*Комашинский Д. В., Котенко И. В., Шоров А. В.*

# ОБНАРУЖЕНИЕ ВРЕДОНОСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НА БАЗЕ МЕТОДОВ ИНТЕЛЛЕКТУАЬНОГО АНАЛИЗА ДАННЫХ

*Санкт-Петербург, ЗАО "Аркадия", СПИИРАН*
*komashinskiy@comsec.spb.ru, ivkote@comsec.spb.ru, ashorov@comsec.spb.ru*

Проблема противодействия вредоносному программному обеспечению (ПО) продолжает оставаться актуальной, несмотря на определенные успехи компаний-производителей антивирусного программного обеспечения в создании технологических решений раннего детектирования эпидемий, углубленного анализа потенциально вредоносных объектов, быстрой поставки баз обновлений. Основной ее причиной является *недостаточно высокое качество методов эвристического детектирования, использующихся для обнаружения ранее неизвестных экземпляров вредоносных программ*.

Представленная на рассмотрение работа посвящена *применению методов интеллектуального анализа данных (Data Mining) для построения средств эвристического детектирования*. Помимо раскрытия общей постановки задачи и анализа результатов предыдущих релевантных исследований в работе представляется собственная точка зрения на возможные пути решения проблемы детектирования вредоносного ПО средствами эвристического анализа. *Предлагаемый подход* основан на скрытном процессе циклического интерактивного сбора поведенческой информации, который в совокупности с результатами обобщения доступной статической информации, извлекаемого из файлового контейнера исследуемого объекта, облегчает принятие решения о степени его опасности. Также при подготовке данного подхода внимание было уделено вопросам интегрированного использования различных методов интеллектуального анализа данных для различных классов вредоносного ПО. В работе реализовано и исследовано семейство различных методов интеллектуального анализа, основанных на Байесовском подходе, деревьях решений, нейронных сетях и др. Предлагается общий интегрированный подход к реализации комплекса методов детектирования вредоносного программного обеспечения.

К настоящему времени опубликовано большое количество работ по детектированию вредоносного ПО. Выделим три из них. Одной из первых работ, посвященных использованию для построения моделей детектирования методов классификации, является работа группы С. Столфо [1]. В ней рассматриваются проблемы, которые на данный момент не имеют столь сильной остроты за счет серьезного рывка в развитии средств защиты программ от статического анализа, однако она продемонстрировала всю мощь рассматриваемого в контексте данной работы подхода. Работа Д. Ванга и др. [2] посвящена вопросам построения модели детектирования, использующей данные, полученные в процессе статического анализа исполняемых файлов формата PE32 (основной формат потенциальных носителей вредоносного кода на момент проведенного исследования). В работе Б.Занга и др. [3] рассматриваются вопросы применения методов Data Mining для построения модели детектирования приложений на основе собираемой в процессе их выполнения поведенческой информации.

Очевидно, что условие наличия некоторого набора данных, обеспечивающих успешное выполнение детектирования конечного набора образцов вредоносного ПО, не является достаточным для успешного поиска его новых образцов. Это обуславливает необходимость поиска более эффективных и гибких методов детектирования. Ясно, что более эффективным, с точки зрения безопасности, является детектирование вредоносного ПО без его выполнения. В данном случае хост, на котором производится проверка, не подвергается деструктивному воздействию. Данное условие породило группу статических методов детектирования. Вместе с тем, существуют мощные средства сокрытия кода (упаковщики, протекторы), наличие которых существенно затрудняет использование статических методов. Ответом на данные трудности явилось появление методов детектирования "на лету", реализующих детектирование при выполнении кода. Решение проблемы безопасности для таких методов выполняется, как правило, за счет формирования изолированной вычислительной среды.

Очевидно, что детектирование вредоносного ПО может основываться на двух основных подходах: обнаружении заданных статических и поведенческих признаков (как правило, различных для разных классов вредоносного ПО) и обнаружении аномальных признаков (отличающихся от типичных заведомо безопасных приложений). Особенности данных подходов указывают на необходимость проведения исследований, посвященных изучению комплексного использования различных методов. Оно обеспечивает разумный компромисс в соблюдении следующих базовых критических требований к детектированию: (1) точность - минимальные значения ошибок первого и второго рода принятия решения о вредоносности приложений, не находившихся в обучающем наборе; (2) своевременность (оперативность) принятия решения; (3) эффективное потребление вычислительных ресурсов. Причем следует иметь четкие оценки применимости тех или иных правил выделения признаков, отбора из их числа наиболее значимых и, в конечном итоге, их использования на основе методов Data Mining. Кроме того, важно обеспечение требований устойчивости выполнения детектирования и скрытности сбора данных, необходимость которого обусловлена наличием разнообразных техник его обнаружения и противодействия. В рамках представляемого подхода фокус смещен на обнаружение заведомо опасных поведенческих признаков. Он основан на использовании методов классификации (отнесения объектов к тому или иному классу на базе формируемой математической модели). Использование методов классификации предполагает проведение предварительного обучения выбранного классификатора с последующим использованием определенного набора настроенных весов. Признаки выделяются в процессе обучения классификатора на выборке, содержащей приложения, отнесенные к целевым классам. Пространство признаков является многомерным и определяется количеством выделенных признаков. Решающая математическая модель представляет собой функцию, определенную на пространстве признаков, оптимально разделяющую вектора, отнесенные к тому или иному классу на этапе обучения.

