International Journal of System Signal Control and Engineering Application 4 (4-6): 74-79, 2011
ISSN: 1997-5422

Stochastic Model on Computer Virus

Upendra Kumar
Department of Information Technology, Birla Institute of Technology, Mesra, Ranchi, India

Abstract: Susceptible (S), Exposed (E), Infectious (I), Quarantine (Q), Recovered (R), Stochastic Compartmental Model for computer virus has been developed and researchers assume that the transmission of virus in computer network are probabilistic. To implement this concept, Reversible Jump Monte Carlo Markovian Chain (MCMC) algorithm has been implemented. The MCMC algorithm basically includes Bayesian inferences, likelihood function, prior distribution, posterior distribution and predictive distribution. The threshold R_0 has been defined to determine the removal of the disease if threshold parameter is at most one, otherwise endemic state exists. The effect of quarantine has been implemented on SEIR Stochastic Model and time dependency has been analyzed by implementing predictive distribution for virus removal time.

Key words: Computer virus, stochastic model, quarantine, MCMC algorithm, Bayesian inferences, India

INTRODUCTION

The electronic mail, secondary storage devices, etc. are the major source of transmission of computer virus in the computer network these days (Newman *et al*., 2002). Transmission of viruses in computer network and the concept of Bayesian inference for stochastic epidemic has been used (Boys and Giles, 2007). The action of malicious objects throughout the network can be studied by using epidemiological model for diseases propagation (Mishra and Saini, 2007; Kermack and McKendrick, 1927; Mishra and Jha, 2007; Gelenbe, 2005, 2007; Gelenbe *et al*., 2004; Piqueira and Cesar, 2008; Piqueira *et al*., 2005; Forest *et al*., 1994; Wang and Wang, 2003). Based on the SIR Classical Epidemic Model, dynamical models for malicious objects propagation were proposed, providing estimation for temporal evaluation of infected nodes depending on network parameter considering topological aspects of network (Allen, 1994; Kermack and McKendrick, 1932, 1933). The Model SEIR assumed that secondary hosts have a permanent immunization period with certain probability which are consistent with real situation. In order to overcome the limitation, a SEIRS Model with latent and temporary immunity period is developed. Extending this property the new compartment quarantine has been introduced to control the spread of virus in computer network by several actions (Mishra and Jha, 2010; Zou *et al*., 2005; Kephart *et al*., 1993; Keeling and Eames, 2005; Williamson and Laeveillae, 2003).

In this study, researchers have developed SEIQR Stochastic Compartmental Model for virus outbreak where focus have been given on time inhomogenity and infectious nodes are removed from the epidemic. The viruses are spread in computer network exponentially and the rate parameters of the distribution are step function. The SEIQR Compartmental Model has different steps of recovery and having different probabilities to transfer between different compartments. The complication arises during partial observation of data for epidemic outbreak. Recent resources have focused on the implementation of MCMC Technique to help the overcome a missing data problem and parameters are estimated using Bayesian perspective (Thommes and Coates, 2005; Yan and Liu, 2006; Gibson and Renshaw, 1998; O'Neill and Roberts, 1999). We have used several parameters like latent period, quarantine time, immunization time for viruses. On all existing parameters reversible jump, MCMC methodology has been used to quarantine and remove viruses from the computer nodes (Gamerman and Hopes, 2006; Green, 1995).

MULTI-TYPE SEIQR MODEL

This model has five compartment and has defined as Susceptible-Exposed-Infected-Quarantined-Recovered in which susceptible compartment is partitioned into m sub categories. The susceptible nodes having m groups are due to different potentiality of viruses. Initially, there are N_i nodes in group i, i = 1, 2, 3,..., m and the epidemic begin as soon as any member of the group contracts the viruses. Having contracted the viruses, nodes enter in a latent period where they show no symptom of the attack of the virus and are unable to infect susceptible nodes before becoming infective. We have considered latent period of fixed length c with cohesion appropriately

according to known properties of the infection (infection due to viruses). The epidemic proceeds according to the following transitions probabilities in the small interval of time (t, t+dt) and the model is:

$$P_r\{S_i(t+dt), E_i(t+dt)\} = \{S_i(t) -1, E_i(t) +1\} = \beta S_i(t) I_i(t)dt+o(dt)$$

