Столыпинский вестник

Научная статья

Original article

УДК 004.056

МОДЕЛИРОВАНИЕ ВРЕДОНОСНОГО ТРАФИКА В ЛОКАЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

SIMULATION OF MALICIOUS TRAFFIC IN LOCAL COMPUTER NETWORKS

Тураев Саиджон Эркинович, Аспирант, Национальный исследовательский университет ИТМО, (Санкт-Петербург)

Заколдаев Данил Анатольевич, научный руководитель, кандидат технических наук, доцент, Национальный исследовательский университет ИТМО, (Санкт-Петербург)

Turaev Saijon Erkinovich, PhD Student, ITMO National Research University, (St. Petersburg)

Zakoldaev Danil Anatolievich, scientific adviser, candidate of technical sciences, associate professor, National Research University ITMO, (St. Petersburg)

Аннотация. Основной целью защиты от вредоносных программ является исследование методов обнаружения вредоносных программ. В данной статье приводится математическое моделирование распространения вредоносного трафика в локальных вычислительных сетях. Проводится факторный анализ распространения трафика через уравнение регрессии. Эти работы не учитывают методы графических моделей обнаружения вредоносных программ и не раскрывают всех недостатков этих методов. В

настоящее время необходимо оценить методы обнаружения вредоносных программ, чтобы определить наиболее эффективный метод обнаружения вредоносных программ.

Annotation. The main task in protection against malicious programs is the study of methods for detecting malicious programs. This article provides mathematical modeling of the spread of malicious traffic in local area networks. A factor analysis of traffic distribution is carried out through the regression equation. In these works, the method of graph models of malware detection is not considered; all the shortcomings of the methods are not revealed. There is currently a need to evaluate malware detection methods in order to determine the most effective malware detection method.

Ключевые слова: моделирование, трафик, программное обеспечение, локальные вычислительные сети, вредоносный трафик.

Keywords: modeling, traffic, software, local area networks, malicious traffic

Пусть есть некоторая система трафика в локальных вычислительных сетях (ЛВС) - передача пакетов данных между устройствами и сетью интернет, состоящая из одного элемента. Время ее безотказной работы ξ имеет распределение [1] $F(x) = \mathbf{P}\{\xi < x\}$, $F(0) = 0$.

Предположим, что система подвергается атакам вредоносного программного обеспечения согласно процессу Пуассона. Как известно, в таком процессе интервалы η между атакам вредоносного программного обеспечения и имеют экспоненциальное распределение с параметром λ:

$$\mathbf{P}\{\eta < x\} = 1 - e^{-\lambda x},\ x \geq 0.$$

Система начинает исправно работать, и в момент старта осуществляется назначение времени начала планового предупредительного обновления трафика в системе ν с распределением $G(x) = \mathbf{P}\{\nu < x\},\ G(0) = 0$.

Особенностью данной модели является предположение, что отказ системы под воздействием вредоносного трафика не проявляется

самостоятельно. Мы узнаем об отказе только тогда, когда наступит момент прекращения работы устройств ЛВС. Если к моменту $\nu = \tau$ система не отказала, начинается плановая профилактика системы антивирусным ПО (профилактика системы антивирусным ПО 1-го типа) продолжительностью γ_1, распределенная следующим образом: [2] $F_1(x) = \mathbf{P}\{\gamma_1 < x\},\ F_1(0) = 0.$

Если же отказ произошел до назначенного момента $\nu = \tau$, то в момент $\nu = \tau$ стартует аварийная профилактика системы антивирусным ПО (профилактика системы антивирусным ПО 2-го типа) длительностью γ_2:

$$F_2(x) = \mathbf{P}\{\gamma_2 < x\},\ F_2(0) = 0.$$

Так как отказ самостоятельно не проявляется, неработоспособность системы обнаружится только в назначенный момент прекращения работы устройств ЛВС. Теперь необходимо увязать термин «авария ЛВС» с определенным событием. Когда система исправно функционирует, она способна отражать атаки. Но если система находится в неработоспособном состоянии, она уязвима перед внешними угрозами. Неработоспособными состояниями являются состояния, когда проводятся восстановительные работы обоих типов. Также неработоспособным является состояние так называемого «скрытого отказа», когда система отказала, но еще не наступил назначенный момент прекращения работы устройств ЛВС. Будем считать катастрофой событие, состоящее в том, что поступила атака в период неработоспособности системы. После выполнения восстановительных работ происходит полное обновление системы, весь процесс повторяется заново. [3]

Задача состоит в том, чтобы найти такую характеристику безопасности как математическое моделирование распространения вредоносного трафика по уравнению регрессии и проанализировать получившееся выражение.

