Опубликовано: Климентьев К.Е. Применение ГИС-технологий при исследовании распространения вредоносных программ // В сб. «Геоинформационные технологи в проектировании и создании корпоративных информационных систем. Межвузовский научн. сборник». – Уфа: изд-во УГАТУ. – 2012. – С. 130-133.

ПРИМЕНЕНИЕ ГИС-ТЕХНОЛОГИЙ ПРИ ИССЛЕДОВАНИИ РАСПРОСТРАНЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ

Климентьев К.Е., Самарский государственный аэрокосмический университет, г. Самара

На кафедре информационных систем и технологий СГАУ силами студентов и преподавателей ведется работа по созданию системы моделирования эпидемий различных видов компьютерных вирусов. Система содержит и использует элементы ГИС.

1. Актуальность создания системы.

Интерес к исследованию распространения вредоносного программного обеспечения возник в 2001-2002 гг. после масштабных эпидемий сетевых червей типа Code Red II, заражавших в течение суток до 0.5 млн компьютеров во всем мире. В настоящее время размеры «ботнетов» (сетей зараженных компьютеров, контролируемых злоумышленниками) могут достигать нескольких миллионов. Работы, посвященные исследованию распространения вредоносных программ в сетях, традиционно решают следующие задачи:

- исследование поведения вредоносной программы в заданных условиях;
- параметрическая идентификация модели поведения, т.е. оценивание условий распространения по зарегистрированным характеристикам поведения.

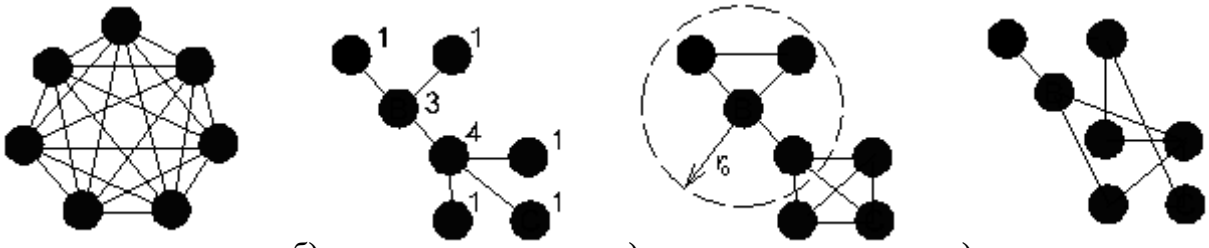
При этом чаще всего объектом рассмотрения является асимптотическое поведение вредоносных программ – т.е. на интервале времени, стремящемся к бесконечности, и в сетях с однородной структурой. При этом игнорируются важные факторы:

- неоднородность сети;
- непостоянство условий распространения (например, структуры сети) во времени.

Эти факторы могут быть учтены введением в модели распространения вредоносных программ «геометрических» и «географических» аспектов.

2. Описание и анализ проблемной области.

2.1. Графовые модели сетей. Самые крупные и быстро развивающиеся эпидемии вызываются сетевыми Интернет-червями. Червь такого типа способен атаковать любой компьютер, имеющий собственный IP-адрес, вне зависимости от его географического местоположения. Этой ситуации соответствует модель в виде «полного» (или «гомогенного») графа, каждый узел которого связан со всеми остальными. Такой граф из N узлов имеет $N(N-1)/2$ ребер, причем степень каждого узла равна $N-1$ (см. рис. 1, а).



а) б) в) г)

Рис. 1. Примеры графов, применяемых для моделирования компьютерных сетей:
а) «Гомогенный »; б) «Scale free»; в) «геометрический»; г) граф Радо

В основе моделей «медленных» эпидемий, вызываемых почтовыми червями и файловыми вирусами, лежат графы «безмасштабного» («scale-free») вида. Это графы, для которых лишь немногие узлы имеют большое количество соседей. И наоборот, степень основной массы узлов невелика. Считается, что граф «безмасштабен», если доля $P_k = N_k / N$ его узлов со степенью k примерно равна $P_k \approx k^{-\gamma}$, где $2 \le \gamma \le 3$. На рис. 1,б лишь один узел имеет четырех и один - трех, зато все остальные - по единственному соседу.

