(t)) \geq 0$ *and* $(\lambda_R(t) - \lambda_S(t)) \geq 0$.

Lemma 1 is proved in the similar way to those in [3], [9] and it is based on the following properties.

**Property 1** *Let* $v(t)$ *be a continuous and piecewise differential function of* $t$. *Let* $v(t_1) = L$ *and* $v(t) > L$ *for all* $t \in (t_1, \ldots, t_0]$. *Then* $\dot{v}(t_1^+) \geq 0$. *Where* $v(t_1^+) = \lim_{x \to 0} v(x)$.

**Property 2** *For any convex and differentiable function* $y(x)$, *which is* $0$ *at* $x = 0$, $y'(x)x - y(x) \geq 0$ *for all* $x \geq 0$.

**Proof of Lemma 1.**

Lets split our proof into two parts. At the first part we will consider the case when $t = T$ and show that derivatives of the functions $(\lambda_R(t) - \lambda_{I_1}(t))$, $(\lambda_R(t) - \lambda_{I_2}(t))$, $(\lambda_R(t) - \lambda_{I_{12}}(t))$ and $(\lambda_R(t) - \lambda_S(t))$ are less or equal to zero to prove that they are non-increasing at $t = T$. In the second part we will use the method of proof by contradiction and show that on the whole interval $[0, T]$ these functions are also non-negative.

**I.** At time moment $T$, we have according to (9):

$$(\lambda_R(T) - \lambda_{I_1}(T)) = 0; \quad (\lambda_R(T) - \lambda_{I_2}(T)) = 0;$$
$$(\lambda_R(T) - \lambda_{I_{12}}(T)) = 0; \quad (\lambda_R(T) - \lambda_S(T)) = 0. \tag{14}$$

From (8) we receive

$$\begin{aligned}
&(\dot{\lambda}_R(T) - \dot{\lambda}_{I_1}(T)) = -\dot{f}_1(I_1(T)) \le 0;\\
&(\dot{\lambda}_R(T) - \dot{\lambda}_{I_2}(T)) = -\dot{f}_2(I_2(T)) \le 0;\\
&(\dot{\lambda}_R(T) - \dot{\lambda}_{I_{12}}(T)) = -\dot{f}_3(I_{12}(T)) \le 0;\\
&(\dot{\lambda}_R(T) - \dot{\lambda}_S(T)) = 0.
\end{aligned} \tag{15}$$

Now we have that at time moment $T$ all functions are equal to 0 and their derivatives are less or equal to 0 then we can obtain that $(\lambda_R(t) - \lambda_{I_1}(t))$, $(\lambda_R(t) - \lambda_{I_2}(t))$, $(\lambda_R(t) - \lambda_{I_{12}}(t))$ and $(\lambda_R(t) - \lambda_S(t))$ are decreasing functions at $t = T$.

**II.** (Proof by contradiction)

Let $0 \le t^* \le T$ be the last instant moment at which one of the inequality constraints are satisfied:

$$(\lambda_R(t) - \lambda_{I_1}(t)) \ge 0, \quad (\lambda_R(t) - \lambda_{I_2}(t)) \ge 0,$$
$$(\lambda_R(t) - \lambda_{I_{12}}(t)) \ge 0, \quad (\lambda_R(t) - \lambda_S(t)) \ge 0.$$

First, suppose that following inequalities are hold

$$\begin{aligned}
&(\lambda_R(t^*) - \lambda_{I_1}(t^*)) = 0, && (\lambda_R(t^*) - \lambda_{I_2}(t^*)) \ge 0,\\
&(\lambda_R(t^*) - \lambda_{I_{12}}(t^*)) \ge 0, && (\lambda_R(t^*) - \lambda_S(t^*)) \ge 0,\\
&(\lambda_S(t^*) - \lambda_{I_1}(t^*)) \ge 0, && (\lambda_S(t^*) - \lambda_{I_2}(t^*)) \ge 0\\
&(\lambda_{I_1}(t^*) - \lambda_{I_{12}}(t^*)) \ge 0, && (\lambda_{I_2}(t^*) - \lambda_{I_{12}}(t^*)) \ge 0.
\end{aligned} \tag{16}$$

From (8) we have:

$$\dot{\lambda}_R(t^*) - \dot{\lambda}_{I_1}(t^*) = -\frac{df_1}{dI_1} - \beta_1 S(\lambda_S - \lambda_{I_1}) - \\ \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) - \\ (\lambda_{I_1} - \lambda_R)\sigma_1 I_1 - u_1(1 - \varepsilon\beta_2)(\lambda_{I_1} - \lambda_R). \tag{17}$$

Since $(\lambda_R(t) - \lambda_{I_1}(t))$ is decreasing function on the interval $[0;T]$ then according to Property 1, we consider a time moment $t^{*+}$ such as:

$$\dot{\lambda}_R(t^{*+}) - \dot{\lambda}_{I_1}(t^{*+}) = -\frac{df_1}{dI_1} - \beta_1 S(\lambda_S - \lambda_{I_1}) - \\ \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}). \tag{18}$$

According to assumption (16), the difference $(\dot{\lambda}_R(t) - \dot{\lambda}_{I_1}(t))$ is negative at the time moment $t^{*+}$, then function $(\lambda_R(t) - \lambda_{I_1}(t))$ decreases. It contradicts Property 1 and proves that our function is not increasing on interval $t \in [0;T]$. We get that $(\lambda_R(t) - \lambda_{I_1}(t)) \geq 0$ for $t \in [0;T]$.

Second part of the proof is similar to presented above and presented in Appendix 1.

The proof of Lemma 1 is completed.

Now we return to the main proposition and consider two cases which depend on the properties of $h_i(u_i)$, $i = 1, \ldots, 4$.

1) $h_i$ **is concave.** Since functions $h_i$ are concave ($h_i'' < 0$, $i = 1, \ldots, 4$), then $(u_i\psi_i - h_i(u_i))$ are convex functions of $u_i$. Hamiltonian $H$ is a strictly convex function according to (7) and for any $t \in [0,T]$ and it reaches its maximum either at $u_i = 1$ or $u_i = 0$, $i = 1, \ldots, 4$.

$$u_i(t) = \begin{cases} 0, & \text{if } \psi_i < h_i(1); \\ 1, & \text{if } \psi_i > h_i(1). \end{cases} \tag{19}$$

2) $h_i$ **is strictly convex.** If functions $h_i$ are strictly convex ($h_i'' > 0$, $i = 1, \dots, 4$) then $(u_i\psi_i - h_i(u_i))$ and Hamiltonian $H$ are concave functions of $u_i$, therefore $\frac{\partial H}{\partial u_i} = -\dot{h}_i(u_i) + \psi_i = 0$ and

$$u_i(t) = \begin{cases} 0 \ , & \text{if } \psi_i \le \frac{dh_i(0)}{du_i}; \\ \frac{dh_i^{-1}(\psi_i)}{du_i} \ , & \text{if } \frac{dh_i(0)}{du_i} < \psi_i \le \frac{dh_i(1)}{du_i}; \\ 1 \ , & \text{if } \psi_i > \frac{dh_i(1)}{du_i}. \end{cases} \tag{20}$$

The proof of the Proposition 1 is completed.

### 3.4 Numerical simulation

In this section, we use numerical simulations to support our theoretical results. We use the following values and parameters: iteration step is $\delta = 0.25$, initial fractions of nodes are $S(0) = 0.55$, $I_1(0) = 0.15$, $I_2(0) = 0.2$, $I_{12}(0) = 0.1$, $R(0) = 0$. Intensive rates of transition from susceptible to infected and recovered are $\beta_1 = 0.6$, $\beta_2 = 0.7$ and $\gamma = 0.1$ respectively. Virus interaction rate is $\varepsilon = 0.8$. Intensive rates of transition from infected to recovered are $\sigma_1 = 0.03$, $\sigma_2 = 0.02$ and $\sigma_3 = 0.01$.We suppose that epidemic duration is $T = 15$ time units.

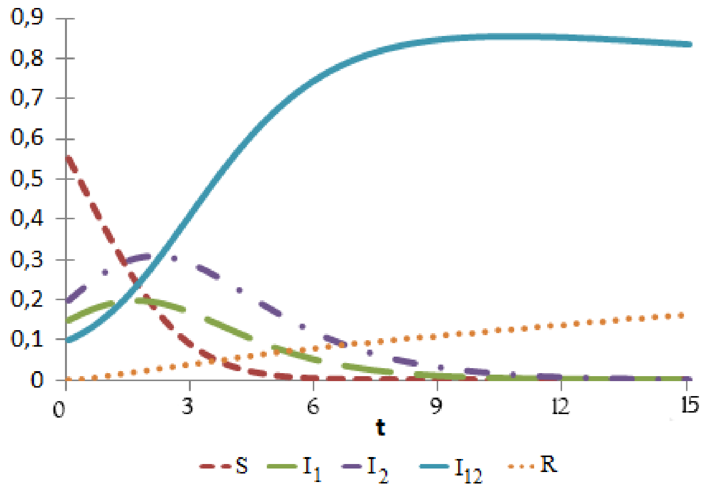


Figure 2: Uncontrolled SIR model.

First, we consider uncontrolled case. Since we do not apply any antivirus patches to reduce the epidemics, then 80 % of the hosts are transferred into the group infected $I_{12}$ at time $T = 15$, this is shown by the solid line in the Fig. 2. 20 % of the nodes are healed of the malware by self-recovery parameters $\sigma_i$ in the system (dot line).

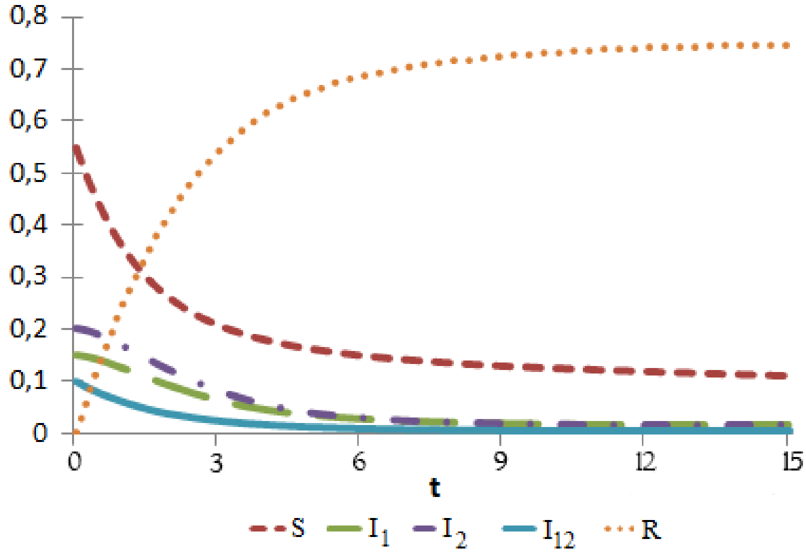


Figure 3: Controlled SIR model. Costs functions $h_i$ are convex.

Also, we illustrate the case when costs for treatment measures $h_i(u_i)$ are convex functions (Fig. 3). We use $h_1(u_1) = 0.35u_1^2$; $h_2(u_2) = 0.4u_2^2$; $h_3(u_3) = 0.5u_3^2$; $h_4(u_4) = 0.2u_4^2$.

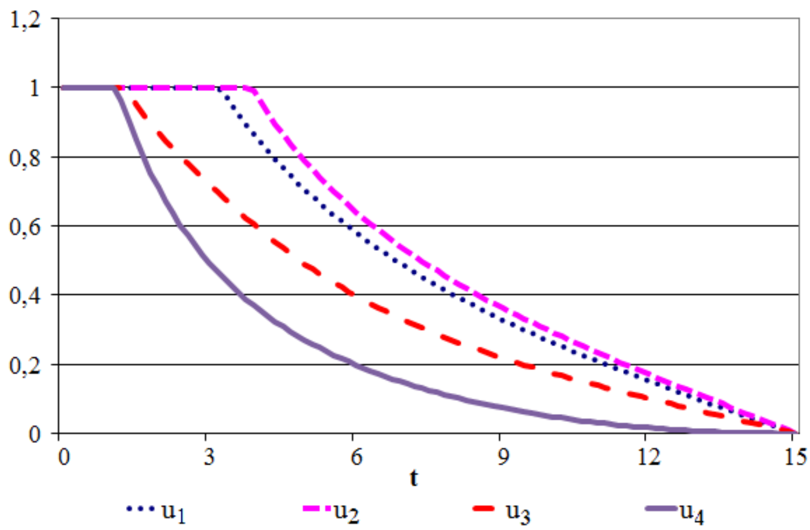


Figure 4: Optimal control in SIR model (functions $h_i$ are convex).

At the last time moment $T = 15$ we can see that there are no infected hosts, most hosts have transferred into recovered group and only some hosts are in a susceptible group. This fraction of susceptible have not been impact to the epidemic process. The structure of optimal control is shown in the Fig. 4.

The next diagram (Fig. 5) illustrates the comparison between aggregated system costs for uncontrolled and controlled system, which are equal to $J = 775.5$ and $J = 112.1$ monetary units respectively.

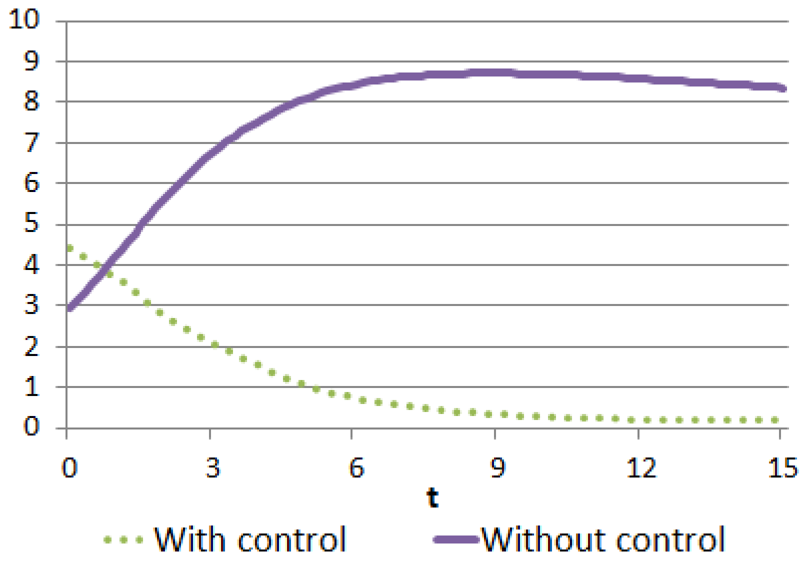


Figure 5: Comparison of aggregated system costs for controlled and uncontrolled SIR models.