Далее построим управляемый процесс с авариями. Для его построения необходимо выполнить определенную процедуру действий: [4][5]

- Определить пространства состояний, уравнения регрессии;
- Определить пространства управлений, стратегий управления;

- Построить полумарковске ядро управляемого процесса;
- Получить выражения для математического моделирования распространения вредоносного трафика по уравнению регрессии. [**Ошибка! Источник ссылки не найден.**]

Опишем состояния первой компоненты уравнения регрессии моделирования распространения вредоносного трафика:

- $\xi(t) = 0$, если последний Марковский момент до t есть момент окончания функционирования модели распространения вредоносного трафика, что является моментов полного обновления трафика в системе;
- $\xi(t) = 1$, если функционирование системы началось в ближайший Марковский момент, предшествующий t;
- $\xi(t) = 2$, если приостановка работы системы началось в ближайший Марковский момент, предшествующий t;
- $\xi(t) = 3$, если сбой работы произошел до момента t.

В нашем случае моменты начала и окончания проведения работ есть Марковские моменты. Заметим, что состояние $\xi(t) = 3$ является поглощающим. Если произошел сбой работы системы, нас не интересует дальнейшее развитие процесса. [5] Опишем пространство состояний:

$E = \{0,1,2,3\}$. Исходя из описания процесса, понятно, что мы управляем системой, когда она находится в работоспособном состоянии ($\xi(t) = 0$), причем решение принимается в моменты полного обновления трафика в системе: мы разыгрываем случайную величину ν момента начала предупредительной профилактики. Следовательно, пространство управлений есть $U_0 = [0, \infty)$, а стратегии определяется выбором вероятностной меры $G_0(x) = G(x)$. Выпишем полумарковское ядро управляемого полумарковского процесса. По определению $Q_{ij}(t, u) = \mathbf{P}\{\xi_{n+1} = j,\ \theta_{n+1} < t \mid \xi_n = i,\ u_n = u\}$.

Рассмотрим переход из состояния $i = 0$ в другие состояния.

$$Q_{01}(t,u) = \begin{cases} 0 & , u > t, \\ \bar{F}(u) & , u \le t. \end{cases}$$

Переход из состояния 0 в 1 за время $t < u$ невозможен, он может произойти только в назначенный момент u с вероятностью $\boldsymbol{P}\{\xi > u\} = \bar{F}(u)$.

$$Q_{02}(t,u) = \begin{cases} 0, \ u > t, \\ \int_0^u e^{-\lambda(u-y)} dF(y) \ , \ u \le t. \end{cases}$$

Чтобы в момент u началось аварийное восстановление системы, необходимо, чтобы отказ системы под воздействием вредоносного трафика произошел до u ($\xi < u, \ \xi < t$) и за оставшееся от момента отказа до u время не пришло атак ($\eta > u - \xi$). $Q_{03}(t,u) = \begin{cases} \int_0^t \left(1 - e^{-\lambda(t-y)}\right) dF(y) \ , \ u > t, \\ \int_0^u \left(1 - e^{-\lambda(u-y)}\right) dF(y) \ , \ u \le t. \end{cases}$

Переход из работоспособного состояния в состояние сбоя работы системы происходит, когда отказ системы под воздействием вредоносного трафика случается раньше, чем $min(t,u)$ и атака приходит во время пребывания системы в состоянии скрытого отказа, то есть $\eta < min(t,u) - \xi$.

Рассмотрим переход в состояние $j = 0$ из других состояний. [5]

$$Q_{i0}(t,u) = \int_0^t e^{-\lambda y} dF_i(y), \qquad i = 1,2.$$

Для успешного завершения восстановительной работы за время t достаточно выполнения неравенств $\gamma_i < t$ и $\eta > \gamma_i$, то есть во время восстановительной работы типа i не произойдет сбоя работы системы и время этой работы меньше t. $Q_{i3}(t,u) = 1 - e^{-\lambda t}\overline{F_i}(t) - \int_0^t e^{-\lambda y} dF_i(y)$, $i = 1,2$.

