УДК 004.49

П.В. Збицкий

Функциональная сигнатура компьютерных вирусов

Рассмотрены проблемы детектирования сложных полиморфных и метаморфных компьютерных вирусов. Предложен и детально рассмотрен новый метод построения функциональной сигнатуры вируса на основе последовательности системных вызовов.
Ключевые слова: вирус, сигнатура, система, защита информации.

Информация и компьютерные системы играют существенную роль в жизни современного общества. Проблема обеспечения подлинности и целостности передаваемой и обрабатываемой информации, а также повышение устойчивости систем обработки информации является необходимым условием успешной экономической деятельности. Этому способствует внедрение существующих методов и подходов [1, 2] к обеспечению информационной безопасности. Однако в то же самое время колоссальное развитие получила индустрия вредоносного программного обеспечения (ПО): рассылка спама, организация атак, направленных на отказ в обслуживании онлайн-сервисов коммерческих предприятий и государственных учреждений, кража идентификационных данных платежных систем и т.д. Таким образом, выработка новых методов противодействия распространению вредоносного ПО является важной задачей при обеспечении информационной безопасности государства, общества, личности.

В настоящее время существует несколько подходов к детектированию вирусов: сигнатурный поиск, эвристический и поведенческий анализ. Сигнатура, как правило, представляет собой некую последовательность байт, идентифицирующую вирусный код. Эвристика направлена прежде всего на выявление неизвестных вирусов и, как правило, включает проверку «корректности» исполняемого файла, вычисление энтропии кода или статистического распределения инструкций процессора, а также выявление последовательностей инструкций, характерных для вирусов. Поведенческий анализ в современных антивирусных решениях представляет собой средство контроля за процессами и состоянием операционной системы. В случае нарушения «нормального» функционирования системы, процесс, вызвавший нарушения блокируется.

Фактически же детектирование *известных* вирусов в коммерческом антивирусном ПО осуществляется на основе классических сигнатур. Однако, как показано в [3–5], выделение такой сигнатуры не всегда возможно. На практике антивирусные компании столкнулись с этим явлением несколько лет назад, когда в антивирусные продукты, помимо сигнатурного сканера, начали включать средства контроля приложений и поведенческие блокираторы.

В последние годы развивается идея функциональных сигнатур компьютерных вирусов. В работе [6] рассмотрен подход так называемой кодовой нормализации – применение методов деобфускации/оптимизации для выделения минимальной функционально эквивалентной формы последовательности инструкций. Основным недостатком предложенного метода является предположение о корректной дизассемблируемости нормализируемого кода, сделанное на основании того, что вирус при построении новой копии должен сам себя дизассемблировать. В работах [3, 4] рассмотрены примеры метаморфных вирусов, которым не требуется производить самодизассемблирование при распространении и генерации новых копий.

Автором предлагается новый подход к выявлению функциональных сигнатур, основанный на анализе последовательностей системных вызовов процесса.

Для проведения исследования написано специальное ПО мониторинга системных вызовов процесса с возможностью интерактивного изменения передаваемых параметров и изменения кода возврата системного вызова в операционной системе Windows XP. Были проанализированы последовательности системных вызовов следующего вредоносного ПО:

1. Метамофные вирусы:
   – MetaPHOR, Evol, Zmist.
2. Сетевые черви, в том числе полиморфные:
   – Mimail, Tanatos, Bagle…
3. Разнообразные троянцы:
   – Zeus, Pinch…

Ниже представлены результаты эксперимента.

Для одиночных программ (червей/троянцев) последовательность вызовов за некий период работы может являться сигнатурой. Обработка лога системных вызовов для этого случая включает в себя отсечение активности операционной системы по созданию или за-

вершению процесса (при мониторинге с момента запуска или до завершения процесса), а также сворачивание вызовов, произведенных в цикле.

Для файловых вирусов выделение сигнатуры требует дополнительной обработки – разделения вызовов вируса и зараженной программы. Данное действие можно выполнить путем «вычитания» лога вызовов незараженной программы из лога вызовов зараженной.