# 4 SIR model with continuous control 2

## 4.1 Mathematical model

As in previous model (1) we study a network of homogenous $N$ nodes, with the assumption that two forms of malicious software spread in the network with different speeds. In contrast to the previous model, we assume that in this case there is no permanent antivirus software ($u_4(t) = 0$). Self-recovery parameters $\sigma_i$ which are describing transferring probabilities from infected to susceptible state. From that reason self-recovered node can be infected again.

Fig. 6 represents the evolution of the propagation malware in the network.

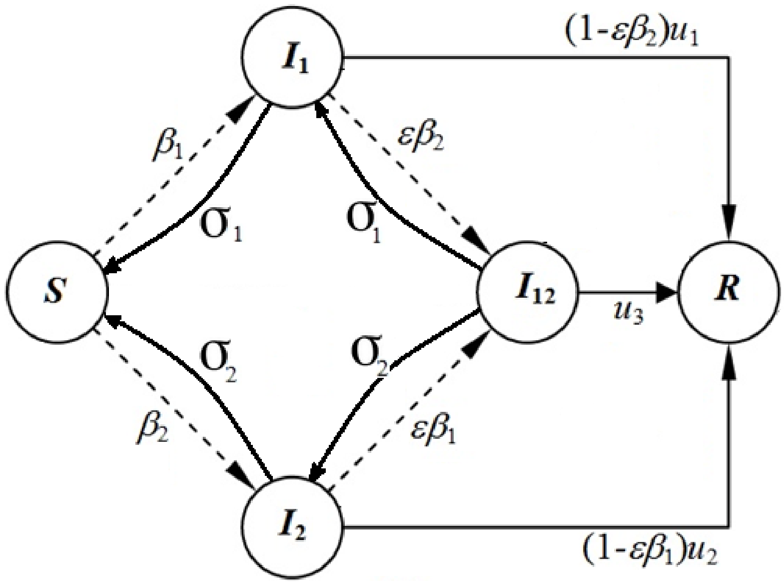


Figure 6: The scheme of the propagation of two forms of malware.

We have extended classical SIR model, and modeled the epidemics of two forms of malware using a system of nonlinear differential equations [5]:

$$
\begin{aligned}
\dot{S} &= -\beta_1 S(I_1+I_{12}) - \beta_2 S(I_2+I_{12}) + \sigma_1 I_1 + \sigma_2 I_2;\\
\dot{I}_1 &= \beta_1 S(I_1+I_{12}) - \varepsilon\beta_2 I_1(I_2+I_{12}) - \sigma_1 I_1 + \sigma_2 I_{12} - (1-\varepsilon\beta_2)u_1 I_1;\\
\dot{I}_2 &= \beta_2 S(I_2+I_{12}) - \varepsilon\beta_1 I_2(I_1+I_{12}) - \sigma_2 I_2 + \sigma_1 I_{12} - (1-\varepsilon\beta_1)u_2 I_2;\\
\dot{I}_{12} &= \varepsilon\beta_1 I_2(I_1+I_{12}) + \varepsilon\beta_2 I_1(I_2+I_{12}) - (\sigma_1+\sigma_2)I_{12} - u_3 I_{12};\\
\dot{R} &= (1-\varepsilon\beta_2)u_1 I_1 + (1-\varepsilon\beta_1)u_2 I_2 + u_3 I_{12}.
\end{aligned}
\tag{21}
$$

### 4.2 Objective function

As in subsection 3.2, we find optimal control parameters which minimize aggregated system costs over the time interval $[0,T]$. At any time moment $t$ we define infected costs $f_i(I_i(t))$ for infected nodes and protection costs $h_i(u_i(t))$, $i=1,2,3$. Functions $f_i$ are non-decreasing and twice-differentiable, such as $f_i(0)=0$, $f_i(I_i)>0$ for $I_i(t)>0$, $i=1,2,3$. Functions $h_i(u_i(t))$ are twice differentiable and increasing in $u_i(t)$ such as $h_i(0)=0$, $h_i(u_i)>0$, when $u_i(t)>0$, $u\in[0,1]$. $g(R)$ is benefit rate, it is non-decreasing and differentiable function and $g(0)=0$.

Aggregated system costs are defined by the following functional

$$
\begin{aligned}
J(u_1,u_2,u_3) = \textstyle\int_0^T (f_1(I_1(t)) + f_2(I_2(t)) + f_3(I_{12}(t)) + h_1(u_1(t)) +\\
h_2(u_2(t)) + h_3(u_3(t)) - g(R))dt,
\end{aligned}
\tag{22}
$$

and the optimal control problem is to minimize the functional on time interval $[0,T]$, i.e.,

$$
\min_{u_1,u_2,u_3} J(u_1,u_2,u_3). \tag{23}
$$

### 4.3 Structure of optimal control

Based on previous work [3] and [9], the optimal program was constructed and it has the following structure:

**Proposition 2** *1. If $h_i$ is strictly convex function then exists time moments $t_0, t_1 \in [0,T]$, $0 \le t_0 \le t_1 \le T$ such as:*

$$u_i(t) = \begin{cases} 0\ , & if\ \psi_i \le \frac{dh_i(0)}{du_i}; \\ \frac{dh_i^{-1}(\psi_i)}{du_i}\ , & if\ \frac{dh_i(0)}{du_i} < \psi_i \le \frac{dh_i(1)}{du_i}; \\ 1\ , & if\ \psi_i > \frac{dh_i(1)}{du_i}; \end{cases} \tag{24}$$

*2. If $h_i$ is concave function then exists time moment $t \in [0,T]$ such as:*

$$u_i(t) = \begin{cases} 0, & if\ \ \psi_i < h_i(1); \\ 1, & if\ \ \psi_i > h_i(1). \end{cases} \tag{25}$$

where functions $\psi_i$ are defined as follows:

$$\begin{aligned} \psi_1 &= (1-\varepsilon\beta_2)(\lambda_R - \lambda_{I_1})I_1, \\ \psi_2 &= (1-\varepsilon\beta_1)(\lambda_R - \lambda_{I_2})I_2, \\ \psi_3 &= (\lambda_R - \lambda_{I_{12}})I_{12}. \end{aligned} \tag{26}$$

**Proof of Proposition 2**

Proof of the Proposition 2 is similar as in previous section model. We use Pontryagin's maximum principle to find optimal control. The proof is in the Appendix 2.

### 4.4 Numerical simulation

In this section, we present numerical examples to study the SIR model. We use the following values and parameters: iteration step is $\delta = 0.15$, initial fractions of nodes are $S(0) = 0.55$, $I_1(0) = 0.15$, $I_2(0) = 0.2$, $I_{12}(0) = 0.1$, $R(0) = 0$. Intensive rates of transition from susceptible to infected and recovered are $\beta_1 = 0.6$, $\beta_2 = 0.7$. Virus interaction rate is $\varepsilon = 0.8$. Intensive rates of transition from infected to recovered are $\sigma_1 = 0.03$, $\sigma_2 = 0.02$ and $\sigma_3 = 0.01$.We suppose that epidemic duration is $T = 15$ time units.

First, we consider uncontrolled case (Fig. 7). Since we do not apply any antivirus patches to reduce the epidemics, then 90 % of the hosts are transferred into the group infected $I_{12}$ at time $T = 15$, this is shown by the solid line in the Fig. 2. There are no recovered nodes like at uncontrolled case in previous model because self-recovery heal nodes and transfer them to susceptible state. After some time they infects again.

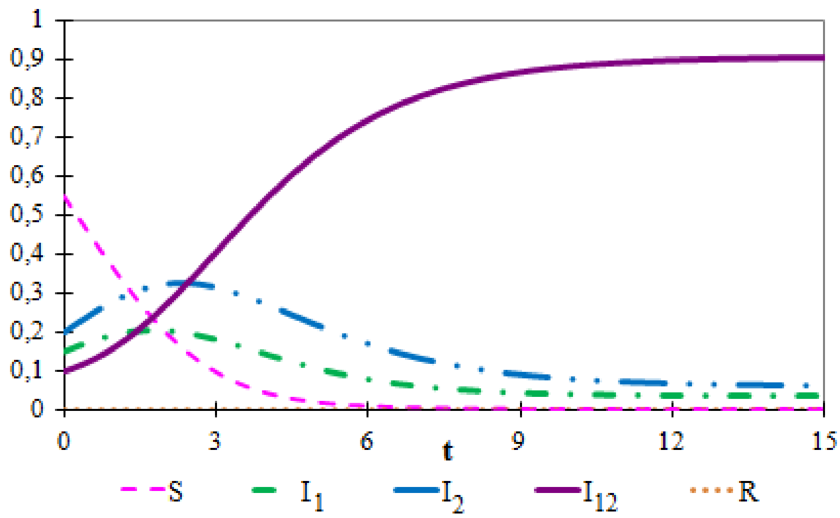


Figure 7: Uncontrolled SIR model.

System costs in uncontrolled case are equal to $J = 302.5$ monetary units (Fig. 8).

In contrast to previous case, we apply antivirus patches to heal in-

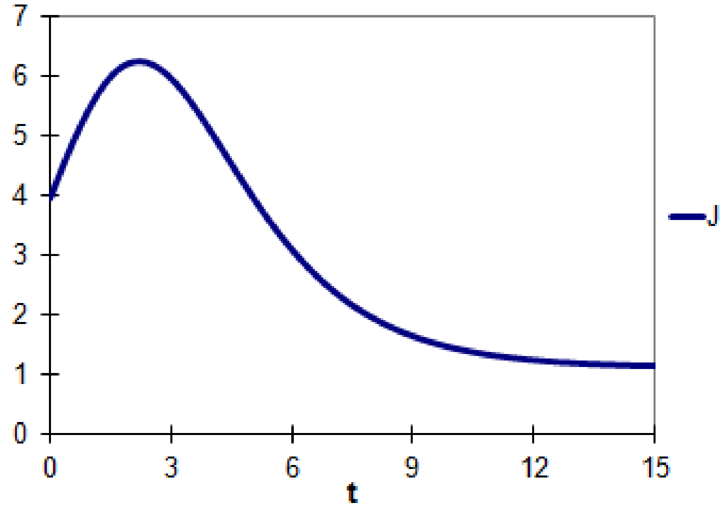


Figure 8: System costs in uncontrolled SIR model. ($J = 302.5$ monetary units)

fected nodes. Simulation represents that almost all nodes are in recovered state and there are no infected nodes (Fig. 9).

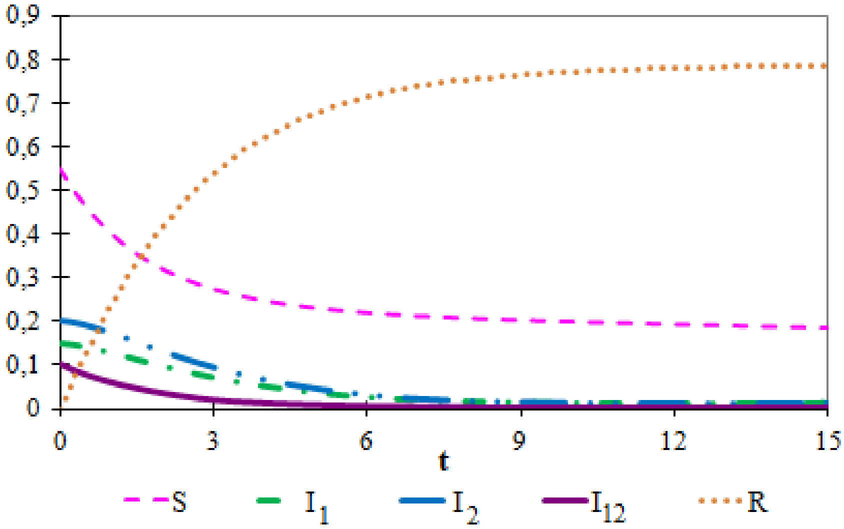


Figure 9: Controlled SIR model ($h_i$ - strictly convex).

Aggregated costs reduces to $J = 130.88$ monetary units (Fig. 10).

The next diagram (Fig. 11) illustrates the comparison between system coasts for uncontrolled and controlled system, which are equal to $J = 302.5$ and $J = 130.88$ monetary units respectively.

Structure of the optimal control is presented in the Fig. 12.

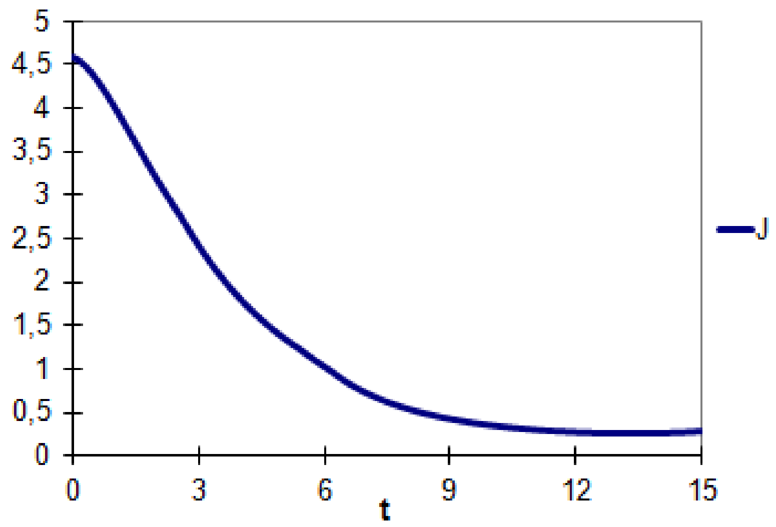


Figure 10: System costs in controlled SIR model. ($J = 130.88$ monetary units)

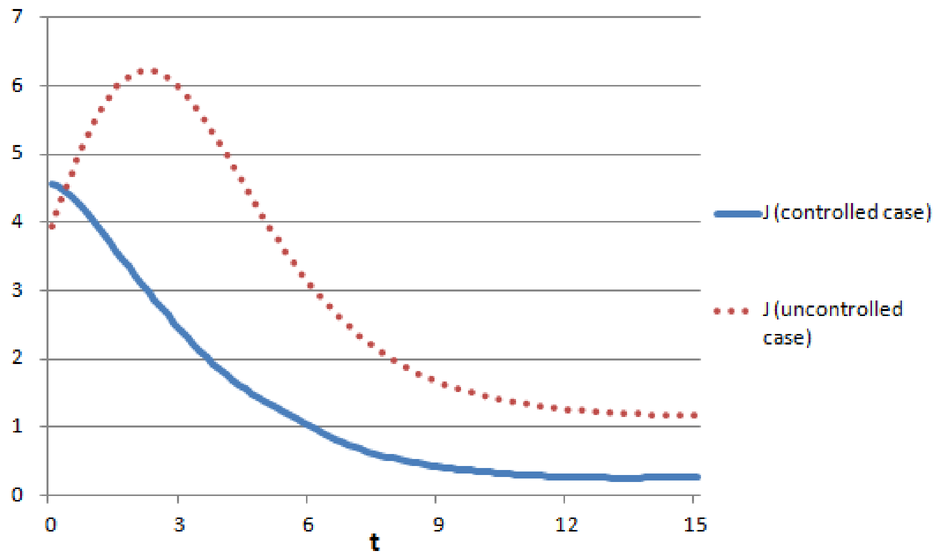


Figure 11: Comparison of system costs with and without spreading of antivirus.