$$P_r \{I_i(t+dt), Q_i(t+dt)\} = \{I_i(t) -1, Q_i(t) +1\} = \gamma I_i(t)dt+o(dt)$$

$$P_r \{Q_i(t+dt), R_i(t+dt)\} = \{Q_i(t) -1, R_i(t) +1\} = \alpha Q_i(t)dt+o(dt)$$

for i = 1, 2, 3,..., m where $S_i(t)$, $E_i(t)$, $I_i(t)$, $Q_i(t)$, $R_i(t)$ denotes number of nodes in group i at time t that are susceptible, exposed, quarantine and recovered, respectively with:

$$I(t) = \sum_{i=1}^{i=m} I_i(t)$$

Each infected node makes infective contract with number of susceptible group i with the rate β_i and pass through latent period before becoming infective. Since the infection time is unobserved, the first sign that the epidemic is in progress is at the time of first removal. For this reason, the time of first removal is set to zero and all infection and removal time are set relative to this reference time. The time dependence of the removal rate is modeled as a step of function with k steps at time s, s = $\{s_1, s_2,\ldots, s_k\}$. The function $\gamma(t)$ takes the values γ_j where $t\in(s_j, s_{j+1})$ for j = 0, 1, 2, …, k, where for convenience $s_0\equiv\infty$ and $s_{k+1}\equiv T$. This formulation is equivalent to a model in which distribution of the infectious period has an exponential immunity α_i within interval (s_o, s_{i+1}).

The end of epidemic is defined as the first time in which no infective, exposed and quarantine nodes remains in the computer network. Initially, all nodes are susceptibles and the infections consist of several groups of susceptibles denoted as N = $\{N_1, N_2, \ldots, N_m\}$. Since each infected groups are either quarantined or recovered, so have been denoted as q_{ij} or r_{ij} (where i = 1, 2, 3, …, m and j = 1, 2, 3, …, n_i), denotes time of quarantine or removal j in group i, n_i denotes total numbers of group of nodes quarantined or recovered.

Researchers have defined τ_{ij} for the time of infection j in group i, so node becomes infective at time τ_{ij}+c at the end of latent period.

Researchers have used r = r_{ij} to denotes the matrix of removal time, q = q_{ij} denotes the matrix of quarantine time and $\tau = \tau_{ij}$ denotes the matrix of infections time. Infection time excluding the time of first infection is $\tau_{i\ min}.1$. Researchers have observed complete infected nodes so total number of quarantined nodes and recovered nodes in each group is equal to n_i and denoted as:

$$n = \sum_{i=1}^{m} n_i$$

Complete data likelihood: The several parametric inferences have been used to define likelihood function, like infection rate of node $\beta = (\beta_1, \beta_2, \ldots, \beta_m)$, the quarantined rate $\gamma = (\gamma_1, \gamma_{2,}\ldots, \gamma_k)$ and recovered rates of nodes $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ because removal rate steps are k and $\bar{\tau}$ = mean of infection time of the nodes has been taken. Researchers assume that the infection, quartine and removal times data are observed. The likelihood function is:

$$\pi\left(\tau\frac{r}{\gamma},\frac{q}{\beta}\alpha,k,s,i_{min},qi_{min.1}\right) = \left\{\prod_{i=1,i\neq i_{min}} \beta Si(\bar{\tau}i_{i1})I(\bar{\tau}_{i1})\right\} g\left\{\prod_{i=1}^{m}\prod_{i=2}^{ni}\beta_i S_i(\bar{\tau}_{i1})I(\bar{\tau}_{i1})\right\} * \left\{\prod_{i=1}^{m}\prod_{i=2}^{ni}\gamma(\bar{q}_{i1})I_i(\bar{q}_{i1})\right\} * \left\{\prod_{i=1}^{m}\prod_{i=1}^{ni}\alpha(\bar{r}_{i1})Q(\bar{r}_{i1})\right\} * \exp\left\{-\sum_{i=1}^{m}\left(\beta_i\int_{T_{i\min.1}}^{T} S_i(t)I(t)dt\right) - \int_{\tau_{i\min.1}}^{T}\gamma(t)I(t)dt - \int_{\tau_{i\min.1}}^{T}\alpha(t)Q(t)dt\right\}$$