Кроме того, метод составления сигнатуры по системным вызовам позволяет выполнить объединение разных штаммов вредоносного ПО в одну сигнатуру. Например, в случае червя Bagle, штаммы Bagle.f, Bagle.g, Bagle.h, Bagle.k, штаммы Bagle.i, Bagle.j, штаммы Bagle.n, Bagle.o, Bagle.p, agle.q, Bagle.r, штаммы Bagle.a, Bagle.t настолько похожи по системным вызовам между собой, что позволяет говорить о простой перекомпиляции исходного кода червя. Для количественной оценки отличий в последовательностях системных вызовов использовалась введенная мера $\mu(x,y)$ – число разных блоков системных вызовов в логах программ x и y. Блок – последовательность вызовов, которая присутствует в первом логе, но отсутствует во втором или наоборот. $\mu(x,y)=0$ означает, что программы x и y в своей работе используют одинаковые системные вызовы. Очевидно, что $\mu(x,y)=0$. Вычисление $\mu(x,y)$ реализуется алгоритмом поиска подстрок в строке.

Основные достоинства функциональных сигнатур на основе системных вызовов:

1. Возможность автоматического выделения сигнатуры

– сокращение времени 0-day;

– получение сигнатуры в клиентском приложении (при должном развитии методов эвристического/поведенческого анализа).

2. Приемлемый размер сигнатур:

– 2-х байтовый унифицированный номер системного вызова;

– 200–600 системных вызовов на сигнатуру.

Однако метод сигнатур на основе последовательностей системных вызовов (трасс выполнения) подвержен достаточно простой атаке – вставке мусорных системных вызовов. По аналогии с классическими сигнатурами и вставкой мусорных инструкций напрашивается решение – разряженные сигнатуры. Следовательно, опять возникает проблема многократных сравнений трасс выполнения, чтобы отсеять мусор. Это не всегда можно сделать в автоматическом режиме.

Таким образом, функциональная сигнатура на основе последовательности системных вызовов (трасс выполнения) может быть использована для детектирования сложных полиморфных и метаморфных вирусов, для которых построение классической сигнатуры затруднено или невозможно. Кроме того, методы автоматического выделения сигнатур на стороне конечного пользователя на основе данных эвристического или поведенческого анализа могут стать основой антивирусного ПО нового поколения.

Литература

1. Белов Е.Б. Основы информационной безопасности: учеб. пособие для вузов / Е.Б. Белов, В.П. Лось, Р.В. Мещеряков, А.А. Шелупанов. – М.: Горячая линия – Телеком, 2006. – 544 с.

2. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах. Екатеринбург: Изд-во УрГУ, 2003. – 328 с.

3. Filiol E. Metamorphism, Formal Grammars and Undecidable Code Mutation // PWASET (World Academy of Science, Engineering and Technology). – 2007. – Vol. 20. – P. 1–7.

4. Zbitskiy P.V. Code mutation techniques by means of formal grammars and automatons // Journal in Computer Virology. – Paris: Springer, 2009. – Vol. 5, Num. 2. – P. 88–99.

5. Варновский Н.П. О применении методов деобфускации программ для обнаружения сложных компьютерных вирусов // Информационное противодействие угрозам терроризма. – М.: ФГПУ НТЦ, 2006 – С. 107–126.

6. Bruschi D. Using Code Normalization for Fighting Self-Mutating Malware / D. Bruschi, L. Martignoni, M. Monga // Technical Report. – Milan: University, 2006. – 14 p.

Збицкий Павел Владимирович
ГОУ ВПО «Южно-Уральский государственный университет»,
аспирант каф. цифровых радиотехнических систем.
Науч. рук.: д.ф.-м.н., проф. А.В. Рожков
Эл. адрес: pavel.zbitskiy@gmail.com

Pavel V. Zbitskiy

Functional signature of computer viruses

Paper describes existing problems of polymorphic and metamorphic viruses detection. New approach of building functional signature based on system calls sequence is suggested and discussed.

Keywords: virus, signature, system, information security.