Next, we change our functions $h_i$ from strictly convex to concave.

$$
\begin{aligned}
h_1(u_1(t)) &= 0.36 - (u_1(t) - 0.6)^2; \\
h_2(u_2(t)) &= 0.49 - (u_2(t) - 0.7)^2; \\
h_3(u_3(t)) &= 0.64 - (u_3(t) - 0.8)^2.
\end{aligned}
\tag{27}
$$

As in previous case, almost all of the nodes have moved into the recovered subgroup (Fig. 13).

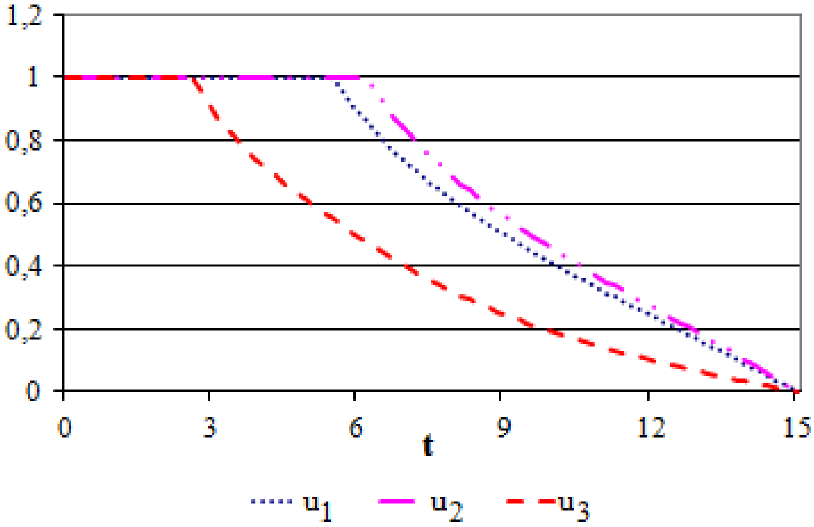


Figure 12: The optimal security response. Protection costs functions are strictly convex $h_1(u_1) = 0.35u_1^2$; $h_2(u_2) = 0.4u_2^2$; $h_3(u_3) = 0.5u_3^2$.

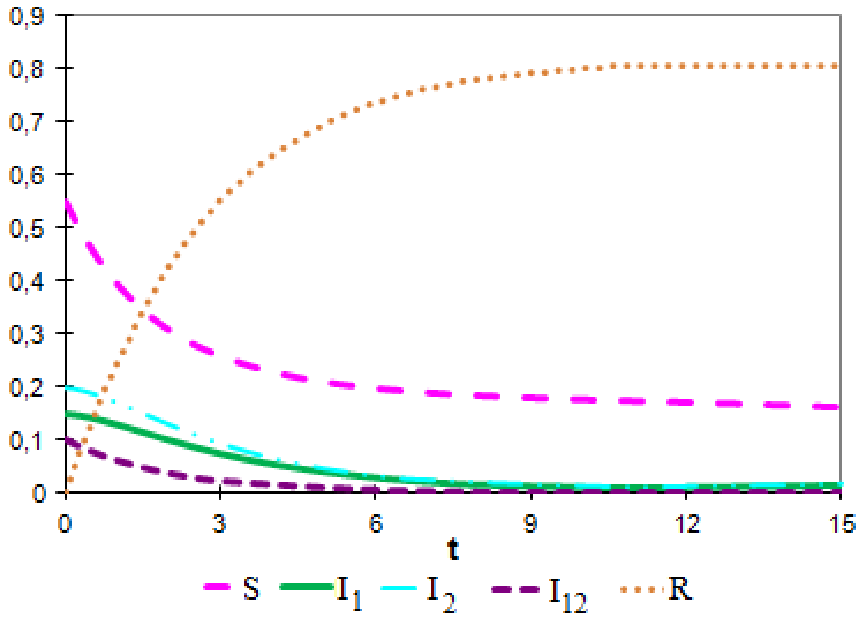


Figure 13: Controlled SIR model ($h_i$ - concave).

Aggregated system costs in case when $h_i$ are concave functions are $J = 193.17$ monetary units (Fig. 14).

Structure of the optimal control is presented in the Fig. 15.

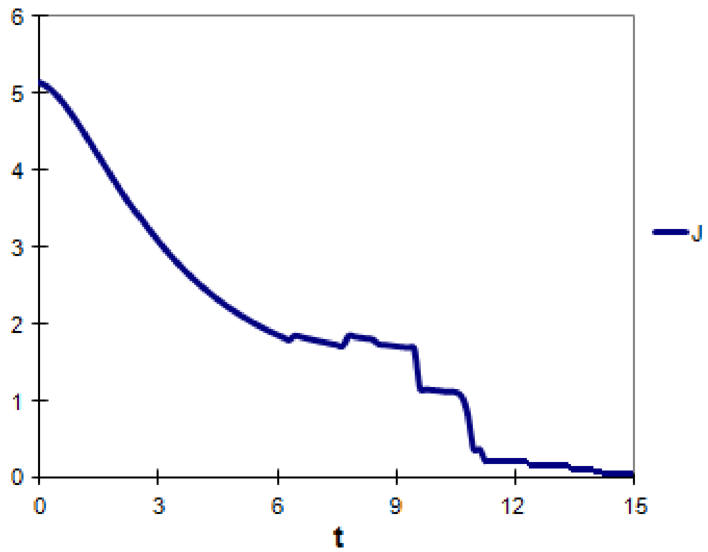


Figure 14: Aggregated costs in controlled SIR model. ($J = 193.17$ monetary units)

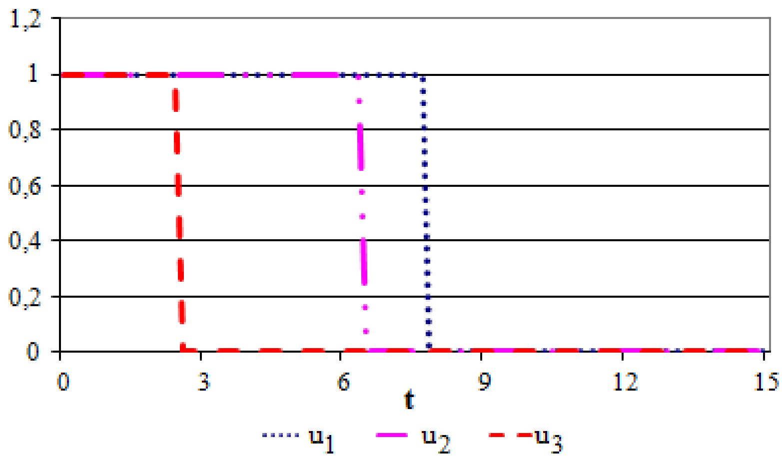


Figure 15: The optimal security response. Protection costs functions are strictly convex $h_1(u_1(t)) = 0.36-(u_1-0.6)^2$; $h_2(u_2(t)) = 0.49-(u_2-0.7)^2$; $h_3(u_3(t)) = 0.64-(u_3-0.8)^2$.

## 4.5 Stability analysis

In this section, we study the stability of the equilibrium states for the uncontrolled system, where $u_i = 0, i = 1, 2, 3$, [18, 22]. At an equilibrium point, all derivatives $\frac{dS}{dt} = \frac{dI_1}{dt} = \frac{dI_2}{dt} = \frac{dI_{12}}{dt} = \frac{dR}{dt} = 0$. We consider system

(6) without control. Therefore, the subgroup of Recovered $R$ is eliminated.

$$\begin{aligned} \dot{S} &= -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) + \sigma_1 I_1 + \sigma_2 I_2; \\ \dot{I}_1 &= \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_1 I_1 + \sigma_2 I_{12}; \\ \dot{I}_2 &= \beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2 + \sigma_1 I_{12}; \\ \dot{I}_{12} &= \varepsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - (\sigma_1 + \sigma_2) I_{12}; \end{aligned} \tag{28}$$

Thus we have three equilibrium points:

- $I_1 = I_2 = I_{12} = 0;$
- $I_1 = I_{12} = 0, \quad I_2 = 1 - \sigma_2/\beta_2;$
- $I_2 = I_{12} = 0, \quad I_1 = 1 - \sigma_1/\beta_1.$

Thus, we can find a simple equation for $I_{12}$:

$$\varepsilon I_1 I_2(\beta_1 + \beta_2) = (\sigma_1 + \sigma_2 + u_3 - \varepsilon\beta_2 I_1 - \varepsilon\beta_1 I_2) I_{12} \tag{29}$$

**Lemma 2** *The number of people infected by both virus 1 and virus 2 will obey the following equation:*

$$\begin{aligned} &\tfrac{dI_{12}}{dt} = \varepsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - (\sigma_1 + \sigma_2) I_{12} = 0; \\ &\textit{hence} \\ &I_{12} = (\varepsilon I_1 I_2(\beta_1 + \beta_2))/(\sigma_1 + \sigma_2 + u_3 - \varepsilon\beta_2 I_1 - \varepsilon\beta_1 I_2) \end{aligned} \tag{30}$$

From (28) we get:

$$\begin{aligned} &\tfrac{dS}{dt} = -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) + \sigma_1 I_1 + \sigma_2 I_2 = 0, \\ &\textit{hence} \\ &\beta_1 S(I_1 + I_{12}) + \beta_2 S(I_2 + I_{12}) = \sigma_1 I_1 + \sigma_2 I_2 \end{aligned} \tag{31}$$

Hence we have the expected three equilibrium points [5]:

- $S = 1$, $I_1 = I_2 = I_{12} = 0$;
- $I_1 = I_{12} = 0, \quad I_2 = 1 - \sigma_2/\beta_2, \quad S = \sigma_1/\beta_1$;
- $I_2 = I_{12} = 0, \quad I_1 = 1 - \sigma_1/\beta_1, \;\; S = \sigma_2/\beta_2$.

To provide the local stability analysis, we linearize the system (28). Together with the condition $S(t) = 1 - I_1(t) - I_2(t) - I_{12}(t)$, and we obtain Jacobi matrices for each equilibrium point, and use the Routh-Hurwitz criterion to determine the stability [10].

According to the criterion, we construct the following characteristic equation

$$z^4 + A_1 z^3 + A_2 z^2 + A_3 z^1 = 0 \tag{32}$$

and examine that all eigenvalues of the characteristic equation (32) has negative real part $Re(z) < 0$. It follows if the determinants of Hurwitz matrices are positive. $A_i$, $i = 1, 2, 3$ are constructed according to the criterion.

**I.** We will check the stability of equilibrium points using Hurwitz criterion. Consider the first state $(1, 0, 0, 0, 0)$. Hurwitz matrix for this case:

$$\begin{bmatrix} 0 & -\beta_1 + \sigma_1 & -\beta_2 + \sigma_2 & -\beta_1 - \beta_2 \\ 0 & \beta_1 - \sigma_1 & 0 & \beta_1 + \sigma_2 \\ 0 & 0 & \beta_2 - \sigma_2 & \beta_2 + \sigma_1 \\ 0 & 0 & 0 & -\sigma_1 - \sigma_2 \end{bmatrix} \tag{33}$$

Let's define $A_i,\ i = 1\dots 4$ as

$$
\begin{aligned}
A_1 &= -a_{11} - a_{22} - a_{33} - a_{44};\\
A_2 &= a_{11}(a_{22} + a_{33} + a_{44}) + a_{22}(a_{33} + a_{44}) + a_{33}a_{44}-\\
&\quad a_{23}a_{32};\\
A_3 &= a_{11}a_{23}a_{32} + a_{23}a_{32}a_{44} - a_{11}a_{22}a_{33}-\\
&\quad a_{11}a_{22}a_{44} - a_{11}a_{33}a_{44}-\\
&\quad a_{22}a_{33}a_{44} - a_{13}a_{21} - a_{32}\\
A_4 &= a_{11}a_{22}a_{33}a_{44} + a_{13}a_{21}a_{31}a_{44}-\\
&\quad a_{11}a_{23}a_{32}a_{44}
\end{aligned}
\tag{34}
$$

where $a_{ij}$ is an element of Hurwitz matrix. We calculate functions $G_i$ as follows:

$$
\begin{aligned}
G_1 &= A_1;\\
G_2 &= A_1A_2 - A_3 :\\
G_3 &= A_1A_2A_3 - (A_1)^2A_4 - (A_3)^2;\\
G_4 &= G_2(A_3A_4) - (A_1A_4)^2.
\end{aligned}
\tag{35}
$$

From calculation we receive that all $G_i$ are positive, then our system in first equilibrium point is stable.

