УДК 621.383

НЕЧЕТКОЕ ПРЕДСТАВЛЕНИЕ КАК СРЕДСТВО ОПИСАНИЯ ПОВЕДЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ

Р. А. Пышкин д. т. н. А. А. Каргин

Донецкий национальный университет
Украина, г. Винница
r.pyshkin@donnu.edu.ua

Исследована возможность использования нечеткой логики, нечеткологических оптических информационных технологий в процессе эвристического обнаружения вредоносного программного обеспечения в больших массивах данных в масштабе времени, близком к реальному.

Ключевые слова: вредоносный объект, нечеткое множество, нейронная сеть, вирус, оптический.

Информация, являющаяся собственностью государства или отдельных лиц, представляет ценность и является целью злоумышленников, которые стремятся ее похитить или повредить. Следовательно, любой информационный ресурс нуждается в защите. Особенно это актуально в информационно-коммуникационных системах, где, как правило, циркулируют большие и сверхбольшие объемы данных, представляющих государственную и коммерческую тайну [1].

Традиционная реализация алгоритмов обработки информации на основе микропроцессорных средств [2] подразумевает последовательную загрузку данных при синтезе управляющих решений. При достаточно большом количестве параметров, которыми оперируют, становится практически невозможным принятие необходимых решений в масштабе времени, близком к реальному.

Если применить оптические информационные технологии [3] для построения нечеткой системы, то можно распараллелить процедуры обработки в нечетком алгоритме за счет использования в качестве носителя информации оптического потока, а также уменьшить время параллельной обработки за счет практически мгновенного выполнения математических преобразований оптическим процессором.

Целью работы является создание математической модели оптоэлектронного процессора [7] нечетких множеств в среде Matlab, для эмуляции простейшей системы, служащей для эвристического анализа вредоносных объектов.

Для обнаружения вредоносных программ существуют два способа анализа – статистические и динамические. Статистические методы изучают текст программы, а динамические – программы во время их работы [4], следовательно, применение нечеткологических оптических средств и систем на их основе будет оправдано для ускорения процесса детектирования вредоносного программного обеспечения. В настоящее время растет интерес к динамическим методам анализа, в том числе поведенческим. Поведение многих вредоносных программ типично: например, пересылка данных пользователя, генерация трафика, саморассылка и саморепликация. Рост объемов вредоносного программного обеспечения ставит задачу автоматизации процесса поведенческого анализа, приближения времени обработки и детектирования практически к мгновенным значениям, особенно при работе с очень большими объемами данных.

Для обнаружения вредоносного программного обеспечения используются эвристические методы – совокупность исследовательских методов, способов по обнаружению ранее не известного [5]. Один из таких методов основан на нейронных сетях. Основным материалом для обучения нейронной сети является набор *n*-грам (множество длиной несколько байт), который указывает на заражение [5].

Детектирование вредоносного кода с использованием эвристики проходит три фазы: фаза декодирования, фаза исследования и фаза оценки. Первые две фазы относятся к технической части, а последняя – к аналитической. Назначение этапа декодирования – эмуляция требуемого количества инструкций, необходимых вирусу. Фаза исследования – эмуляция хотя бы единожды всех достигнутых

в программе секций кода, которые могут содержать вирусы. Предназначение оценки – анализ подозрительных действий, которые были обнаружены во время предыдущих этапов, для определения инфицированности системы, путем составления списка всех наблюдаемых поведений программы. Для этого используется база правил нечетких представлений.

Так как поведение вредоносного ПО в зараженной системе может вызывать целую цепь событий, каждое из них содержит следующую информацию: ID-события, объект события, субъект события, вызываемая функция ядра, параметры (значение реестра, путь к файлу, IP-адрес) и статус действия. Все это представляет из себя множество параметров, которые служат для построения классификации [5], на основе большого количества объектов из базы правил (данных), содержащей множества вредоносных действий.

При создании математической модели работы предполагаемой нечеткой системы в среде Matlabиспользовался компонент FuzzyLogicToolbox. За основу были взяты ключевые особенности поведения вредоносного ПО, такие как запись данных, пересылка, изменение кода зараженной программы, изменение реестра, удаление файлов и т. д., они использовались в качестве входных переменных системы. Затем для этих переменных были выбраны соответствующие термы и функции принадлежности, диапазон изменения входных переменных. Создана база правил для нечеткой системы, опирающаяся на характеристики поведения вредоносного ПО. Результатом работы модели является вероятность того, что некоторый объем исследуемых системой данных содержит признаки вредоносных программ. На рисуке представлен результат работы моделируемой системы в FIS-редакторе среды Matlab.

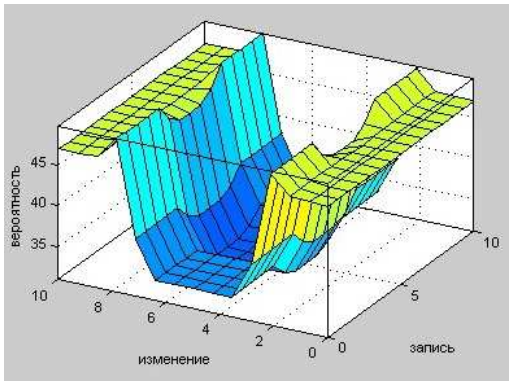

График поверхности

Полученная модель очень проста, использует всего четыре входные переменные, такие как запись, изменение, удаление, пересылка, а в качестве выходных данных – вероятность инфицированности, на основе небольшой базы правил. Таким образом, при достаточно большом количестве входных данных и соответственно гораздо большей базе правил, в реальных условиях становится оправданным применение нечетких оптических информационных технологий [7] для ускорения процесса обработки больших массивов информации в масштабе времени, близком к реальному.

ИСПОЛЬЗОВАННЫЕ ИСТОЧНИКИ

1. Гайворонский М. В., Новиков О. М. Безопасность информационно-коммуникационных систем.– Винница: Издательская группа BHV, 2009.

2. Мелихов А. Н. Боронец В. Д. Проектирование микропроцессорных средств обработки нечеткой информации. – Ростов-на-Дону: Издательство Ростовского университета, 1990.

3. Акаев А. А., Майоров С. А. Оптические методы обработки информации. – Москва: Высшая школа, 1988.

4. Молдавская А. В., Рувинская В. М. Байесовские сети как средство представления сценариев поведения вредоносных программ // Труды XV МНПК «Современные информационные и электронные технологии».– Украина, г. Одесса.– Т. I.– 2014.– С. 125.

5. Рувинская В.М., Беркович, Е. Л., Лотоцкий А. А. Эвристические методы детектирования вредоносных программ на основе сценариев // Искусственный интеллект.– 2008.– № 3.– С.197–207.

6. SecureList – Новые описания детектируемых объектов. //Режим доступа: http://www.securelist.com

7. Пышкин Р. А., Данилов В. В. Оптоэлектронные нечеткологичесие устройства и защита информации // Труды XV МНПК «Современные информационные и электронные технологии».– Украина, г. Одесса.– Т. I.– 2014.– С. 133.

R. A. Pyshkin, A. A. Kargin

Fuzzy representation as a means of describing the behavior of malicious programs.

The article focuses on the possibility of using fuzzy logic, fuzzy logic optical information technologies in the process of heuristic detection of malicious software in large amounts of data in near real time.

Keywords: *malware, fuzzy sets, neural network, virus, optical*.