Применяемые в работе математические модели классификации основаны на использовании следующих групп методов: статистических, используются Naive Bayes, и его специализации, уменьшающие влияние изначального предположения о взаимной независимости атрибутов; индуктивных, используются деревья решений; классификаторов, базирующихся на основе разделимости множеств (на данной фазе исследований был использован классификатор на основе многослойного персептрона).

Сбор исходных данных в подходе основан на мониторинге вызовов низкоуровневых функций операционной системы. Это позволяет получать хронологически корректную последовательность о фактах использования приложением критических системных ресурсов, без которых представляется трудным создать полноценный вариант работоспособного вредоносного ПО. Описание терминального инцидента (события) включает в себя: имя вызванной функции; значения переданных на вход функции операндов; значения возвращенных функцией результатов. Формирование пространства признаков с учетом последних двух наборов и факт мониторинга низкоуровневых функций ОС выгодно отличает предлагаемый подход от описанного в [3].

Процесс выделения из набора хронологически упорядоченных инцидентов неоптимального множества признаков учитывает: количество вызовов каждой функции; количество обращений к ресурсам, обладающим одинаковым рангом значимости; факты запросов специфических ресурсов (попытки обращения к разделам системного реестра, файловой системе и т.д.); наличие и количество специфических цепочек вызовов и др.

Скрытность мониторинга основывается на намеренной модификации структур ядра операционной системы, недоступных для приложений, функционирующих в пользовательском режиме. Тем самым, функционирующее приложение лишено возможности по явным признакам вынести вердикт о наличии факта слежения за ним. Выполнение требований устойчивости и оперативности в общем случае зависит от качества программной реализации модулей перехвата и анализа. Кроме того, выполнение требования оперативности опосредованно определяется особенностями используемых классификаторов (к примеру, возможностью их переобучения).

Программный комплекс сбора данных и оценки результатов построен на базе Windows XP (NT5.1). Сбор данных поведенческой информации основан на внедрении программных перехватчиков функций Native API. Для формирования исходных данных и данных для проверки сформированных классификаторами моделей используется набор вредоносных приложений и типовых безопасных приложений, входящих в состав операционной системы. Для проведения обучения, кросс-проверок, контрольных проверок и визуализации результатов используется специализированный программный комплекс Weka Classifier. Запуск вредоносных программных приложений, необходимый для получения «трасс» их выполнения, производится в изолированной вычислительной среде под максимально привилегированной учетной записью. Для проведения экспериментов была сформирована выборка вредоносных приложений, реализующих свой жизненный цикл с помощью функциональных рутин файлового перебора, сохранения своих резервных копий и автоматического запуска.

Минимальной неделимой информационной единицей при формировании пространства признаков является терминальный инцидент (событие). Описание терминального инцидента включает в себя: идентификатор вызванной функции, значения переданных на вход функции операндов, значения возвращенных функцией результатов. Процесс выделения из набора хронологически упорядоченных инцидентов множества признаков учитывает: количество вызовов каждой функции; количество обращений к ресурсам, обладающим одинаковым рангом значимости; факты запросов специфических ресурсов (загрузка системных библиотек, попытки обращения к разделам системного реестра, системным файлам и т.д.); наличие и количество специфических цепочек вызовов. Выделение специфических для вредоносного ПО вызовов обеспечивается на основе использования формулы расстояния Хэмминга.

Анализ промежуточных результатов прояснил как некоторые вопросы оценки эффективности использования методов Data Mining, так и направления дальнейшего расширения программного комплекса моделирования: доступные исходные данные (файлы вредоносного ПО) имеют достаточно низкий уровень качества, не позволяющий без дополнительных усилий сформировать релевантную обучающую/тестовую выборку; полученные трассы выполнения вредоносного ПО подтверждают применимость более простых технологий детектирования, ориентированных на выбранный для начальных опытов класс вредоносных программ (контроль выделенных областей системного реестра, механизмов межпроцессного взаимодействия, работы с файловой системой); на начальном пространстве признаков и текущем (нерелевантном) тестовом наборе используемые классификаторы демонстрируют возможность обнаружения до 80% неизвестного вредоносного ПО классов троян/файловый вирус/почтовый вирус (платформа Win32) при 15% показателе ложных срабатываний; используемая группа статистических классификаторов удовлетворяет требованиям оперативности в силу возможности инкрементальной дообучаемости. Вместе с тем, текущие результаты экспериментов показали необходимость более тщательного выделения наборов поведенческих признаков, специфичных для каждого класса вредоносного ПО, расширения программного комплекса моделирования и продолжения их усложнения за счет учета дополнительных данных.

Работа выполнена при финансовой поддержке РФФИ (проект №07-01-00547), программы фундаментальных исследований ОНИТ РАН и при частичной финансовой поддержке, осуществляемой в рамках проекта Евросоюза RE-TRUST (контракт № 021186-2).

**Литература**

1. Schultz M.G., Eskin E., Zadok E., Stolfo S.J. Data Mining Methods for Detection of New Malicious Executables // Informatics and Computer Science, 2005, Vol.172, Issue 1-2. P.241-261.
2. Wang J.-H., Deng P.S., Fan Y.-S., Jaw L.-J., Liu Y.-C. Virus Detection using Data Mining Techniques // Proceedings of IEEE 37th Annual 2003 International Carnahan Conference, 2003. P.71-76.
3. Zhang B.-Y., Yin J.-P., Hao J.-B., Zhang D.-X., Wang S.-L. Using Support Vector Machine to Detect Unknown Computer Viruses // International Journal of Computational Intelligence Research, 2006, Vol.2, No.1. P.100-104.