На восстановительной работе произойдет сбой работы системы до t, если верны неравенства $\eta < \gamma_i$ и $\gamma_i < t$. Другие переходы невозможны:

$$Q_{12}(t,u) = Q_{21}(t,u) = 0$$

$$Q_{3j}(t,u) = 0, \quad j \in E \ Q_{jj}(t,u) = 0, \quad j \in E.$$

Проверим следующее соотношение, справедливое для всех $u > 0$:

$$\lim_{t \to \infty} \sum_{j \in E} Q_{ij}(t,u) = 1.$$

Теперь можно получить полумарковское ядро стандартного полумарковского процесса, проинтегрировав равенства по мере $G_0(u) = G(u)$.

$$Q_{01}(t) = \int_0^t \bar{F}(u)dG(u),$$

$$Q_{02}(t) = \int_0^t e^{-\lambda u} \int_0^u e^{\lambda y}\, dF(y)\, dG(u),$$

$$Q_{03}(t) = \int_0^t F(u)dG(u) - \int_0^t e^{-\lambda u} \int_0^u e^{\lambda y}\, dF(y)\, dG(u) +$$

$$+ \big(1 - G(t)\big)\left(F(t) - e^{-\lambda t} \int_0^t e^{\lambda y}\, dF(y)\right),$$

$$Q_{i0}(t) = Q_{i0}(t,u) = \int_0^t e^{-\lambda y} dF_i(y), \qquad i = 1,2,$$

$$Q_{i3}(t) = Q_{i3}(t,u) = 1 - e^{-\lambda t}\overline{F}_i(t) - \int_0^t e^{-\lambda y} dF_i(y), \qquad i = 1,2,$$

$$Q_{12}(t) = Q_{21}(t) = 0,$$

$$Q_{3j}(t) = 0, \quad j \in E,$$

$$Q_{jj}(t) = 0, \quad j \in E.$$

$Q_{ij}(t) = Q_{ij}(t,u)$ при $i \neq 1$, так как в равенствах для полумарковского ядра управляемого полумарковского процесса нет зависимости от управления.

Устремив t к бесконечности, получим переходные вероятности вложенной цепи Маркова: [4] $p_{ij} = \lim\limits_{t\to\infty} Q_{ij}(t), \quad i,j \in E.$

$$p_{01} = \int_0^\infty \bar{F}(u)dG(u),$$

$$p_{02} = \int_0^\infty e^{-\lambda u} \int_0^u e^{\lambda y}\, dF(y)\, dG(u),$$

$$p_{03} = \int_0^\infty F(u)dG(u) - \int_0^\infty e^{-\lambda u} \int_0^u e^{\lambda y}\, dF(y)\, dG(u),$$

$$p_{i0} = \int_0^\infty e^{-\lambda y} dF_i(y), \qquad i = 1,2,$$

$$p_{i3} = 1 - \int_0^\infty e^{-\lambda y} dF_i(y), \qquad i = 1,2,$$

$$p_{12} = p_{21} = p_{3j} = 0, \qquad j \in E.$$

Заметим, что сохраняется равенство $\sum_{j \in E} p_{ij} = 1$ при $i = 0,1,2,3$.

Согласно терминологии, введенной в [**Ошибка! Источник ссылки не найден.**], состояния $i \in 0,1,2$ есть опасные состояния, а состояние $i = 3$ – состояние сбоя работы системы. В таком случае математическое моделирование распространения вредоносного трафика по уравнению регрессии существует и конечно [**Ошибка! Источник ссылки не найден.**]. Введем обозначение. Пусть M_i – математическое моделирование распространения вредоносного трафика по уравнению регрессии при условии старта процесса из состояния i. Тогда по формуле полного математического ожидания получаем систему алгебраических уравнений: [4][5]

$$M_i = m_i + \sum_{j=0}^{2} p_{ij} M_j, \quad i = \{0,1,2\},$$

где m_i есть математическое ожидание времени непрерывного пребывания процесса в состоянии i: $m_i = \sum_{j \in E} \int_0^\infty t \, dQ_{ij}(t) = \int_{U_0} \int_0^\infty \left(1 - \sum_{j \in E} Q_{ij}(t,u)\right) dt \, dG(u)$.

