OPTIMAL CONTROL MODEL FOR COMPUTER VIRUSES

Digbalay Bose
Roll-143070026
M.Tech 1st year, Control and Computing
Electrical Engineering Department
I.I.T Bombay

May 5, 2015

1 Introduction to problem

Computer viruses are malicious programs that can adversely affect the computer systems by erasing data, creating multiple copies, thus hampering the normal operation. Development of a suitable mathematical model for the propagation of viruses is of paramount importance, as it helps in creation of effective strategies for stopping their growth. The greater similarity between biological and computer viruses have resulted in the SIRA (Susceptible-Infected-Removed-antidotal) model \[1\]. The modification of the SIRA model and its formulation as an optimal control problem have been discussed in this article.

2 Model

The original SIRA model is governed by the following system of equations:

\[
\begin{align*}
\frac{dS}{dt} &= N - \alpha_{SA}SA - \beta SI - \mu S + \sigma R \\
\frac{dI}{dt} &= \beta SI - \alpha_{IA}AI - \delta I - \mu I \\
\frac{dR}{dt} &= \delta I - \sigma R - \mu R \\
\frac{dA}{dt} &= \alpha_{SA}SA + \alpha_{IA}AI - \mu A
\end{align*}
\]

(1)

Here the entire population is divided into 4 categories:

- **S**: susceptible computers
- **I**: infected computers
- **R**: recovered computers
- **A**: antidotal computers

The modifications include change of the parameters \(\alpha_{IA}\) into \(u_1(t)\) and \(\delta\) into \(u_2(t)\). \(N\) represents the influx rate i.e. the rate at which new computers are added to the network. \(\mu\) is the proportion coefficient for the mortality rate (i.e. not due to virus action). Considering the modification of the parameters as discussed above, the infected computers can be fixed using anti-virus programs with
a rate proportional to $AI$ and a proportion factor given by $u_1(t)$. The useless infected computers are removed via a rate given by $u_2(t)$.

$N$ is usually kept to be 0 indicating that the propagation of the virus is much faster than the incorporation of new computers in the network. Also the mortality rate is considered to be zero. The modified system of equations is given by:

\[
\begin{align*}
\frac{dS}{dt} &= N - \alpha S A - \beta SI + \sigma R \\
\frac{dI}{dt} &= \beta SI - u_1(t) AI - u_2(t) I \\
\frac{dR}{dt} &= u_2(t) I - \sigma R \\
\frac{dA}{dt} &= \alpha S A + u_1(t) AI
\end{align*}
\] 

(2)

### 2.1 Optimal control formulation

The optimal control problem is formulated as:

\[
\min J(u) = \int_0^T [I(t) + \epsilon u_1^2(t) + \tau u_2^2(t)]
\]

(3)

subject to the state dynamics given by (2). The set of inputs $(u_1, u_2)$ must belong to the class given by:

\[
U = \{u = (u_1, u_2) : u_i \text{ measurable}, 0 \leq u_i(t) \leq 1, t \in [0, T] \text{ for } i = 1, 2\}
\]

(4)

### 3 Simulation and results

The initial conditions $(S_0, I_0, R_0, A_0)$ and parameters $(\alpha_{SA}, \beta, \sigma, \epsilon, \tau)$ are listed:

<table>
<thead>
<tr>
<th>$S_0$</th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$\alpha_{SA}$</th>
<th>$\beta$</th>
<th>$\sigma$</th>
<th>$\tau$</th>
<th>$\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>0.6</td>
<td>0.1</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: List of initial conditions and parameters

Simulation was performed using GPOPS (Gauss Pseudospectral Optimization Software)[2]. The final time $T$ was taken to be 25. The codes are attached in Appendix A. The associated figures are given below:
3.1 Results of simulation (with control inputs $u_1(t)$ and $u_2(t)$)

Figure 2: Number of susceptible computers under application of control

Figure 3: Number of infected computers under application of control

Figure 4: Number of recovered computers under application of control
Figure 5: Number of antidotal computers under application of control

Figure 6: Control variable $u_1(t)$

Figure 7: Control variable $u_1(t)$ (expanded)
3.2 Results of simulation without control inputs

Here the simulation results (without control) are considered. The blue line indicates the mode of operation without any control. The following figures are taken from [1].
Figure 10: Number of susceptible computers (with and without control)

Figure 11: Number of infected computers (with and without control)

Figure 12: Number of recovered computers (with and without control)
4 Discussion

The results obtained using the optimal control model are compared with the cases when no control is present. Considering the figures 11 and 3 (for infected computers), the number of nodes infected with computer virus become less upon application of optimal control strategy\((u_1(t) \text{ and } u_2(t))\). In figure 11, it is clearly seen that the population of the infected nodes without control is larger, which decreases significantly in figure 3. Similarly in the case of recovered computers, the population becomes larger upon application of optimal controls as seen in the figures 12 and 4. The number of antidotal computers when compared between figures 5 and 13, show a dramatic increase in figure 5, when the optimal controls are applied. Thus greater number of nodes are being converted from infected domain to recovered or susceptible. Hence the optimal control strategy \((u_1(t) \text{ and } u_2(t))\)is effective in controlling the number of infected computers in the system.

References


