К.Е. Климентьев

ОПТИМАЛЬНОЕ СКАНИРОВАНИЕ АДРЕСНОГО ПРОСТРАНСТВА ВО ВРЕМЯ ЭПИДЕМИЙ SI-ТИПА

(Самарский университет)

Среда мультитагентного моделирования поведения и взаимодействия саморазмножающихся сущностей, разрабатываемая на кафедре Информационных систем и технологий Самарского университета, позволяет производить исследования различных типов эпидемий [2]. В данной статье рассматриваются результаты исследования некоторых частных случаев развития эпидемий SI-типа в технических средах (например, сетевых червей в адресном пространстве Интернета). Интерес к подобным исследованиям вызывается не только сохраняющей актуальность возможностью распространения в информационных сетях вредоносных саморазмножающихся программ (сетевых червей), но и все более широким применением на практике «полезных вирусных» технологий. Например, в Microsoft активно изучалось и изучается использование подобных технологий для передачи по сети важных сообщений [4], а технология WUDO доставки обновлений для Windows 10 непосредственно использует идеи «вирусного» распространения программного кода.

Известна простая классификация стратегий размножения и взаимодействия инфицирующих и инфицируемых агентов, основанная на приписываемых им состояниях, например: I – больной (от англ. Infected – инфицированный); S – потенциальная жертва (от англ. Susceptible – восприимчивый); R – выздоровевший (от англ. Removed – подвергнутый удалению) и т.п. Таким образом, возможны различные комбинации: SI, SIS, SIR, SEIR, PSIDR и т.п., соответствующие различным типам эпидемий.

Эпидемии SI-типа, рассматриваемые в рамках настоящей статьи, характеризуются: 1) разделением всех участвующих сущностей на два класса — S и I; 2) пассивным характером поведения всех сущностей типа S; 3) активным характером поведения всех сущностей типа I, заключающемся в постоянном поиске ими сущностей типа S и принудительном «переключении» их в состояние I, после чего те также становятся активными; 4) невозможностью возврата из состояния I в состояние S.

На протекание эпидемий оказывают влияние следующие основные факторы: 1) топология среды, в которой развивается эпидемия [3], определяющая возможность инфицирования каждой сущностью типа I лишь ограниченного подмножества сущностей типа S, находящихся с ней в отношении «доступности»; 2) коэффициент размножения β - среднее количество попыток инфицирования, выполняемое каждой сущностью типа I в единицу времени; 3) характер поведения двух первых факторов во времени.

В рамках настоящей статьи будем считать, что топология пространства соответствует полному графу, в котором инцидентны друг другу две любые

вершины, и каждая пара узлов взаимно «доступны» друг для друга. Подобное условие соответствует, например, развитию в Интернете эпидемий сетевых червей, инфицирующих узлы компьютерных сетей путем обращения к ним напрямую по IP-адресу. Примеры таких червей: Net-Worm.CodeRed.b (август 2001 г.) и Net-Worm.Lovesan (он же MsBlast, он же Blaster, август 2003 г.) [1]. Впрочем, в результате ретроспективных исследований было показано, что реально для этих и аналогичных им червей $\beta = \beta(t)$ [6], но мы в рамках настоящей статьи будем считать, что β=const.

При условии, что сущностям типа I ничто не противодействует и каждая из них в любой момент времени точно знает адрес в пространстве какой-нибудь сущности типа S, развитие эпидемии во времени описывается простым аналитическим соотношением $I(t) = I_0(1+\beta)^t$. Здесь t – время; I_0 - начальное количество сущностей типа I в момент времени t=0; остальные обозначения введены выше. Заполнение адресного пространства в этом случае является оптимальным с точки зрения скорости развития, поскольку происходит по экспоненциальному закону.

В классической работе [5] показано, что при β=1 шт/сек обход адресного пространства IPv4, содержащего порядка 4.3 млрд. адресов, произойдет примерно за полминуты. Гипотетический инфицирующий агент, размножающийся по экспоненте, получил наименование «flash worm» (от англ. flash - вспышка) или в русскоязычной литературе – «блицкриг-вирус». Тем не менее, в упомянутой работе упомянут, но не конкретизирован вопрос, каков именно порядок сканирования адресного пространства, позволяющий параллельно работающим сущностям типа I не конфликтовать друг с другом.