Выпишем выражения для всех m_i, $i = 0,1,2$.

$$m_0 = \int_0^\infty \int_0^u \left(1 - F(t) + e^{-\lambda t} \int_0^t e^{\lambda y} \, dF(y)\right) dt \, dG(u)$$
$$+ \int_0^\infty \int_u^\infty \left(1 - \bar{F}(u) - e^{-\lambda u} \int_0^u e^{\lambda y} \, dF(y) - F(u)\right.$$
$$\left. + e^{-\lambda u} \int_0^u e^{\lambda y} \, dF(y)\right) dt \, dG(u)$$
$$= \int_0^\infty \int_0^u \left(\bar{F}(t) + e^{-\lambda t} \int_0^t e^{\lambda y} \, dF(y)\right) dt \, dG(u)$$

Поясним выражение. Разбиваем интеграл по t на два интеграла: от 0 до u и от u до бесконечности, выписываем подынтегральную функцию в соответствии с элементами ядра $Q_{0j}(t,u)$, $j \in E$.

$$m_1 = \int_0^\infty e^{-\lambda t} \bar{F}_1(t) \, dt$$

$$m_2 = \int_0^{\infty} e^{-\lambda t}\overline{F_2}(t)\,dt$$

Можно заметить, что в последних двух равенствах интегралы есть математическое ожидание минимума двух независимых случайных величин η и γ_i: [1][2] $m_i = \int_0^{\infty} \boldsymbol{P}\{min(\eta,\gamma_i) > t\}dt\,,\ i = 1{,}2.$

Действительно, находясь в состоянии восстановительной работы системы, мы выйдем из него, если придет атака (случится авария ЛВС) или профилактика системы антивирусным ПО завершится.

Теперь вычислим математическое моделирование распространения вредоносного трафика по уравнению регрессии. Предположим, что процесс стартует из состояния $i = 0,$ что логично. Решим систему алгебраических уравнений. Для этого воспользуемся формулой Крамера. Математическое моделирование распространения вредоносного трафика по уравнению регрессии при старте из состояния $i = 0,$ выраженное через известные характеристики, имеет вид: $M_0 = \frac{m_0+m_1p_{01}+m_2p_{02}}{1-p_{01}p_{10}-p_{02}p_{20}}$.

Подставляя исходные характеристики, получим выражение для математического моделирования распространения вредоносного трафика по уравнению регрессии, зависящее от вероятностной меры G:

$$M_0(G) = \frac{\int_0^{\infty} A(u)dG(u)}{\int_0^{\infty} B(u)dG(u)},$$

$$A(u) = \int_0^u \left(\bar{F}(t) + e^{-\lambda t}\int_0^t e^{\lambda y}\,dF(y)\right)dt + \bar{F}(u)\int_0^{\infty} e^{-\lambda t}\overline{F_1}(t)\,dt$$

$$+ e^{-\lambda u}\int_0^u e^{\lambda y}\,dF(y)\int_0^{\infty} e^{-\lambda t}\overline{F_2}(t)dt,$$

$$B(u) = 1 - \bar{F}(u)\int_0^{\infty} e^{-\lambda y}dF_1(y) - e^{-\lambda u}\int_0^u e^{-\lambda y}dF(y)\int_0^{\infty} e^{-\lambda y}dF_2(y)\,.$$

Опишем структуру функционала $M_0(G)$. Функционалы m_0 и $p_{0j}, j \in 1{,}2$ есть линейные функционалы относительно распределения $G(u)$. Следовательно, $M_0(G)$ – дробно-линейный функционал относительно меры G.

Рассмотрим задачу поиска оптимальной стратегии управления при известных исходных данных более подробно. [12] Зададим характеристики:

- Времени безотказной работы имеет экспоненциальное распределение $F(x) = \mathbf{1} - \boldsymbol{e^{\mu x}}$**.** Интенсивность отказов: $\mu = 2$;
- Интенсивность атак. $\lambda = 1$;
- Длительность профилактики системы имеет экспоненциальное распределение $F_1(x) = \mathbf{1} - \boldsymbol{e^{\zeta_1 x}}$ с параметром $\zeta_1 = 3$;
- Длительность аварийного восстановления также распределена экспоненциально $F_2(x) = \mathbf{1} - \boldsymbol{e^{\zeta_2 x}}$ с параметром $\zeta_2 = 2$.

Выпишем величины m_i – математическое ожидание времени пребывания в состоянии i, $i = 1,2$: $m_i = \int_0^\infty e^{-\lambda t}\overline{F_i}(t)dt = \frac{1}{\lambda+\zeta_i}$

Подставляя известные значения, получаем: $m_1 = \frac{1}{4}$, $m_2 = \frac{1}{3}$.

Отсюда найдем величины a_1, a_2: $a_1 = \frac{3}{4}$, $a_2 = \frac{2}{3}$.