Prior model: In Bayesian statistical inference, a prior probability distribution often called prior of an uncertain quantity. Researchers have used prior distribution for infection rate (β_i), the group and time for the first infection and step function for the quarantine and removal rate. The quarantine rate γ_i, removal rate α_i all are imposed on gamma prior:

$$\beta_i \cong \Gamma\left(g_{\beta i}, h_{\beta i}\right),\ i = 1,2,\ldots,m$$

$$\frac{\gamma_j}{k} \cong \left(g_\gamma, h_\gamma\right),\ j = 0,1,\ldots,k$$

$$\frac{\alpha_j}{k} \cong \left(g_\alpha, h_\alpha\right),\ j = 0,1,\ldots,k$$

The number of steps k in the interval (0, T) has been used to define Poisson distribution. The parameter λ with probability function given by:

$$\prod(k) \propto \frac{\lambda^k}{k!}\quad k = 0,1,\ldots K_{max}$$

having probability function with maximum steps (K_{max}). The K_{max} has been used to define maximum number of change point in the quarantine rate and removal rate. Since K is not constant for informative prior specification for the γ_j/k and α_j/k so it define posterior distribution. The concept of joint distribution has been used to every ath order statistic for size {a(k+1)-1} points sampled uniformly over (0, T).

Since in case of prior distribution no jump occurs in step function at time zero, so consecutive change has been defined S_{j+1} - S_j/T by implementing β(a, ak) distribution. Thus, α has reflect prior benefit about change point in the quarantine rate γ and removal rate α with large value of an essential putting step position on a fixed grid.

The small value of α can result in the MCMC algorithm accepting intervals containing no data (removal times), define prior option. Its impact has implemented on posterior distribution.

Posterior distribution: The posterior distribution is the combination of likelihood function and the prior distribution work on the observed data or unknown parameter and are combined using Baye's theorm:

$$\pi\left(k,\beta,\gamma,\alpha,s,i_{min},\tau_{i_{min.1}}q_{i_{min.1}},\frac{\tau}{q},\frac{q}{r}\right)$$
$$\propto\pi\left(k,\tau,\gamma,\alpha,s,i_{min},\tau_{i_{min.1}}q_{i_{min.1}},\frac{r}{\gamma},\frac{q}{\beta}\right)*$$
$$\pi\left(k,\beta,\gamma,\alpha,s,i_{min},\tau_{i_{min.1}}q_{i_{min.1}}\right)$$

To define equilibrium distribution of posterior inference, Markov Chain Monte Carlo (MCMC) process has been used.

Initially, MCMC is used with some suitable parameters to initialize Markovian Chain. Further Markovian Chain has been simulated by successively updating parameter values with the help of either Gibbs sampler or Metropolish-Hasting sampler which define equilibrium distribution of Markovian Chain used in posterior distribution.

The updated parametric values has also changed dimensionality of Markovian Chain. The MCMC algorithm has constructed and employed Gibbs sampler updates and appropriates Metropolish-Hastings updates, both are used to simulate from posterior distribution. This technique use to explore the spaces of infections rate, quarantines rate, hidden infection and quarantine times for uncertainty of steps function k and to minimize complication on number of steps k, reversible jump has been used.

Inferences for the basic reproduction number: For the epidemiological interest, researchers have defined the basic reproduction number R_0. To define R_0, the inferences about model parameters such as the unobserved infection rates, infection times, quarantine rate, quarantine time and removal rate steps functions is used.

Heterogeneity in infectivity across the population has been studied using a model which allows nodes to have different reproduction number, each of which is drawn from a distribution with mean R_0, however we have focused on estimating the nodes level reproduction R_0 (Becker and Hopper, 1983).

The epidemic will occur if $R_0>1$ and the model is virus infection free if $R_0 \leq 1$. Initially epidemic has been defined by Poisson process with rate:

$$\sum_{i=1}^{m}\beta_i N_i$$

infective remain so far as mean period is $1/\gamma_0$. The basic reproduction number for the model is:

$$R_o = \frac{1}{\gamma_o}\sum_{i=1}^{m}\beta_i N_i$$

MCMC SCHEME

In this MCMC scheme, researcher's thought are related to make infection time feasible. At the end of epidemic no infectious, exposed or quarantine nodes exist in the network. The MCMC scheme is cycled with the help of following steps:

Update infection time : In this step, the infection time has been continuously updated by randomly selected infection time τ′ sampled from a U (-d, T) distribution. The new time for which nodes becomes infective is τ′+c which will also update i_{min}.