Наиболее сложным является изучение эпидемий мобильных червей, распространяющихся на небольшие расстояния при помощи беспроводных интерфейсов типа WiFi или Bluetooth. Сети, характеризующие систему связей в этом случае, называются «специальными» («ad hoc»). Наиболее близкими абстракциями для подобных сетей считаются «случайные» графы: граф Радо (RRG) и граф со случайной геометрией (RGG). В построении графа со случайной геометрией участвуют «координаты» узлов. Например, если считать, что граф размещен в единичном квадрате, тогда каждому узлу приписываются координаты, представляющие собой равномерно распределенные на интервале [0..1] случайные величины. Затем, для каждой пары узлов с индексами i и j рассчитывается «расстояние» r_{ij}, и ребро между узлами проводится в том случае, если расстояние между ними не превышает заранее назначенного «радиуса» r_0. Особенностью топологии подобных графов является значительное количество «тесных миров». Случайный же граф Радо получается из «полного» графа, каждое ребро которого остается с заранее выбранной вероятностью p и удаляется с вероятностью $1-p$.

2.2. Модели размножения. Модели размножения вредоносных программ восходят к прототипам, разрабатываемым с начала XX в. для изучения эпидемий и эпизоотий среди людей и животных. В общем случае, они описываются системами дифференциальных уравнений вида:

$$\left\{\begin{array}{l} \partial I / \partial T = F(I,S,R) \\ \partial S / \partial T = G(I,S,R) \\ \partial R / \partial T = H(I,S,R) \end{array}\right\},$$

где T – время; I – множество инфицированных узлов сети; S – множество неинфицированных узлов; R – множество «вылеченных» или «выздоровевших» узлов; F, G и H – некоторые функции. Наиболее простые модели на «гомогенных» графах (например, с уравнениями Мальтуса и Фергюльста) имеют аналитические решения, однако уже модели типа Кермака-МакКендрика могут быть исследованы только численно. Для исследования распространения в сетях «безмасштабного» и, особенно, «случайного» видов целесообразно применение методов имитационного моделирования. В этом случае и вредоносные программы, и процессы обнаружения/лечения моделируются в дискретном времени при помощи перемещения вдоль ребер и взаимодействия «мобильных агентов» - специальных меток различных типов.

3. Учет «геометрического» и «географического» аспектов.

В разрабатываемой системе предусматривается использование не только «полностью случайных» графов, но и графов, в основе которых лежит пространственное размещение узлов на конкретных территориях. Исходными данными для построения моделей является:

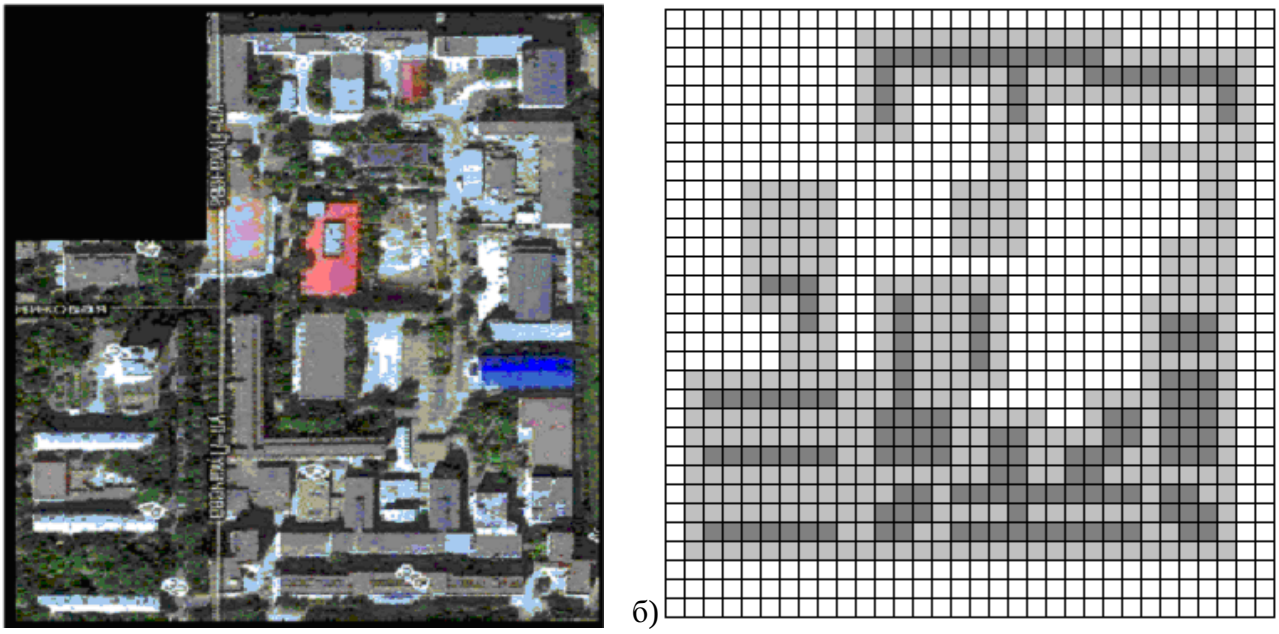


Рис 2. Этапы построения случайного графа, привязанного к территории кампуса СГАУ: а) исходная карта местности; б) «карта плотностей» уязвимых узлов

- графическое представление территории (интерьер помещения, карта местности, схема и т.п.) с нанесенной на нее сеткой – см. рис. 2,а;
- априорные сведения о средней «плотности» присутствия уязвимых узлов (например, данные для крупнейших городов планеты – см. на сайте www.securelist.com);
- технические характеристики сети (радиус взаимодействия r_0).

Результатом сопоставления этих данных, выполняемого системой, является «карта плотностей» (см. рис. 2,б). Последним шагом построения графа служит генерация узлов и ребер в соответствии с алгоритмом «геометрических»

графов, применяемым каждой локации с учетом характерной для него «плотности». Пример результата приведен на рис. 3. Таким образом, становится возможным исследование распространения вредоносных программ на конкретных территориях.

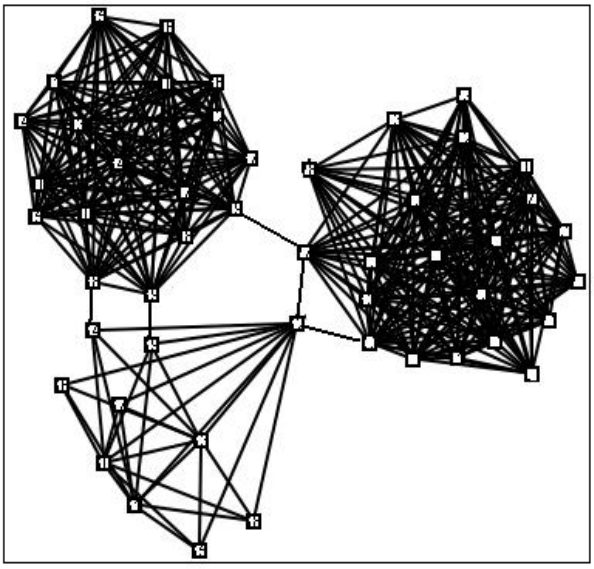


Рис. 3. Итоговый «географический» граф

Технические характеристики исследовательского прототипа системы:

- «сетка» для указания плотности уязвимых узлов - 32×32;
- количество узлов сети – до 10000 (планируется довести до 10^6);
- графы - «гомогенный», «scale free», RRG, RGG, «географический»;
- поиск в графе – случайный, адресный, «морской бой», «контратака»;
- способы взаимодействия мобильных агентов – удаление, безусловное замещение, условное замещение, «вакцинация».

Список литературы:

1. Блинкова О.А., Климентьев К.Е. Упоров А.Н., Шубин Ю.В. Интернет для географов. – Харьков: Kharkiv University Press, 2003.
2. Колчин В.Ф. Случайные графы. – М.:Физматлит, 2004.
3. Братусь А.С., Новожилов А.С., Платонов А.П. Динамические системы и модели биологии. - М.: Физматлит, 2010.
4. Serazzi G., Zanero S. Computer virus propagation models // Performance tools and applications to networked systems, vol. 2965 – Berlin, 2004.
5. Nekovee M. Worm epidemics in wireless ad hoc networks // New Journal of Physics, vol. 9/189 – 2007.