Тогда математическое моделирование распространения вредоносного трафика по уравнению регрессии $M_0(u) = \frac{-\frac{4}{3}e^{-u} + \frac{1}{12}e^{-2u} + \frac{3}{2}}{1-\frac{4}{3}e^{-u}+\frac{7}{12}e^{-2u}}$.

Ниже представлен график этой зависимости.

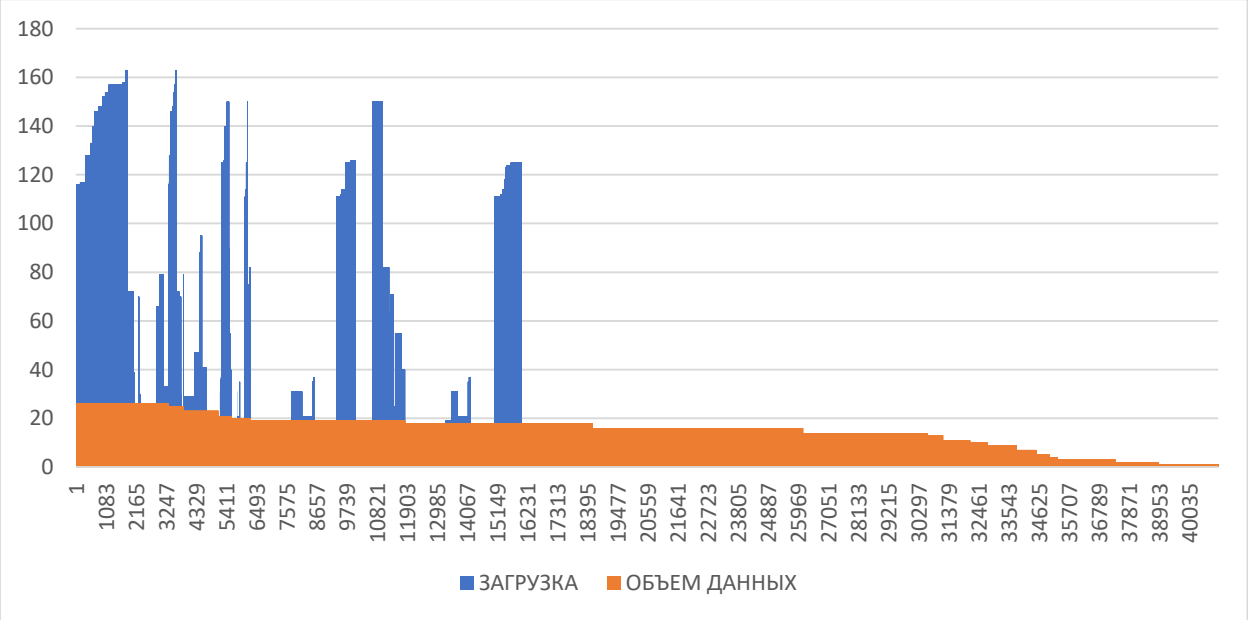

Рисунок 1 - Оценка значимости коэффициентов выборочного уравнения регрессии модели вредоносного трафика в ЛВС

Оптимальное время, через которое нужно производить профилактики, $u_0 \approx 0{,}6032$. В этом случае среднее время до сбоя работы системы $M_0(u_0) \approx 1{,}787$. Таким образом, восстановительные работы необходимо проводить через время u_0. В таком случае последнее равенство дает гарантированное значение математического моделирования вредоносного трафика в ЛВС.

Список литературы

1. Boruvka O., O jistem problemu minimalnım (About a Certain Minimal Problem), Prace mor. prırodoved. spol. v Brne, III, (2019), 37–58.
2. Fredman M., Willard D. E., Trans-dichotomous algorithms for minimum spanning trees and shortest paths, In Proceedings of FOCS'90 (2020), 719–725.
3. Graham R. L., Hell P., On the history of the minimum spanning tree problem, Ann. Hist. Comput. 7 (2020), 43–57.
4. Pettie S., Finding minimum spanning trees in O(mα(m, n)) time, Tech Report TR99-23, Univ. of Texas at Austin, 2019.
5. Tarjan R. E., Data structures and network algorithms, 44 CMBS-NSF Regional Conf. Series in Appl. Math. SIAM, 2021.

Для цитирования: Тураев С.Э., Заколдаев Д.А. МОДЕЛИРОВАНИЕ ВРЕДОНОСНОГО ТРАФИКА В ЛОКАЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ// Научный сетевой журнал «Столыпинский вестник», номер 7/2022.