If $\tau' \leq \tau'_{i_{min.1}}$ and randomly selected infective is different to i_{min} having the move with probability min {1, A} and includes quarantine time q where:

$$A = \frac{\pi\left(\tau', q/\beta, \gamma, k, s, i'_{min}, \tau'_{i_{min.1}}\right)}{\pi\left(\tau, q/\beta, \gamma, k, s, i_{min}, \tau_{i_{min.1}}\right)}$$

Update infection parameter: Sample a new value for each infection rate parameters β_i(i = 1, 2, ..., m) from the conditional posterior distribution gives all other states of the of the chain as:

$$\beta_i \left| \cong \Gamma \left(g\beta_i + n_i - \delta_{ijmin}, h\beta_i + \int_{\tau_{i_{min}1}}^{T} S_i(t) I(t) dt \right) \right.$$

where, i = 1, 2..., m, δ_{ij} is the kronecker's delta function.

Update quarantine time: The update sequence of hidden quarantine time by randomly (uniformly) chosen quarantine time to new time q′ sampled from a U(-d, T) distribution shift the associated time at which nodes becomes quarantined to a new time q′+c which will also update i_{min} if the group of randomly selected quarantine is different to i_{min} and $q' \le q_{i_{min}1}$ accept the proposed move with the probability min {1, A} where:

$$A = \frac{\pi\left(q', r/\gamma, \alpha, k, s, i_{min}, q'_{i_{min}1}\right)}{\pi\left(q, r/\gamma, \gamma, k, s, i_{min}, q_{i_{min}1}\right)}$$

Update quarantine parameter: Sample a new value for each quarantine rate parameter γ(i = 1, 2,…, m) from its conditional posterior distribution given all other status of chain as:

$$\gamma_j \left| \cong \Gamma \left(g\gamma_i + n_i - \delta'_{ijmin}, h\gamma_i + \int_{q_{i_{min}1}}^{T} \gamma Q(t) dt \right) \right.$$

where, i = 1, 2, …, m and δ'_{ij} is Kronecker's delta function.

Update removal rate step function: The removal rate step function have been used and includes different move for each aspect of the step function. At each interaction one of the following step must be involved:

Updates removal rates step with probability p_1: For each removal rate α_j, j = 0, 1, 2, …, k condition on k, s and other states of chain has been included and new values of simulation have been defined as:

$$\alpha_j \left| k, s \cong \Gamma \left(g_\alpha + w_j, h_\alpha + \int_{s_j}^{s_{j+1}} Q(t) dt \right) \right. \quad j = 0, 1, \ldots, k$$

where, α_j is total number of removal in (s_j, s_{j+1}).

Update existing step position with probability p_2: Select one of the step positions s_j uniformly from the k existing steps, propose moving s_j to a new positions S'_{ji} sampled from a $U(s_{J-1}, s_{J+1})$ distribution. Accepts the position with probability min {1, A} where A has been given as:

$$A = \frac{\pi\left(q, r/\gamma, \alpha, k, s', i_{min}, q_{i_{min}1}\right)}{\pi\left(q, r/\gamma, \alpha, k, s, i_{min}, q_{i_{min}1}\right)} * \left\{ \frac{\left(s_{j+1} - s'_j\right)\left(s'_j - s_{j-1}\right)}{\left(s_{j+1} - s_j\right)\left(s_j - s_{j-1}\right)} \right\}^{\alpha-1}$$

Increase number of steps with probability p_3: Researchers have used an additional step at a position s′ sampled uniformly over (0, T) and $s' \in (s_j, s_{J-1})$. The current removal rate α_j over (s_j, s_{j+1}) must be split into removal rate α'_j over the removal interval (s_j, s') and α'_{j+1} over (s', s_{j+1}). Adding a new step induce an enlargement of the parameter sub-space by two new parameter s′ and an additional removal rate. A proposed division of α_j into $(\alpha'_j, \alpha'_{j+1})$ is achieved using a stochastic innovation and preserving a weighted geometric mean as:

$$\left(s' - s_j\right)\log\alpha'_j + \left(s_{j+1} - s'\right)\log\alpha'_{j+1} = \left(s_{j+1} - s_{ji}\right)\log\alpha_j$$