**II.** Jacobi matrix at $E_1(S, I_1, I_2, I_2, R)$, $I_1 = I_{12} = 0,\ \ I_2 = 1-\sigma_2/\beta_2$, $S = \sigma_1/\beta_1$ is

$$
\begin{pmatrix}
-\beta_1\mu_1 & 0 & -\frac{\beta_2\sigma_1}{\beta_1} + \sigma_2 & -\eta_1\sigma_1\\
\beta_1\mu_1 & 0 & -\varepsilon\beta_2\mu_1 & \sigma_1 - \varepsilon\mu_1 + \sigma_2\\
0 & 0 & (\frac{\beta_2}{\beta_1} - 1)\sigma_1 - \varepsilon\beta_1\mu_1 & \eta_1\sigma_1\\
0 & 0 & \varepsilon(\beta_2 + \beta_1)\mu_1 & \varepsilon\beta_2\mu_1 - \sigma_1 - \sigma_2
\end{pmatrix}.
$$

To simplify the notation we denote $\eta_1 = (\frac{\beta_2}{\beta_1} + 1)$, $\mu_1 = (1 - \frac{\sigma_1}{\beta_1})$.

All determinants are positive $G_i > 0$, $i = 1, 2, 3, 4$, and then this equilibrium state is stable.

**III.** Jacobi matrix at $E_2(S, I_1, I_2, I_2, R)$, $I_2 = I_{12} = 0$, $\quad I_1 = 1 - \frac{\sigma_1}{\beta_1}$, $S = \frac{\sigma_1}{\beta_1}$ is

$$\begin{pmatrix} -\beta_2\mu_2 & -\frac{\beta_1\sigma_2}{\beta_2} - \sigma_1 & 0 & -\frac{\beta_1\sigma_2}{\beta_2} - \sigma_2 \\ 0 & (\frac{\beta_1}{\beta_2} - 1)\sigma_2 - \varepsilon\beta_2\mu_2 & 0 & -\frac{\beta_1\sigma_2}{\beta_2} + \sigma_2 \\ \beta_2\mu_2 & -\varepsilon\beta_1\mu_2 & 0 & \sigma_2 - \varepsilon\beta_1\mu_2 + \sigma_1 \\ 0 & \varepsilon(\beta_1 + \beta_2)\mu_2 & 0 & \varepsilon\beta_1\mu_2 - \sigma_1 - \sigma_2 \end{pmatrix},$$

where $\mu_2 = (1 - \frac{\sigma_2}{\beta_2})$, $\eta_2 = (\frac{\beta_1}{\beta_2} + 1)$.

From calculation we get that all determinants are positive $G_i > 0$, $i = 1, \ldots, 4$, and hence this equilibrium state is stable.

Using direct Lyapunov's method [13], we can check the global stability of $E_0$. We consider the Lyapunov function as follows:

$$L = S + I_1 + I_2 + I_{12} + R. \tag{36}$$

The derivative $\frac{dL}{dt}|_{(1)}$ under the system (28) is

$$\begin{aligned} \frac{dL}{dt} &= \dot{S} + \dot{I}_1 + \dot{I}_2 + \dot{I}_{12} + \dot{R} = -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) + \\ &\sigma_1 I_1 + \sigma_2 I_2 + \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \\ &\sigma_1 I_1 + \sigma_2 I_{12} - (1 - \varepsilon\beta_2)u_1 I_1 + \beta_2 S(I_2 + I_{12}) - \\ &\varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2 + \sigma_1 I_{12} - \\ &(1 - \varepsilon\beta_1)u_2 I_2 + \varepsilon\beta_1 I_2(I_1 + I_{12}) + \\ &\varepsilon\beta_2 I_1(I_2 + I_{12}) - (\sigma_1 + \sigma_2)I_{12} - u_3 I_{12} + \\ &(1 - \varepsilon\beta_2)u_1 I_1 + (1 - \varepsilon\beta_1)u_2 I_2 + u_3 I_{12} \leq 0. \end{aligned} \tag{37}$$

Hence according to Lyapunov's stability theorem we have that stationary state is stable but not asymptotically stable.

**Proposition 3** *Stationary states* $E_0(S, 0, 0, 0, R)$, $E_1(S, I_1, I_2, I_2, R)$, $E_2(S, I_1, I_2, I_2, R)$ *of the model (28) are stable.*

# 5 SIR model with impulse control

## 5.1 Mathematical model

Based on the model (1) we describe a model which include the system of differential equation to describe behavior of viruses and discrete system of impulses. As in previous research [9], [20] we consider a multi-virus case in which two forms of malicious software spread in the network with different speeds. Viruses can generate epidemic process in population periodically, for example epidemics of influenza can reach its peak two times during one epidemic period. Many examples of several waves of spreading the identical malicious software in computer and wireless networks are well-known [21]. In current work we formulate the conditions for eradication of epidemics of malwares for different cases of protection policies and compare costs and effectiveness of impulse actions and standard method of resistance.

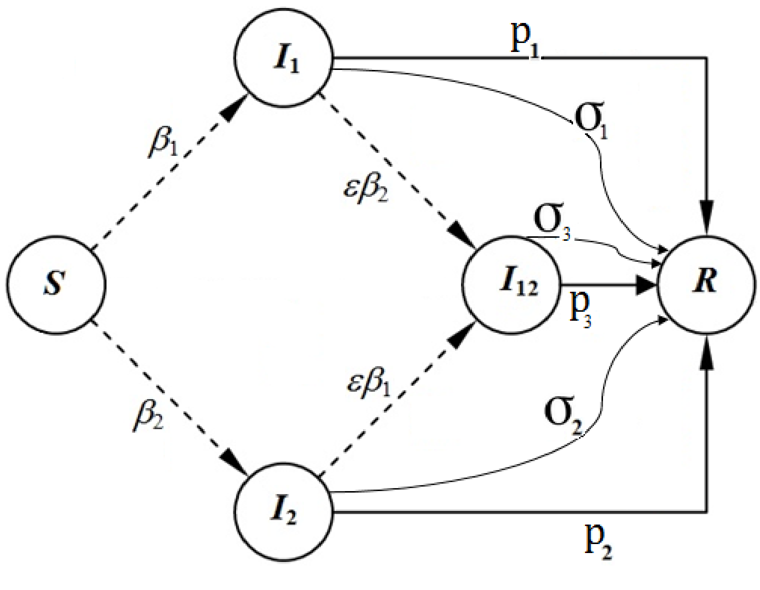


Figure 16: The scheme of the propagation of the system 38.

As in previous work [20] we model the epidemics of two forms of malware using a system of nonlinear differential equations (38).

$$
\begin{aligned}
&\dot{S} = -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}),\\
&\dot{I}_1 = \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_1 I_1,\\
&\dot{I}_2 = \beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2,\\
&\dot{I}_{12} = \epsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_3 I_{12},\\
&\dot{R} = \sigma_1 I_1 + \sigma_2 I_2 + \sigma_3 I_{12},
\end{aligned}
\tag{38}
$$

together with the condition

$$
S(t) + I_1(t) + I_2(t) + I_{12}(t) + R(t) = 1,\ t \in [0; T]. \tag{39}
$$

However, the analysis of the behavior of computer viruses has shown that a small shares of infected nodes might be survived in the Internet and if the local network has a connection with the Global Network then the epidemics resumes [21]. Then we can say the IT-security deals with the repeated waves of epidemics of malicious software. The iterative epidemic process can be formulated as a combined multi-virus model with series of impulses which control the shares of Infected nodes. As a basis for complex model can be used the system (38) or its closed modification i.e. Susceptible-Infected-Recovered-Susceptible (SIRS) model.

$$
\begin{aligned}
&\dot{S} = -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) + \mu R,\\
&\dot{I}_1 = \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_1 I_1,\\
&\dot{I}_2 = \beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2,\\
&\dot{I}_{12} = \epsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_3 I_{12},\\
&\dot{R} = \sigma_1 I_1 + \sigma_2 I_2 + \sigma_3 I_{12} - \mu R,\\
&\qquad S + I_1 + I_2 + I_{12} + R = 1.
\end{aligned}
\tag{40}
$$

Here $\mu$ is a rate at which recovered nodes becomes new susceptible nodes in the network. This SIRS model can be helpful for describing epidemics that are repeated periodically. Application of antivirus in impulse control form allows us to reduce the amount of infected nodes.

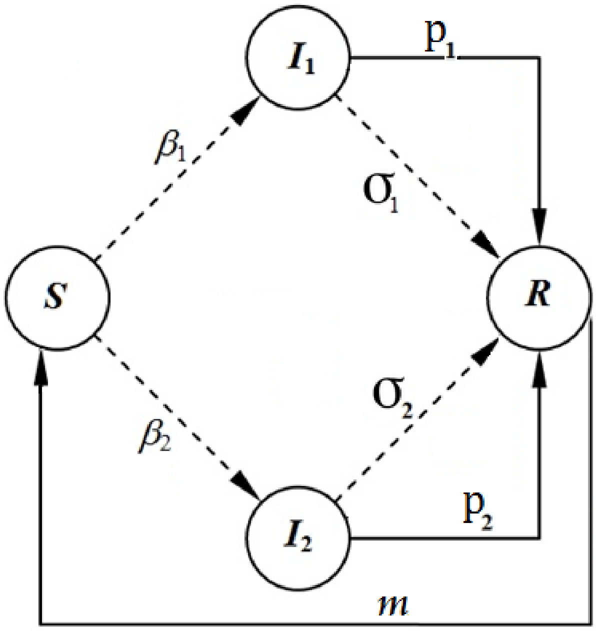


Figure 17: The scheme of the propagation of the system 40.

### 5.2 Impulse control and objective function

Together with initial systems (38), (40), which describe the behavior of viruses on the time intervals $(\tau_{p-1}, \tau_p)$, we formulate a model with application of impulse control strategies. To protect the network from repeated epidemics special patches can be applied as series of impulses at certain time moments. By using these patches as the control impulses at time moments $\tau_1, \ldots, \tau_k$ we receive the extended system of differential equation to describe the case of spreading of two malwares for all time period except the sequence of time moments $\tau_p$. States of the system after time moments $\tau_p$ are

$$
\begin{aligned}
&\dot{S}(\tau_p^+) = \dot{S}(\tau_p);\\
&\dot{I}_1(\tau_p^+) = \dot{I}_1(\tau_p) - u_1(\tau_p)I_1(\tau_p),\\
&\dot{I}_2(\tau_p^+) = \dot{I}_2(\tau_p) - u_2(\tau_p)I_2(\tau_p),\\
&\dot{I}_{12}(\tau_p^+) = \dot{I}_{12}(\tau_p) - u_3(\tau_p)I_3(\tau_p),\\
&\dot{R}(\tau_p^+) = \dot{R}(\tau_p) + u_1(\tau_p)I_1(\tau_p) + u_2(\tau_p)I_2(\tau_p) + u_3(\tau_p)I_{12}(\tau_p).
\end{aligned}
\tag{41}
$$

We define $u_i(\cdot)$, $i = 1, 2, 3$ as a control parameter which corresponds to the application of special antivirus patches at time moments $\tau_1, \ldots, \tau_k$. At each time moment $u_i$ is a fraction of treated nodes. Here $u_1 = (u_1^1, \ldots, u_1^k)$, $u_2 = (u_2^1, \ldots, u_2^k)$, $u_3 = (u_3^1, \ldots, u_3^k)$ are components of control vectors correspond to the set of time moments $\tau_1, \ldots, \tau_k$,
$u_1^j \in [0, \overline{u}_1^j]$, $u_2^j \in [0, \overline{u}_2^j]$, $u_3^j \in [0, \overline{u}_3^j]$, where $\overline{u}_1^j$, $\overline{u}_2^j$, $\overline{u}_3^j$ are maximum values of control.

The objective function of the combined system (41) is constructed to evaluate the aggregated costs on a time interval $[0, T]$ including the costs of control impulses. Aggregated costs for continuous systems (38) and (40) are defined as follows: at any given $t \neq \tau_p$, $p = 1, \ldots, k$, the overall system costs include infected costs $f_1(I_1(t))$, $f_2(I_2(t))$, $f_3(I_{12}(t))$. Functions $f_i$ are non-decreasing and twice-differentiable, such as $f_i(0) = 0$, $f_i(I_j(t)) > 0$ for $I_i(t) > 0$ for $t \in (\tau_{p-1}, \tau_p)$ for all $j$, $i = 1, 2, 3$. For system (41) we define infected costs as functions $h_i(u_i^p(\tau_p^+))$, $p = 1, \ldots, k$, where $h_i(u_i^p(\tau_p^+)) > 0$, $u_i^p(\tau_p^+) > 0$, $u_i \in [0, \overline{u}_i^p]$ for $i = 1, 2, 3$ which are generated by the consumption of resources for the application of antivirus or stationary security patches. Infected costs are consist of damages caused by viruses.

Aggregated system costs are defined as the functional:

$$J(u_1, u_2, u_3) = \int_0^T f_1(I_1(t)) + f_2(I_2(t)) + f_3(I_3(t))dt + \sum_{p=1}^{k} h_1(u_1(\tau_p^+)) + \sum_{p=1}^{k} h_2(u_2(\tau_p^+)) + \sum_{p=1}^{k} h_3(u_3(\tau_p^+)). \quad (42)$$

To illustrate the special properties of impulse treatment strategies we examine some simple examples of SIR and SIRS models which assess the process of propagation of malicious softwares. Pulse treatment is effective if we succeed to keep the number of Susceptible below a critical value which is generated by the envelope curves $L_j(t)$ (index $j$ corresponds to the enumeration of viruses) or certain critical values [1].