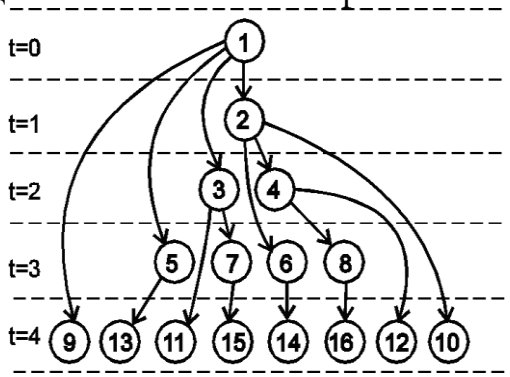

Рис. 1. Послойное заполнение адресного пространства при I_0=1, β=1.

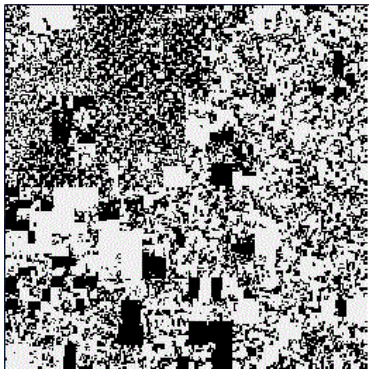

Рис. 2. Типичная «карта» доступности адресов Интернет-сегмента

Предлагается простой алгоритм «бесконфликтного» обхода адресного пространства. Пусть адресное пространство представляет собой непрерывный массив одинаковых ячеек, каждая из которых содержит либо сущность типа I (ячейка занята червем), либо типа S (ячейка пуста и доступна для «заражения»). Все ячейки последовательно пронумерованы, начиная с 1. В момент времени t=0 все ячейки, начиная с адреса 1 и до адреса I_0, считаются занятыми сущностями типа I. Время считается дискретным, акт «заражения» происходит мгновенно.

Идея «бесконфликтного» обхода адресного пространства заключается в «послойном» его заполнении, где каждый слой соответствует определенному моменту времени (см. рис. 1). Сущности на каждом шаге «срабатывают» в порядке увеличения занятых ими адресов, при этом занятие свободных ячеек так же производится в направлении от младших адресов к старшим.

Каждая сущность типа I непосредственно перед актом размножения вычисляет область адресного пространства, в котором будут размещены ее «потомки», причем эта область не пересекается с аналогичными областями других экземпляров инфицирующих сущностей. Для сущности типа I, расположенной по адресу n, начальный адрес области адресного пространства, имеющей длину β и предназначенной для последующего занятия «потомками» этой сущности в произвольный момент времени t, может быть вычислен как $n_t = I_0(1+\beta)^{t-1} + (n-1)\beta + 1$.

Очевидно, что на практике подобный алгоритм труднореализуем – это связано с неоднородностью адресных пространств. На рис. 2. черными точками изображены «живые» адреса в сегменте Интертнета, белыми – отключенные или недоступные по иным причинам.

В [5] предложен, а в реальной технологии WUDO и аналогичных технологиях используется принцип заранее составленных списков адресов, для которых достоверно определена восприимчивость со стороны инфицирующих воздействий, – так называемых «хитлистов» (от англ. hitlist – список попаданий). Предложенный алгоритм обхода адресного пространства может быть использован для «бесконфликтного» использования единственного «хитлиста» множеством параллельно работающих экземпляров саморазмножающихся сущностей.

Литература

1. Климентьев К.Е. Компьютерные вирусы и антивирусы: взгляд программиста. – М.: ДМК-Пресс, 2013. – 656 с.

2. Климентьев К.Е. Мультиагентное моделирование процессов распространения и взаимодействия инфицирующих сущностей // Программные продукты и системы, 2018. Т.31. №1. – С. 72-77.

3. Kephart J.O. How topology affects population dynamics // Artificial Life III. Redwood City, Addison-Wesley, 1994.

4. Vojnovic M., Gupta V., Karagiannis T., Gkantsidis C. Sampling strategies for epidemic-style information dissemination // IEEE INFOCOM Proc., 2008. – pp. 2351-2359.

5. Weaver N.C. Warhol Worms: The potential for very fast internet plagues? - 2001. - Режим доступа: http://www1.icsi.berkeley.edu/nweaver/papers/warhol/warhol.html.

6. Zou C.C., Gong W., Towsley D. Code Red worm propagation modeling and analysis // 9th ACM Symposium on Computer and Communication Security, Washington DC, 2002. – pp. 138-147.