and $$\frac{\alpha'_{j+1}}{\alpha'} = \frac{1-u}{u}$$

where, u is sampled from a u(0, 1) distribution. The move is accepted with probability min (1, A):

$$A = \frac{\pi\left(q, r/\gamma, \alpha', k', s', i_{min}, q_{i_{min}1}\right)}{\pi\left(q, r/\gamma, \alpha, k, s', i_{min}, q_{i_{min}1}\right)} * \frac{\pi(k+1)}{\pi(k)} * \frac{\left[\alpha(k+2)-1\right]}{T^{\alpha}\left[\alpha(k+1)-1\right]!(\alpha-1)!} \left[\frac{\left(s' - s_j\right)\left(s_{j+1} - s_j\right)^{\alpha-1}}{\left(s_{j+1} - s_j\right)} \right] * \frac{h_\alpha g_\alpha}{\Gamma(g_\alpha)} * \left(\frac{\alpha'_j \alpha'_{j+1}}{\alpha_j} \right)^{g_{\alpha-1}} * \exp\left\{ -h_\alpha\left(\alpha'_j + \alpha'_{j+1} - \alpha_j\right) \right\} * \frac{p_4 T}{p_3^{(k+1)}} * \frac{\left(\alpha'_j + \alpha'_{j+1}\right)^2}{\alpha_j}$$

The increase in the number of steps beyond k_{max} can not be accepted as $\pi(k = {}^{k}_{max}+1) = 0$.

Decrease number of steps with probability $p_4 = 1 - p_1 - p_2 - p_3$: Researchers have used removal of step position s_j. It has been uniformly selected k existing steps. The removal rate (α_{j-1}, α_j) are replaced by a new removal rate α'_{j-1} which satisfies the weighted geometric mean condition:

$$\left(s_j - s_{j-1}\right)\log\alpha_{j-1} + \left(s_{j+1} - s_j\right)\log\alpha_j = \left(s_{j+1} - s_{j-1}\right)\log\alpha'_{j-1}$$

The move is accepted with probability min {1, A} where A is:

$$A = \frac{\pi\left(q, r/\gamma, \alpha', k', s', i_{min}, q_{i_{min1}}\right)}{\pi\left(q, r/\gamma, \alpha, k, s', i_{min}, q_{i_{min1}}\right)} * \frac{\pi(k-1)}{\pi(k)} * \frac{T^{\alpha}(\alpha k-1)!(\alpha-1)!}{\left[\alpha(k+1)-1\right]!} \left[\frac{\left(s_{j+1}-s_j\right)}{\left(s_{j+1}-s_j\right)\left(s_j-s_{j-1}\right)}\right]^{\alpha-1} * \frac{\Gamma(g_{\alpha})}{h_{\alpha}g_{\alpha}} * \left(\frac{\alpha'_{j-1}}{\alpha_{j-1}\alpha_j}\right)^{g_{\alpha-1}} * \exp\left\{-h_{\alpha}\left(\alpha'_{j-1}+\alpha'_{j-1}-\alpha_j\right)\right\} * \frac{p_3 k}{p_4 T} * \frac{\alpha'_{j-1}}{\left(\alpha_{j-1}+\alpha\right)^2}$$

The reduction in the number of step below zero can not be accepted.

CONCLUSION

This model has been formulated to define different states of viruses in which quarantine and removal rates are time inhomogenou. The quarantine and removal times of infective nodes are available for analysis. The complete data likelihood has been defined on observed data and further concept of prior model has been implemented. The quarantine and removal rates changes have been formulated using posterior distribution.

In this model predictive distribution for quarantine and removal times have been used as improved fit. R_0 has been constructed for dynamic spread of viruses. The MCMC algorithm has been formulated for computer network viruses outbreak. Further reversible jump has used for the models with time dependency for infection rate, quarantine and removal rate on more improved data.

REFERENCES

Allen, L.G.S., 1994. Some discrete time *SI, SIR* and *SIS* epidemic model. Math. Biol., 24: 83-105.

Becker, N.G. and J.L. Hopper, 1983. Accessing the heterogeneity of disease spread through a community. Am. J. Epidem. 117: 362-374.

Boys, R.J. and P.R. Giles, 2007. Bayesian inferences for stochastic epidemic models with time-inhogogeneous removal rates. J. Math. Biol., 55: 223-247.