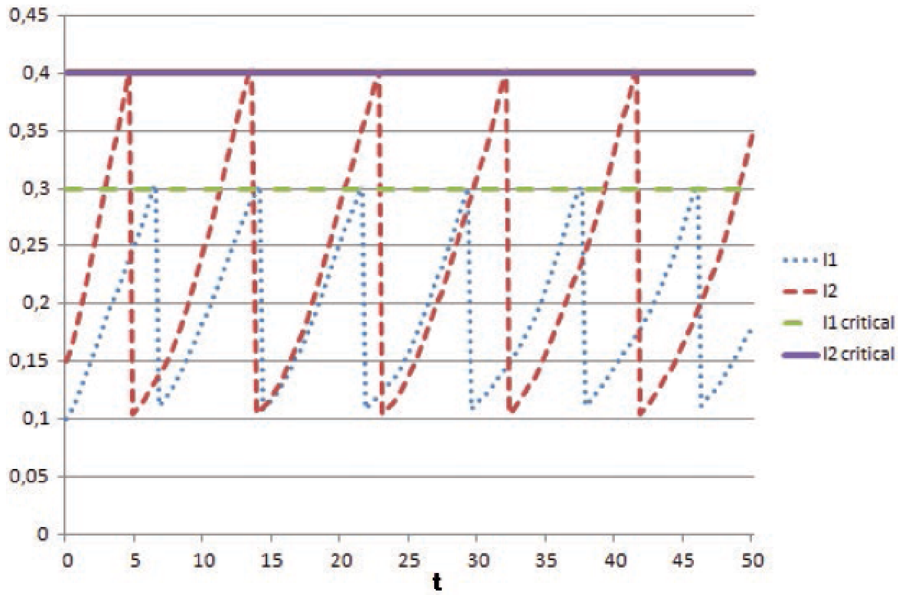


Figure 18: Behaviour of the infected nodes and critical values for infected in corresponding subgroups $I_1$, $I_2$. System parameters are: step $\delta = 0.25$, initial states $S(0) = 0.75$, $I_1(0) = 0.1$, $I_2(0) = 0.15$, $I_12 = 0$, $\varepsilon = 0.5$, $\beta_1 = 0.35$, $\beta_1 = 0.4$, $\sigma_1 = 0.002$, $\sigma_2 = 0.004$, $\mu = 0.4$. Maximal fraction of treated nodes are: $\overline{u}_1 = 0.2$, $\overline{u}_2 = 0.3$.

Figures 18, 19 illustrates the SIRS model (40) in case of impulse treatment spreading which preserve a number of infected of viruses $V_1$ and $V_2$ below critical values.

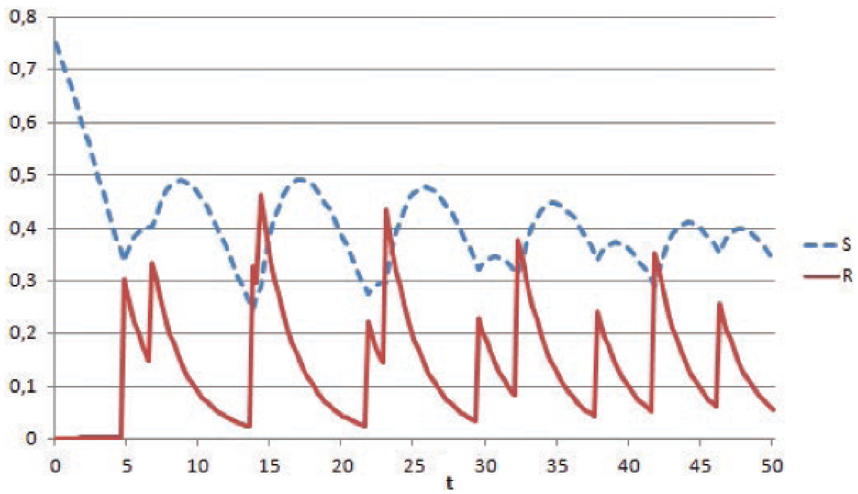


Figure 19: Fractions of $S$ and $R$ in SIRS model.

## 5.3 Structure of the optimal control

We construct the envelope based on the concept of basic reproduction number which is defined as the expected number of new infections from one infected individual in a fully susceptible population through the entire duration of the infectious period [1], [2]. This metric is significant as well it helps to determine whether or not an infectious disease can spread through a population. The replacement number is defined as the expected number of secondary infections that one infected person would produce through the entire duration of the infectious period [15]. Reproduction numbers for virus $V_1$, $V_2$ and both viruses simultaneously are defined respectively:

$$R_{01} = \frac{\beta_1}{\sigma_1}, R_{02} = \frac{\beta_2}{\sigma_2},$$
$$R_{03} = \frac{\varepsilon(\beta_1 + \beta_2)}{\sigma_3}. \tag{43}$$

Then envelop curves we define as follows on the interval $[0, T]$:

$$\begin{aligned} L_1(t) &= R_{01}S(t), \\ L_2(t) &= R_{02}S(t), \\ L_3(t) &= R_{03}S(t). \end{aligned} \tag{44}$$

Pulse treatment is effective to keep the the stable distribution of Susceptible, Infected and Recovered in population below envelope curves that define critical values. From that reasons according to system (38) and from the definition of envelops $L_j(t)$ we have:

$$\begin{aligned} I_1(t) &< \frac{\beta_1}{\sigma_1}S(t), \\ I_2(t) &< \frac{\beta_2}{\sigma_2}S(t), \\ I_{12}(t) &< \frac{\varepsilon(\beta_1+\beta_2)}{\sigma_3}S(t). \end{aligned} \tag{45}$$

In current model, we define the series of time moments $\tau_p, p = 1, \ldots, k$ at which we apply the special patches to treat the infected nodes. As functions $I_j(t)$ are increasing for all $j$ (in the case without any control measures), we turn impulse of control on for the first time when $I_j(t) = L_j(t)$, for all $j$. Series of control impulses switch on at time moments $\tau_p$, $p = 1, \ldots, k$ according to following conditions:

$$u_1(\tau_p) = \begin{cases} 0, & I_1(t) < L_1(t), \\ \overline{u}_1^p, & I_1(t) \geq L_1(t), \end{cases} \tag{46}$$

$$u_2(\tau_p) = \begin{cases} 0, & I_2(t) < L_2(t), \\ \overline{u}_2^p, & I_2(t) \geq L_2(t), \end{cases} \tag{47}$$

$$u_3(\tau_p) = \begin{cases} 0, & I_{12}(t) < L_3(t), \\ \overline{u}_3^p, & I_{12}(t) \geq L_3(t). \end{cases} \quad (48)$$

Here $\overline{u}_i^p \in [0,1]$, $i = 1,2,3$, $p = 1,\dots,k$ is defined as maximal value of applied control impulses.

### 5.4 Numerical simulation

In this section we present numerical simulations that corroborate our results. In the example we set following values of parameters: iteration step is $\delta = 0.25$, initial fractions of nodes are $S(0) = 0.45$, $I_1(0) = 0.2$, $I_2(0) = 0.3$, $I_{12}(0) = 0.05$, $R(0) = 0$. Intensive rates of transition from susceptible to infected and recovered are $\beta_1 = 0.35$, $\beta_2 = 0.4$ and $\gamma = 0.1$ respectively. Virus interaction rate is $\varepsilon = 0.5$. Intensive rates of transition from infected to recovered are $\sigma_1 = 0.05$, $\sigma_2 = 0.04$ and $\sigma_3 = 0.03$. We suppose that epidemic duration is $T = 25$ time units. Infected costs for infected nodes are $f_1(I_1(t)) = 5I_1(t)$, $f_2(I_2(t)) = 6I_2(t)$ and $f_3(I_{12}(t)) = 10I_{12}(t)$.

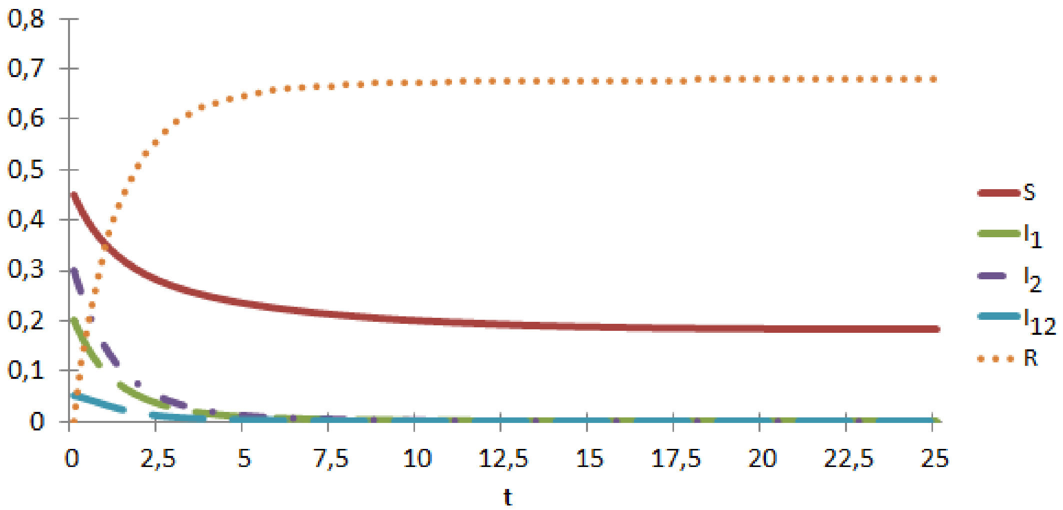


Figure 20: Multi-virus SIR model with continuous control.

In continuous control case, we use as protection costs following functions: $h_1(u_1(t)) = 1.3u_1(t)^2$, $h_2(u_2(t)) = 1.5u_2^2(t)$, $h_3(u_3(t)) = 2u_3^2(t)$ and $h_4(u_4(t)) = u_4^2(t)$, aggregated system costs is $J = 65.36$ monetary units (Fig. 22). At the time $T = 25$ we can see that there are no infected hosts (Fig 20). The distribution of nodes in population shows that the most of nodes have been transferred into Recovered subgroup and a small shares of nodes is in a susceptible subgroup, which have no influence on the epidemics. Optimal control for this case is shown in the Fig 21.

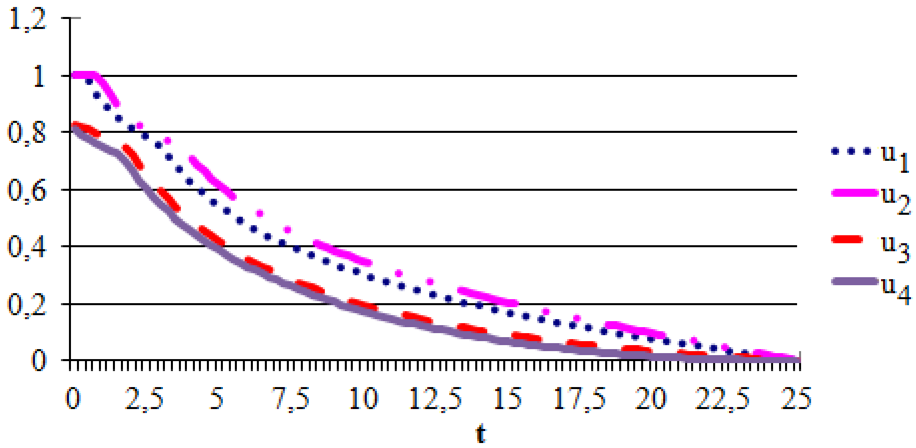


Figure 21: Optimal control for continuous case.

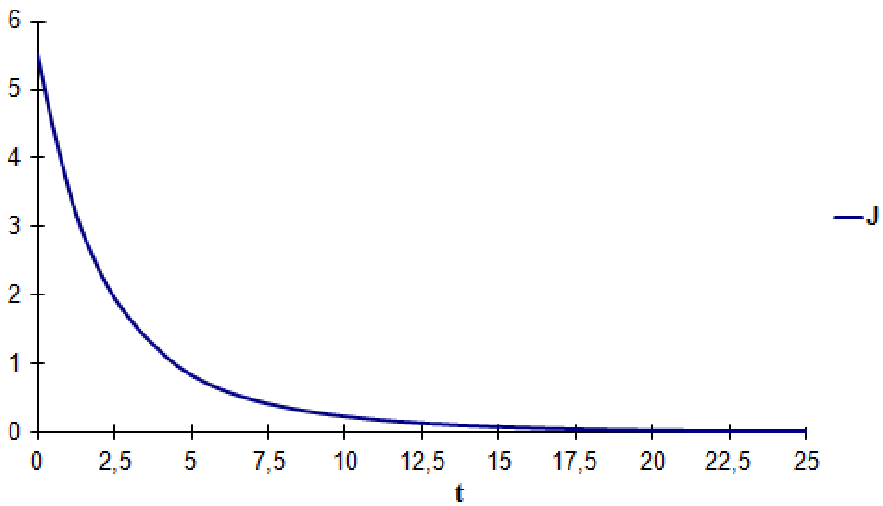


Figure 22: System costs for continuous case ($J = 65.36$).

In case with impulse control, protection costs are $h_1(u_1(\tau_p)) = 10u_1(\tau_p)$,

$h_2(u_2(\tau_p)) = 15u_2(\tau_p)$, $h_3(u_3(\tau_p)) = 18u_3(\tau_p)$. At the end of time interval $T = 25$ we can see that the most of infected hosts become recovered (Fig. 23) but in contract to continuous case aggregated costs is bigger. This fact appears from the formulation of combined problem and means that in continuous case we treat the infected nodes only once while in impulse case we deal with repeated procedures.

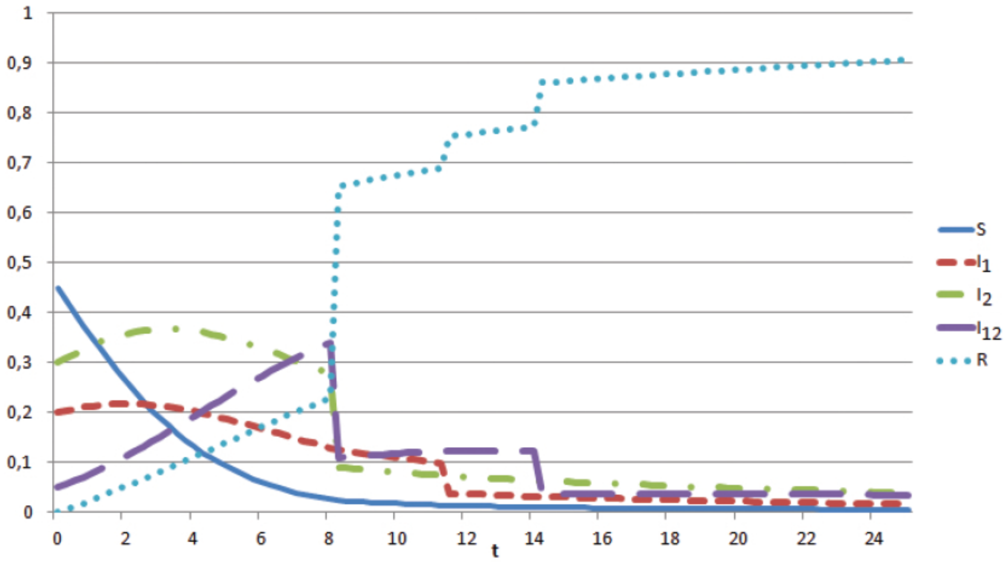


Figure 23: Multi-virus SIR model with series of impulses.

System costs $J = 252.51$ monetary units are illustrated in the Fig.24.

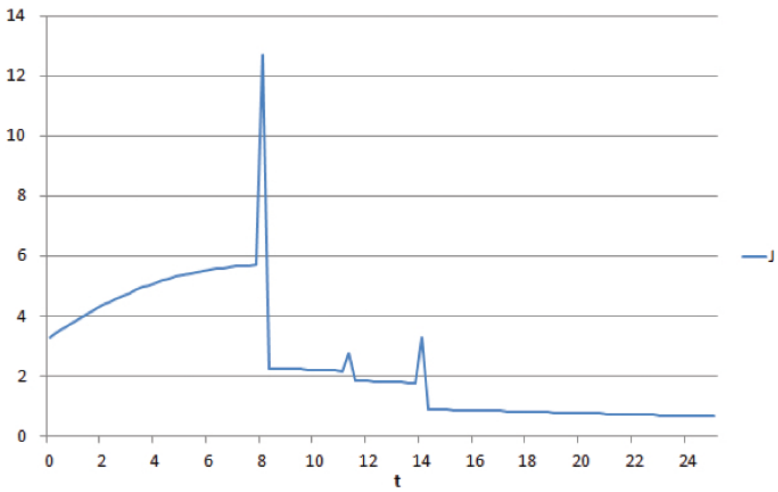


Figure 24: Aggregated system costs for the combined system with impulse controls.

In the next diagram (Fig. 25) presented curves $L_i,\ i = 1 \ldots 3$ and fractions of infected nodes $I_1$, $I_2$ and $I_{12}$. We apply pulse treatment at time moments $\tau_p$ when $I_i(t) = L_i(t)$.

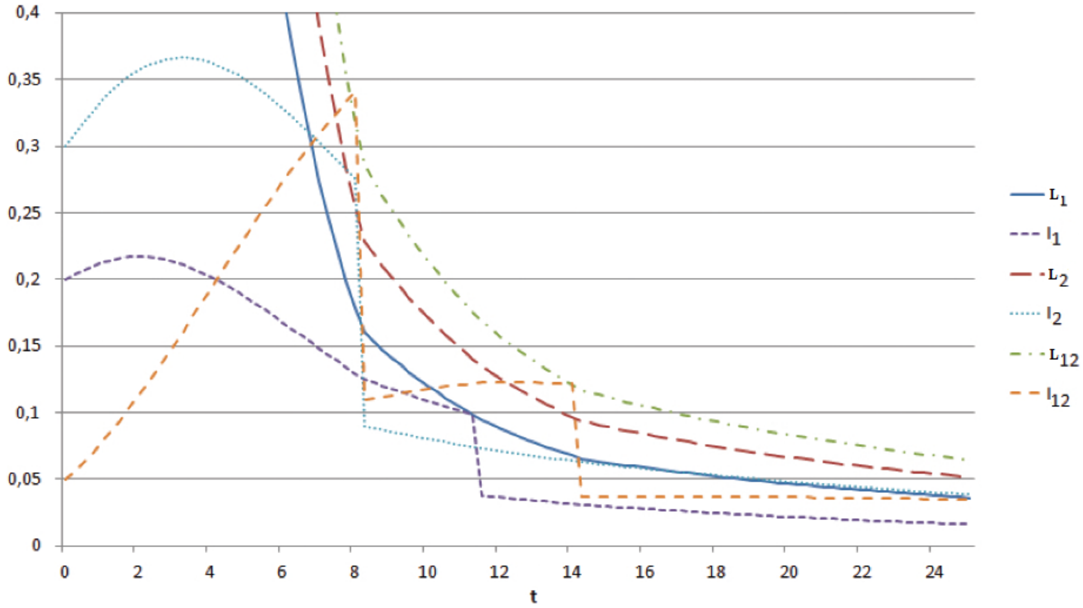


Figure 25: Envelope curves $L_i$ for the proportion of infected nodes $I_i$.

# 6 Conclusion

In this work we have studied four modifications of multi-viruses Susceptible-Infected-Recovered and Susceptible-Infected-Recovered-Susceptible models which describe propagation of two types of malware in computer networks. These extended models take into account the coexistence of heterogeneous malware and the exposure of computer systems to multiple vulnerabilities.

Firstly, we study case when we spread antivirus in continuous form. We have formulated an optimal control problem to study the tradeoffs between security risks and the control investment. By using Pontryagin's maximum principle, we have obtained different control policies structure that minimize the aggregated cost. The structure of the control depends on properties of costs functions. Numerical simulations were performed using a specially written procedures.

In second part, we have reformulated the SIR and SIRS models under the impulse control. In contrast to the previous statement we analyze the conditions for application the series of impulses that protect the network during periodic waves of epidemics of malwares instead of continuous control. This case is also supported with numerical simulations.

In subsection 5.4, according to the numerical simulation for the set of initial data and parameters, it has been shown that the aggregated system costs in continuous case are less than in impulse treatment case. Due to selected conditions of applying treatment, protection strategies in impulse form starts to treat our system not immediately but after some time from the beginning of epidemics, while protection strategies of continuous form of spreading starts to heal infected nodes immediately at time moment

$t = 0$. As a result malicious software has time to infect susceptible nodes in system which provokes additional system costs.

In further researches, continuous control SIR and SIRS models can be reformulated as impulse optimal control problem. These models will cover the studying of stability and finding of the optimal control in model with pulse treatment.

## References


1. Agur Z., Cojocaru L., Mazor G., Anderson R. M., Danon Y. L. Pulse mass measles vaccination across age cohorts // Proceedings of the National Academy of Sciences of the United States of America, 1993. Vol. 90, P. 11698—11702.

2. Agur Z., Stone L., Shulgin B. Pulse Vaccination Strategy in the SIR Epidemic Model // Bulletin of Mathematical Biology, 1998. Vol. 60, P. 1123—1148.

3. Altman E., Khouzani M., Sarkar S. Optimal control of epidemic evolution // Proceedings of INFOCOM, 2011. P. 1683—1691

4. Behncke H. Optimal control of deterministic epidemics // Optimal Control Applications and Methods, 2000. Vol. 21, P. 269—285.

5. Beutel A., Faloutsos C., Prakash B., Rosenfeld R. Interacting Viruses in Networks: Can Both Survive? // KDD-2012, 2012. P.426—434.

6. Chitnis N. Introduction to Mathematical Epidemiology: The Basic Reproductive Number // Universitat Basel Mathematisches Institut, 2011.

7. Evans A. S., Kaslow R. A. Viral Infections of Humans: Epidemiology and Control // New York: Springer, 1997.

8. Funke B., Tortora G., Gerard J., Christine L. Case. Microbiology: An Introduction. San Francisco, CA: Benjamin Cummings, 2004. P. 848.

9. Gubar E., Zhu Q. Optimal Control of Influenza Epidemic Model with Virus Mutations // 12th Biannual European Control Conference, IEEE Control Systems Society, 2012. P. 3125—3130.

10. Hurwitz A. Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt // Mathematische Annalen, 1895. Vol. 46, P. 273—-284.

11. Kermack W., Mc Kendrick A. A contribution to themathematical theory of epidemics // Proceedings of the Royal Society, 1927. P. 700—721.

12. Khouzani M. H. R., Sarkar S., Altman E. Maximum damage malware attack mobile wireless networks // Proc. of the 29th International Conference on Computer Communications (INFOCOM), 2010. P. 749—757.

13. Lyapunov A. M. The General Problem of the Stability of Motion. Doctoral dissertation, Univ. Kharkov, 1892. P. 250.

14. Moore D., Shannon C. Code-Red: a Case Study on the Spread and Victims of an Internet Worm // Proceedings of the 2002 ACM SICGOMM Internet Measurement Workshop, 2002. PP. 273—284.

15. Nakul C. Introduction to Mathematical Epidemiology: The Basic Reproductive Number // Einfuhrung in die Mathematische Epidemiologie, 2011. P. 5.

16. Pontryagin L., Boltyanskii V., Gamkrelidze R., Mishchenko E. The Mathematical Theory of Optimal Processes, Russia: Interscience, 1962. P. 360.

17. Rohloff K. R., Başar T. Deterministic and stochastic models for the detection of random constant scanning worms // ACM Transactions on Modeling and Computer Simulation (ACM TOMACS), 2008. Vol. 18, PP. 8:1—8:24.

18. Sharma S., Samanta G. P. Stability analysis and optimal control of an epidemic model with vaccination // International Journal of Biomathematics, 2015. Vol. 8, No. 3, P. 28.

19. Taylor P. D., Jonker L. B. Evolutionarily stable strategies and game dynamics // Mathematical Biosciences, 1978. Vol. 40, PP.145—156.

20. Taynitskiy V.A., Gubar E.A., Zhitkova E.M. Structure of optimal control in the model of propagation of two malicious softwares // International conference "Stability and Control Processes" in memory of V.I. Zubov (SCP), 2015. PP. 261-264.

21. Pastor-Satorras R., Vespignani A. Epidemic Spreading in Scale-Free Networks // Physical Review Letters, 2001. Vol. 86, PP.3200—3203.

22. Weibull J. Evolutionary Game Theory. The M.I.T. Press, Cambridge, MA, 1995. P. 265.

23. Wickwire K. H. A note on the optimal control of carrier-borne epidemics // Journal of Applied Probability, 1975. Vol. 12, PP. 565—568.

24. Wu Q., Small M., Liu H.. Superinfection Behaviors on Scale-Free Networks with Competing Strains // Journal of Nonlinear Science, 2013. Vol. 23, PP. 113—127.

# Appendix 1

Now we have to prove that $(\lambda_R(t) - \lambda_S(t)) \geq 0$. Analogously we assume:

$$\begin{aligned}
&(\lambda_R(t^*) - \lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_R(t^*) - \lambda_{I_2}(t^*)) \geq 0, \\
&(\lambda_R(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_R(t^*) - \lambda_S(t^*)) = 0, \\
&(\lambda_S(t^*) - \lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_S(t^*) - \lambda_{I_2}(t^*)) \geq 0 \\
&(\lambda_{I_1}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_{I_2}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0.
\end{aligned} \tag{49}$$

From (8) we obtain:

$$\begin{aligned}
\dot{\lambda}_R(t^*) - \dot{\lambda}_S(t^*) = -\beta_1(\lambda_S - \lambda_{I_1})(I_1 + I_{12}) - \\
-\beta_2(\lambda_S - \lambda_{I_2})(I_2 + I_{12}) + \gamma u_4(\lambda_R - \lambda_S).
\end{aligned} \tag{50}$$

According to assumption (81), the difference $\dot{\lambda}_R(t^{*+}) - \dot{\lambda}_S(t^{*+}) \leq 0$ is negative at the time moment $t^{*+}$, then function $(\lambda_R(t) - \lambda_S(t))$ decrease. It contradicts Property 1. At time moment $t = T$ $(\lambda_R(t) - \lambda_S(t))$ is equal to zero and on the whole interval it is decreasing. We proved that $(\lambda_R(t) - \lambda_S(t)) \geq 0$ for $t \in [0; T]$.

Analogously we can prove that $(\lambda_R(t) - \lambda_{I_2}(t)) \geq 0$ and $(\lambda_R(t) - \lambda_{I_{12}}(t)) \geq 0$.

We get that on the interval $t \in [0; T]$:

$$\begin{aligned}
&(\lambda_R(t) - \lambda_{I_1}(t)) \geq 0; \quad (\lambda_R(t) - \lambda_{I_2}(t)) \geq 0; \\
&(\lambda_R(t) - \lambda_{I_{12}}(t)) \geq 0; \quad (\lambda_R(t) - \lambda_S(t)) \geq 0.
\end{aligned} \tag{51}$$

# Appendix 2

**Proof of Proposition 2.**

We use Pontryagin's maximum principle [4], [16] to find the optimal control $u = (u_1, u_2, u_3)$ which yields the minimum solution to the functional (22) for the problem described above. Consider the Hamiltonian

$$\begin{aligned} H = &-(f_1(I_1) + f_2(I_2) + f_3(I_{12}) + h_1(u_1)+ \\ &h_2(u_2) + h_3(u_3) - g(R)) + \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S)+ \\ &\beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) + \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1})+ \\ &\varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) + \sigma_1(\lambda_S - \lambda_{I_1})I_1+ \\ &\sigma_2(\lambda_S - \lambda_{I_2})I_2 + \sigma_1(\lambda_{I_2} - \lambda_{I_{12}})I_{12}+ \\ &\sigma_2(\lambda_{I_1} - \lambda_{I_{12}})I_{12} + (1 - \varepsilon\beta_2)u_1 I_1(\lambda_R - \lambda_{I_1})+ \\ &(1 - \varepsilon\beta_1)u_2 I_2(\lambda_R - \lambda_{I_2}) + u_3(\lambda_R - \lambda_{I_{12}})I_{12}. \end{aligned} \tag{52}$$

Let be $\lambda_0 = 1$. We construct the adjoint system as follows:

$$\frac{d\lambda_S}{dt} = \beta_1(\lambda_S - \lambda_{I_1})(I_1 + I_{12}) + \beta_2(\lambda_S - \lambda_{I_2})(I_2 + I_{12});$$

$$\begin{aligned} \frac{d\lambda_{I_1}}{dt} &= f_1' + \beta_1 S(\lambda_S - \lambda_{I_1}) + \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \sigma_1(\lambda_{I_1} - \lambda_S)+ \\ &\quad \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) + u_1(1 - \varepsilon\beta_2)(\lambda_{I_1} - \lambda_R); \\ \frac{d\lambda_{I_2}}{dt} &= f_2' + \beta_2 S(\lambda_S - \lambda_{I_2}) + \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \sigma_2(\lambda_{I_2} - \lambda_S)+ \\ &\quad \varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}}) + u_2(1 - \varepsilon\beta_1)(\lambda_{I_2} - \lambda_R); \\ \frac{d\lambda_{I_{12}}}{dt} &= f_3' + \beta_1 S(\lambda_S - \lambda_{I_1}) + \beta_2 S(\lambda_S - \lambda_{I_2}) + \sigma_1(\lambda_{I_{12}} - \lambda_{I_2})+ \\ &\quad \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \sigma_2(\lambda_{I_{12}} - \lambda_{I_1})+ \\ &\quad +u_3(\lambda_{I_{12}} - \lambda_R); \\ \frac{d\lambda_R}{dt} &= -g'(R), \end{aligned} \tag{53}$$

together with condition $R(t) = 1 - S(t) - I_1(t) - I_2(t) - I_{12}(t)$, along with the transversality conditions

$$\lambda_S(T) = \lambda_{I_1}(T) = \lambda_{I_2}(T) = \lambda_{I_{12}}(T) = \lambda_R(T) = 0. \tag{54}$$

According to Pontryagin's maximum principle, there exist continuous and piecewise continuously differentiable co-state functions $\lambda_i$ that at every point $t \in [0; T]$ where $u_1$, $u_2$ and $u_3$ are continuous, satisfy (53) and (54).