Forest, S., S. Hofmeyr, A. Somayaji and T. Longstaff, 1994. Self-nonself discrimination in a computer. Proceedings of the Symposium on Computer Security and Privacy, May 16-18, 1994, IEEE Computer Society, USA., pp: 202-202.

Gamerman, D. and H.F. Hopes, 2006. Markov chain Monte Carlo: Stochastic Simulation for Bayesian Inference. 2nd. Ed. Taylor and Francis, USA., ISBN: 1584885874, Pages: 323.

Gelenbe, E., 2005. Keeping viruses under control. Comput. Inf. Sci. ISCIS, 3733: 304-311.

Gelenbe, E., 2007. Dealing with software viruses: A biological paradigm. Inf. Secur. Tech. Rep., 12: 242-250.

Gelenbe, E., V. Kaptan and Y. Wang, 2004. Biological metaphors for agent behavior. Comput. Inf. Sci. ISCIS, 3280: 667-675.

Gibson, G.J. and E. Renshaw, 1998. Estimating parameters in stochastic compartmental model using Markov Chain methods. Math. Med. Biol., 15: 19-40.

Green, P.J., 1995. Reversible jump markov chain computation and bayesian model determination. Biometrika, 82: 711-712.

Keeling, M.J. and K.T.D. Eames, 2005. Networks and epidemic models. J. R. Soc. Interface, 2: 295-307.

Kephart, J.O., S.R. White and D.M. Chess, 1993. Computers and epidemiology. IEEE Spectrum, 30: 20-26.

Kermack, W.O. and A.G. McKendrick, 1927. Contributions of mathematical theory to epidemics. Proc. Royal Soc. London Ser. A, 115: 700-721.

Kermack, W.O. and A.G. McKendrick, 1932. Contributions of mathematical theory to epidemics. Proc. R. Soc. London Ser. A, 138: 55-83.

Kermack, W.O. and A.G. McKendrick, 1933. Contributions of mathematical theory to epidemics. Proc. R. Soc. London Ser. A, 141: 94-122.

Mishra, B.K. and D.K. Saini, 2007. SEIRS epidemic model with delay for transmission of malicious objects in computer network. Applied Math. Comput., 188: 1476-1482.

Mishra, B.K. and N. Jha, 2007. Fixed period of temporary immunity after run of anti-malicious software on computer nodes. Applied Math. Comput., 190: 1207-1212.

Mishra, B.K. and N. Jha, 2010. SEIQRS model for the transmission of malacious objects in computer network. Applied Math. Modell., 34: 710-715.

Newman, M.E.J., S. Forrest and J. Balthrop, 2002. Email networks and the spread of computer viruses. Phys. Rev. E, 66: 035101-1-035101-4.

O'Neill, P.D. and G.O. Roberts, 1999. Bayesian inference for partially observed stochastics epidemic. J. R. Statist. Soc., 62: 121-129.

Piqueira, J.R.C. and F.B. Cesar, 2008. Dynamical models for computer viruses propagation. Math. Problems Eng., 2008: 1-11.

Piqueira, J.R.C., B.F. Navarro and L.H.A. Monteiro, 2005. Epidemiological models applied to viruses in computer networks. J. Comput. Sci., 1: 31-34.

Thommes, R.W. and M.J. Coates, 2005. Modeling virus propagation in peer to peer networks. Proceedings of the 5th International Conference on Information, Communications and Signal Processing, December 6-9, 2005, Bangkok, pp: 981-985.

Wang, Y. and C.X. Wang, 2003. Modeling the effects of timing parameters on virus propagation. Proceedings of the ACM Workshop on Rapid Malcode, October 27-30, 2003, ACM Press, Washington DC, USA., pp: 61-66.

Williamson, M.M. and J. Laeveillae, 2003. An epidemiological model of virus spread and cleanup. Technical Reports HPL-2003-39. http://www.hpl.hp.com/techreports/2003/HPL-2003-39.html.

Yan, P. and S. Liu, 2006. SEIR epidemic model with delay. ANZIAM J., 48: 119-134.

Zou, C.C., W.B. Gong, D. Towsley and L.X. Gao, 2005. The monitoring and early detection of internet worms. IEEE/ACM Trans. Networking, 13: 961-974.