In addition, we have $\overline{\lambda}(t) = (\lambda_S(t), \lambda_{I_1}(t), \lambda_{I_2}(t), \lambda_{I_{12}}(t), \lambda_R(t))$, and

$$(u_1, u_2, u_3) \in \arg\max_{\underline{u_1}, \underline{u_2}, \underline{u_3}} H(\overline{\lambda}, S, I_1, I_2, I_{12}, R, \underline{u_1}, \underline{u_2}, \underline{u_3}). \tag{55}$$

Here, $\underline{u_1}, \underline{u_2}, \underline{u_3}$ are feasible controls.

Rewrite Hamiltonian in terms of function $\psi$, and we obtain

$$\begin{aligned} H = & -(f_1(I_1) + f_2(I_2) + f_3(I_{12}) - g(R)) + \\ & \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) + \beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) + \\ & \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) + \varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) + \\ & \sigma_1(\lambda_S - \lambda_{I_1})I_1 + \sigma_2(\lambda_S - \lambda_{I_2})I_2 + \sigma_1(\lambda_{I_2} - \lambda_{I_{12}})I_{12} + \\ & \sigma_2(\lambda_{I_1} - \lambda_{I_{12}})I_{12} + (-h_1(u_1) + u_1\psi_1) + \\ & (-h_2(u_2) + u_2\psi_2) + (-h_3(u_3) + u_3\psi_3). \end{aligned} \tag{56}$$

For any admissible control $u_1$, $u_2$ and $u_3$ and according to (56) for all $t \in [0, T]$ we arrive at

$$\begin{aligned} & (-h_1(u_1) + u_1\psi_1 - h_2(u_2) + u_2\psi_2 - h_3(u_3) + u_3\psi_3) \geq \\ & (-h_1(\underline{u_1}) + \underline{u_1}\psi_1 - h_2(\underline{u_2}) + \underline{u_2}\psi_2 - h_3(\underline{u_3}) + \underline{u_3}\psi_3). \end{aligned} \tag{57}$$

Hence we obtain

$$(u_1(t), u_2(t), u_3(t)) \in arg \max_{x,y,z\in[0,1]}(-h_1(x) + x\psi_1 - \\ -h_2(y) + y\psi_2 - h_3(z) + z\psi_3). \tag{58}$$

Then,

$$\max_{u_1,u_2,u_3}(-h_1(u_1) + u_1\psi_1 - h_2(u_2) + u_2\psi_2 - h_3(u_3) + u_3\psi_3) \\ = \max_{u_1}(-h_1(u_1) + u_1\psi_1) + \max_{u_2}(-h_2(u_2) + u_2\psi_2) \\ + \max_{u_3}(-h_3(u_3) + u_3\psi_3). \tag{59}$$

According to the algorithm of Pontryagin's maximum principle to determine the optimal control structure, we consider derivatives $\frac{\partial H}{\partial u}$:

$$\frac{\partial H}{\partial u_i} = -\dot{h}_i(u_i) + \psi_i = 0, \tag{60}$$

As $h_i(u_i)$, $i = 1, 2, 3$ are increasing functions and $I_i \geq 0$ then Hamiltonian reaches its maximum if $\dot{h}_i(u_i) = \psi_i \geq 0$.

Let us calculate time derivatives of functions $\psi_i$:

$$\dot{\psi}_1 = (1 - \varepsilon\beta_2)[(\lambda_R - \lambda_{I_1})\dot{I}_1 + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_1})I_1] = \\ (1 - \varepsilon\beta_2)[(\lambda_R - \lambda_{I_1})(\beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \\ \sigma_1 I_1 + \sigma_2 I_{12} - (1 - \varepsilon\beta_2)u_1 I_1) + \\ (-g'(R) - (f_1' + \beta_1 S(\lambda_S - \lambda_{I_1}) + \\ \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) + \\ \sigma_1(\lambda_{I_1} - \lambda_S) + u_1(1 - \varepsilon\beta_2)(\lambda_{I_1} - \lambda_R)))I_1]. \tag{61}$$

$$\dot{\psi}_2 = (1-\varepsilon\beta_1)[(\lambda_R - \lambda_{I_2})\dot{I}_2 + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_2})I_2] =$$
$$(1-\varepsilon\beta_2)[(\lambda_R - \lambda_{I_2})(\beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) -$$
$$\sigma_2 I_2 + \sigma_1 I_{12} - (1-\varepsilon\beta_1)u_2 I_2) +$$
$$(-g'(R) - (f_2' + \beta_2 S(\lambda_S - \lambda_{I_2}) + \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) +$$
$$\sigma_2(\lambda_{I_2} - \lambda_S) +$$
$$\varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}}) + u_2(1-\varepsilon\beta_1)(\lambda_{I_2} - \lambda_R)))I_2].$$
$$\dot{\psi}_3 = [(\lambda_R - \lambda_{I_{12}})\dot{I}_{12} + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_{12}})I_{12}] =$$
$$(\lambda_R - \lambda_{I_{12}})(\varepsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) -$$
$$(\sigma_1 + \sigma_2)I_{12} - u_3 I_{12}) +$$
$$(-g'(R) - (f_3' + \beta_1 S(\lambda_S - \lambda_{I_1}) + \beta_2 S(\lambda_S - \lambda_{I_2}) +$$
$$\varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \sigma_2(\lambda_{I_{12}} - \lambda_{I_1}) +$$
$$\sigma_1(\lambda_{I_{12}} - \lambda_{I_2}) + u_3(\lambda_{I_{12}} - \lambda_R)))I_{12}.$$

After arrangement:

$$\begin{aligned}
\dot{\psi}_1 &= (1-\varepsilon\beta_2)[(\lambda_R - \lambda_{I_1})\dot{I}_1 + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_1})I_1] \\
&(1-\varepsilon\beta_2)[I_1(-f_1' - g' + \beta_1 S(\lambda_R - \lambda_S) + \\
&\varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_R) + \varepsilon\beta_1 I_2(\lambda_{I_{12}} - \lambda_{I_2}) + \\
&\sigma_1(\lambda_S - \lambda_R)) + I_{12}(\lambda_R - \lambda_{I_1})(\beta_1 S + \sigma_2)] \\
\dot{\psi}_2 &= (1-\varepsilon\beta_1)[(\lambda_R - \lambda_{I_2})\dot{I}_2 + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_2})I_2] = \\
&(1-\varepsilon\beta_1)[I_2(-f_2' - g' + \beta_2 S(\lambda_R - \lambda_S) + \\
&\varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_R) + \varepsilon\beta_2 I_1(\lambda_{I_{12}} - \lambda_{I_1}) + \\
&\sigma_2(\lambda_S - \lambda_R)) + I_{12}(\lambda_R - \lambda_{I_2})(\beta_2 S + \sigma_1)] \\
\dot{\psi}_3 &= [(\lambda_R - \lambda_{I_{12}})\dot{I}_{12} + (\dot{\lambda}_{I_R} - \dot{\lambda}_{I_{12}})I_{12}] = \\
&I_{12}[-f_3' - g' + (\lambda_R - \lambda_{I_2})(\varepsilon\beta_1 I_2 - \sigma_1) + \\
&(\lambda_R - \lambda_{I_1})(\varepsilon\beta_2 I_1 - \sigma_2) + S(\beta_1(\lambda_{I_1} - \lambda_S) + \\
&\beta_2(\lambda_{I_2} - \lambda_S))] + \varepsilon I_1 I_2(\beta_1(\lambda_R - \lambda_{I_{12}}) + \\
&\beta_2(\lambda_R - \lambda_{I_{12}})).
\end{aligned} \tag{62}$$

**Lemma 3** *$\psi_i(t), i = 1, 2, 3$ are decreasing functions of $t \in [0, T)$.*

Lemma 3 is proved using following the similar methodology to those presented in [3], [9]. We can consider two cases:

1) $h_i$ **is concave.**

Since functions $h_1$, $h_2$ and $h_3$ are concave ($h_i'' < 0, \quad i = 1, 2, 3$), then $(u_i\psi_i - h_i(u_i))$, $i = 1, 2, 3$ are convex functions of $u_i$. Hamiltonian $H$ is a strictly convex function according to (56) and for any $t \in [0, T]$ and it reaches its maximum either at $u_i = 1$ or $u_i = 0$, $i = 1, 2, 3$

$$u_i(t) = \begin{cases} 0, & \text{if } \psi_i < h_i(1); \\ 1, & \text{if } \psi_i > h_i(1). \end{cases} \tag{63}$$

2) $h_i$ **is strictly convex.**

If functions $h_i$ are strictly convex ($h_i'' > 0$, $i = 1, 2, 3$) then $(-h_i(u_i) + u\psi_i, \ i = 1, 2, 3)$ and Hamiltonian is concave function, then $(\frac{dH}{du_i} = -\dot{h}_i(u_i) + \psi_i = 0$, $u_i \in [0, 1]$, $i = 1, 2, 3)$. Then,

$$u_i(t) = \begin{cases} 0\ , & \text{if } \psi_i \leq \frac{dh_i(0)}{du_i}; \\ \frac{dh_i^{-1}(\psi_i)}{du_i}\ , & \text{if } \frac{dh_i(0)}{du_i} < \psi_i \leq \frac{dh_i(1)}{du_i}; \\ 1\ , & \text{if } \psi_i > \frac{dh_i(1)}{du_i}, \end{cases} \tag{64}$$

functions $\psi_i$, $h_i'$, $u_i$ are continuous at all $t \in [0, T]$. In this case $h_i$ is strictly convex and $h_i'$ is strictly increasing function then $h_i'(0) < h_i'(1)$. Thus there exists such moments $t_0$, $t_1$ $(0 < t_0 < t_1 < T)$ such as conditions (64) are satisfied if $\psi_i$ is as described above.

To complete the proof of proposition we consider auxiliary lemma. It follows from (60) that Hamiltonian reaches maximum if and only if the following conditions are satisfied: $(\lambda_R - \lambda_{I_1}) \geq 0$, $(\lambda_R - \lambda_{I_2}) \geq 0$ and $(\lambda_R - \lambda_{I_{12}}) \geq 0$.

**Lemma 4** *For all* $0 \leq t \leq T$*, we have* $(\lambda_R - \lambda_{I_1}) \geq 0$*,* $(\lambda_R - \lambda_{I_{12}}) \geq 0$ *and* $(\lambda_R - \lambda_{I_2}) \geq 0$.

Lemma 4 is proved in the similar way to those in [3], [9] and it is based on the following two properties described before (Property 1 and 2).

**Proof of Lemma 4**

**I.** At time $T$, we have $(\lambda_R(T) - \lambda_{I_1}(T)) = 0$, $(\lambda_R(T) - \lambda_{I_2}(T)) = 0$ and $(\lambda_R(T) - \lambda_{I_{12}}(T)) = 0$ according to (54).

Consider the derivatives

$$\begin{aligned} (\dot{\lambda}_R(T) - \dot{\lambda}_{I_1}(T)) &= -\dot{f}_1(I_1(T)) \leq 0, \\ (\dot{\lambda}_R(T) - \dot{\lambda}_{I_2}(T)) &= -\dot{f}_2(I_2(T)) \leq 0, \\ (\dot{\lambda}_R(T) - \dot{\lambda}_{I_{12}}(T)) &= -\dot{f}_3(I_{12}(T)) \leq 0. \end{aligned} \tag{65}$$

Moreover we may say that $\lambda_i(t) \geq 0$, because $\dot{\lambda}_{I_i}(t) \geq 0$ and $(\lambda_R - \lambda_{I_1}) \geq 0$, $(\lambda_R - \lambda_{I_{12}}) \geq 0$ are positive on the open interval $[0, T]$.

Now we have that functions $\psi_i(T) = 0$ and also we may say that $\psi_i(t)$ are positive in $t \in [0; T]$.

**II.**(Proof by contradiction).

Let $0 \leq t^* \leq T$ be the last instant moment at which one of the inequality constraints are satisfied, i.e.:

$$(\lambda_R - \lambda_{I_1}) \geq 0, (\lambda_R - \lambda_{I_2}) \geq 0, (\lambda_R - \lambda_{I_{12}}) \geq 0$$

and

$$(\lambda_R - \lambda_{I_1}) = 0 \text{ or } (\lambda_R - \lambda_{I_2}) = 0 \text{ or } (\lambda_R - \lambda_{I_{12}}) = 0.$$

First, suppose that following inequality satisfy

$$\begin{aligned}&(\lambda_R(t^*) - \lambda_{I_1}(t^*)) = 0, \quad (\lambda_R(t^*) - \lambda_{I_2}(t^*)) \geq 0,\\&(\lambda_R(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_S(t^*) - \lambda_{I_1}(t^*)) \geq 0,\\&(\lambda_{I_1}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_{I_2}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0.\end{aligned} \tag{66}$$

We have to prove that $(\lambda_R(t) - \lambda_{I_1}(t))$ are non-decreasing function on the interval $[0; T]$. According to Property 1, we consider a time moment $t^{*+}$:

$$\begin{aligned}\dot{\lambda}_R(t^{*+}) - \dot{\lambda}_{I_1}(t^{*+}) = -f_1' - \beta_1 S(\lambda_S - \lambda_{I_1}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}})\\ -u_1(1 - \varepsilon\beta_2)(\lambda_{I_1} - \lambda_R) - \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}).\end{aligned} \tag{67}$$

As well as $f_1(I_1)$ is non-decreasing function and all parameters are non-negatives hence we have $\frac{d}{dt}(\lambda_R(t^{*+}) - \lambda_{I_1}(t^{*+})) \leq 0$. This contradicts Property 1.

The system of ODE is autonomous, i.e., Hamiltonian and the constraints on the control do not have an explicit dependency on the independent variable $t$

$$\begin{aligned}&H(S(t), I_1(t), I_2(t), I_{12}(t), R(t), u_1(t), u_2(t),\\&\quad \lambda_S(t), \lambda_{I_1}(t), \lambda_{I_2}(t), \lambda_{I_{12}}(t), \lambda_R(t)) = \text{const}.\end{aligned} \tag{68}$$

From (52) and (54) we obtain

$$\begin{aligned}H = H(T) = -(f_1(I_1(T)) + f_2(I_2(T)) + f_3(I_{12}(T)) +\\ h_1(u_1(T)) + h_2(u_2(T)) + h_3(u_3(T))).\end{aligned} \tag{69}$$

Since $f_i$ and $h_i$ are non-decreasing functions and according to (52), we receive

$$\lambda_R(t) - \lambda_{I_1}(t) = \frac{1}{(1-\varepsilon\beta_2)u_1 I_1}(H + f_1(I_1) + f_2(I_2) + f_3(I_{12}) +$$
$$h_1(u_1) + h_2(u_2) + h_3(u_3) - \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) -$$
$$\beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) - \tag{70}$$
$$(1-\varepsilon\beta_1)(\lambda_R - \lambda_{I_2})u_2 I_2 - u_3(\lambda_{I_{12}} - \lambda_R) -$$
$$\varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}))$$

Moreover $f_i(I)$ is a non-decreasing function, then $f_i(I_i(T)) > f_i(I_i(t)) > 0$, where $I_i(T) > 0$ according to general assumptions and we have

$$H + f_2(I_2(t)) + f_3(I_{12}(t)) + h_1(u_1(t)) + h_2(u_2(t)) +$$
$$h_3(u_3(t)) - \beta_1 S(t)(I_1(t) + I_{12}(t))(\lambda_{I_1} - \lambda_S) -$$
$$\beta_2 S(t)(I_2(t) + I_{12}(t))(\lambda_{I_2} - \lambda_S) - \tag{71}$$
$$\varepsilon\beta_2 I_1(t)(I_2(t) + I_{12}(t))(\lambda_{I_{12}} - \lambda_{I_1}) -$$
$$\varepsilon\beta_1 I_2(t)(I_1(t) + I_{12}(t))(\lambda_{I_{12}} - \lambda_{I_2}) -$$
$$\psi_2 u_2(t) - \psi_3 u_3(t) \leq -f_1(I_1(T)) + \psi_3 u_1(T) \leq 0.$$

This follows from assumptions for functions $f_i(I_i)$, $i = 1,2,3$ and $h_i(u_i)$, $i = 1,2,3$, such as $I_i(T) > 0$ then $f_i(I_i) > 0$ and $u_i(t) > 0$ then $h_i(u_i) > 0$.

From (54) and (71) we have

$$\dot{\lambda}_R(t^{*+}) - \dot{\lambda}_{I_1}(t^{*+}) = -f_1' - \beta_1 S(\lambda_S - \lambda_{I_1}) - \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) - \frac{1}{I_1}(H + f_1(I_1) + f_2(I_2) + f_3(I_{12}) + h_1(u_1) + h_2(u_2) + h_3(u_3) - \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) - \beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) - \varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) - \psi_2 u_2 - \psi_3 u_3) = \frac{1}{I_1}(-\dot{f}_1 I_1 + f_1) + \frac{1}{I_1}(H + f_2(I_2) + f_3(I_{12}) + h_1(u_1) + h_2(u_2) + h_3(u_3) - \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) - \beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) - \varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) - \psi_2 u_2 - \psi_3 u_3) - \beta_1 S(\lambda_S - \lambda_{I_1}) - \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}). \tag{72}$$

Here the infected cost function $f_1(I_1)$ is convex increasing function and $f_1(0) = 0$, $I_1 > 0$, also we have $f_1(I_2) - \dot{f}_1(I_1)I_1 \leq 0$ by Property 2. From (54), (1) we can show that $(\dot{\lambda}_R(t) - \dot{\lambda}_{I_1}(t)) \leq 0$ and it contradicts Property 1, hence part **II** of the lemma follows.

In this part we will prove that $(\lambda_R(t) - \lambda_{I_2}(t)) > 0$. Analogously to the part **II** of the Lemma 4 we suppose that

$$\begin{aligned} &(\lambda_R(t^*) - \lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_R(t^*) - \lambda_{I_2}(t^*)) = 0, \\ &(\lambda_R(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_S(t^*) - \lambda_{I_1}(t^*)) \geq 0, \\ &(\lambda_{I_1}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_{I_2}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0. \end{aligned} \tag{73}$$

We can see that

$$
\begin{aligned}
&H + f_1(I_1) + f_3(I_{12}) + h_1(u_1(t)) + h_2(u_2(t)) + h_3(u_3)- \\
&-\beta_1 S(t)(I_1(t) + I_{12}(t))(\lambda_{I_1} - \lambda_S) - \beta_2 S(t)(I_2(t)+ \\
&+I_{12}(t))(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(t)(I_2(t) + I_{12}(t))(\lambda_{I_{12}} - \lambda_{I_1})- \\
&-\varepsilon\beta_1 I_2(t)(I_1(t) + I_{12}(t))(\lambda_{I_{12}} - \lambda_{I_2}) - (1-\varepsilon\beta_2)(\lambda_R- \\
&-\lambda_{I_1})u_1 I_1 - (\lambda_R - \lambda_{I_{12}})u_3 I_{12} \le -f_2(I_2(T))+ \\
&+(1-\varepsilon\beta_1)(\lambda_R(T) - \lambda_{I_2}(T))u_2 I_2(T) \le 0.
\end{aligned}
\tag{74}
$$

At the time moment $t^{*+}$ we have:

$$
\begin{aligned}
&\dot\lambda_R(t^{*+}) - \dot\lambda_{I_2}(t^{*+}) = -f_2' - \beta_2 S(\lambda_S - \lambda_{I_2})- \\
&\quad \varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}}) - \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}})+ \\
&\quad \tfrac{1}{I_2}(H + f_1(I_1) + f_2(I_2) + f_3(I_{12}) + h_1(u_1)+ \\
&\quad h_2(u_2) + h_3(u_3) - \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S)- \\
&\quad \beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1})- \\
&\quad \varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) - \psi_1 u_1 - \psi_3 u_3 =
\end{aligned}
\tag{75}
$$

$$
\begin{aligned}
&\dot\lambda_R(t^{*+}) - \dot\lambda_{I_2}(t^{*+}) = \tfrac{1}{I_2}(-\dot f_2 I_2 + f_2) + \tfrac{1}{I_2}(H + f_1(I_1)+ \\
&\quad f_3(I_{12}) + h_1(u_1) + h_2(u_2) + h_3(u_3)- \\
&\quad \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S) - \beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S)- \\
&\quad \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1}) - \varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2})- \\
&\quad \psi_1 u_1 - \psi_3 u_3) - \beta_2 S(\lambda_S - \lambda_{I_2}) - \varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}})- \\
&\quad \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}).
\end{aligned}
\tag{76}
$$

Here $f_2(I_2)$ is convex increasing function and $f_2(0) = 0$, $I_2 > 0$, we have $f_2(I_2) - \dot f_2(I_2)I_2 \le 0$ by Property 2. From (54) and similar to (71) we also have that $\dot\lambda_R(t) - \dot\lambda_{I_2}(t) \le 0$ and it contradicts Property 1, then part **III** of the lemma follows.

**IV**. Similar to previous parts we have proved that $(\lambda_R(t)-\lambda_{I_{12}}(t)) > 0$. In this case we assume that

$$\begin{aligned}
&(\lambda_R(t^*)-\lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_R(t^*)-\lambda_{I_2}(t^*)) \geq 0, \\
&(\lambda_R(t^*)-\lambda_{I_{12}}(t^*)) = 0, \quad (\lambda_S(t^*)-\lambda_{I_1}(t^*)) \geq 0, \\
&(\lambda_{I_1}(t^*)-\lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_{I_2}(t^*)-\lambda_{I_{12}}(t^*)) \geq 0.
\end{aligned} \tag{77}$$

Also we can see that

$$\begin{aligned}
&H + f_1(I_1) + f_2(I_2) + h_1(u_1(t)) + h_2(u_2(t)) + h_3(u_3)- \\
&-\beta_1 S(t)(I_1(t)+I_{12}(t))(\lambda_{I_1}-\lambda_S) - \beta_2 S(t)(I_2(t)+ \\
&+I_{12}(t))(\lambda_{I_2}-\lambda_S) - \varepsilon\beta_2 I_1(t)(I_2(t)+I_{12}(t))(\lambda_{I_{12}}-\lambda_{I_1})- \\
&-\varepsilon\beta_1 I_2(t)(I_1(t)+I_{12}(t))(\lambda_{I_{12}}-\lambda_{I_2}) - (1-\varepsilon\beta_2)(\lambda_R- \\
&-\lambda_{I_1})u_1 I_1 - (1-\varepsilon\beta_1)(\lambda_R-\lambda_{I_2})u_2 I_2 \leq -f_3(I_{12})+ \\
&+(\lambda_R(T)-\lambda_{I_{12}}(T))u_3 I_{12}(T) \leq 0.
\end{aligned} \tag{78}$$

According to Property 1, we have to check time moment $t^{*+}$:

$$\begin{aligned}
&\dot\lambda_R(t^{*+}) - \dot\lambda_{I_{12}}(t^{*+}) = -\tfrac{df_3}{dI_{12}} - \beta_1 S(\lambda_S-\lambda_{I_1}) - \beta_2 S(\lambda_S-\lambda_{I_2})- \\
&\quad \varepsilon\beta_2 I_1(\lambda_{I_1}-\lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2}-\lambda_{I_{12}}) + \tfrac{1}{I_{12}}(H + f_1(I_1) + f_2(I_2) \\
&\quad +f_3(I_{12}) + h_1(u_1) + h_2(u_2) + h_3(u_3) - \beta_1 S(I_1+I_{12})(\lambda_{I_1}-\lambda_S)- \\
&\quad \beta_2 S(I_2+I_{12})(\lambda_{I_2}-\lambda_S) - \varepsilon\beta_2 I_1(I_2+I_{12})(\lambda_{I_{12}}-\lambda_{I_1})- \\
&\quad \varepsilon\beta_1 I_2(I_1+I_{12})(\lambda_{I_{12}}-\lambda_{I_2}) - (1-\varepsilon\beta_2)(\lambda_R-\lambda_{I_1})u_1 I_1- \\
&\quad (1-\varepsilon\beta_1)(\lambda_R-\lambda_{I_2})u_2 I_2) =
\end{aligned} \tag{79}$$

$$= \frac{1}{I_{12}}(-\dot{f}_3 I_{12} + f_3) + \frac{1}{I_{12}}(H + f_1(I_1) + f_2(I_2)+$$
$$h_1(u_1) + h_2(u_2) + h_3(u_3) - \beta_1 S(I_1 + I_{12})(\lambda_{I_1} - \lambda_S)-$$
$$\beta_2 S(I_2 + I_{12})(\lambda_{I_2} - \lambda_S) - \varepsilon\beta_2 I_1(I_2 + I_{12})(\lambda_{I_{12}} - \lambda_{I_1})- \quad (80)$$
$$\varepsilon\beta_1 I_2(I_1 + I_{12})(\lambda_{I_{12}} - \lambda_{I_2}) - \psi_1 u_1 - \psi_2 u_2 - \beta_1 S(\lambda_S - \lambda_{I_1})-$$
$$\beta_2 S(\lambda_S - \lambda_{I_2}) - \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) - \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}).$$

From (54), (71) and Property 2 we receive that $(\dot{\lambda}_R(t) - \dot{\lambda}_{I_{12}}(t)) \leq 0$ that contradicts Property 1 and hence part **IV** follows, then the time moment $t^*$ does not exist. This is completed the proof of lemma 4.

Now we have to prove that $(\lambda_R(t) - \lambda_S(t)) \geq 0$. Analogously we assume:

$$\begin{aligned}
&(\lambda_R(t^*) - \lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_R(t^*) - \lambda_{I_2}(t^*)) \geq 0,\\
&(\lambda_R(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_R(t^*) - \lambda_S(t^*)) = 0,\\
&(\lambda_S(t^*) - \lambda_{I_1}(t^*)) \geq 0, \quad (\lambda_S(t^*) - \lambda_{I_2}(t^*)) \geq 0\\
&(\lambda_{I_1}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0, \quad (\lambda_{I_2}(t^*) - \lambda_{I_{12}}(t^*)) \geq 0.
\end{aligned} \quad (81)$$

From (53) we obtain:

$$\begin{aligned}
\dot{\lambda}_R(t^*) - \dot{\lambda}_S(t^*) = -\beta_1(\lambda_S - \lambda_{I_1})(I_1 + I_{12})-\\
-\beta_2(\lambda_S - \lambda_{I_2})(I_2 + I_{12}) + \gamma u_4(\lambda_R - \lambda_S).
\end{aligned} \quad (82)$$

According to assumption (81), the difference $\dot{\lambda}_R(t^{*+}) - \dot{\lambda}_S(t^{*+}) \leq 0$ is negative at the time moment $t^{*+}$, then function $(\lambda_R(t) - \lambda_S(t))$ decrease. It contradicts Property 1. At time moment $t = T$ $(\lambda_R(t) - \lambda_S(t))$ is equal to zero and on the whole interval it is decreasing. We proved that $(\lambda_R(t) - \lambda_S(t)) \geq 0$ for $t \in [0; T]$.

We get that on the interval $t \in [0; T]$:

$$\begin{aligned}
&(\lambda_R(t) - \lambda_{I_1}(t)) \geq 0; \\
&(\lambda_R(t) - \lambda_{I_2}(t)) \geq 0; \\
&(\lambda_R(t) - \lambda_{I_{12}}(t)) \geq 0; \\
&(\lambda_R(t) - \lambda_S(t)) \geq 0.
\end{aligned} \tag{83}$$

The proof of the main proposition is completed.