Раннее обнаружение эпидемий сетевых червей в высокоскоростных каналах передачи данных

Булгаков И. А., Гамаюнов Д. Ю.
лаборатория Вычислительных комплексов
факультет ВМК МГУ имени М. В. Ломоносова
{writer, gamajun}@lvk.cs.msu.su

В работе предложен метод раннего обнаружения эпидемий сетевых червей на основе анализа частоты встречаемости участков вредоносного исполнимого кода в сетевом трафике. Идея метода заключается в выявлении таких участков исполнимого кода архитектуры x86, частота встречаемости которых в наблюдаемом канале изменяется в соответствии с простой эпидемиологической моделью. Предложенный метод обладает линейной вычислительной сложностью, а экспериментальные исследования на программном прототипе демонстрируют пропускную способность порядка 1Гбит/с на типовом оборудовании.

За последние десять лет арсенал «киберпреступности» качественно изменился: в область, где ещё в 90-е годы преобладали ручные методы удалённого взлома уязвимого программного обеспечения, пришла автоматизация. Сегодня значительная часть нелегальной активности в Интернете так или иначе связана с так называемыми «ботнетами» - крупными сетями заражённых вредоносным программным обеспечением компьютеров из разных регионов мира, которыми управляет злоумышленник. Можно вспомнить такие известные ботнеты как StormNet или Torpig, использовавшиеся для самой разнообразной вредоносной деятельности: организации DDoS-атак, рассылки спама, хостинга вредоносного контента для XSS-атак и так далее. Ботнеты, как правило, распределены по многим автономным системам и не ограничены географией какой-то одной страны. По статистике Team Cymru рекордсменами по количеству управляющих ботнетами серверов являются США, Китай, Россия и Германия, при этом США занимают первое место с существенным отрывом от остальных [1].

Распространённая схема построения ботнета включает в себя внедрение «бота» на пользовательские компьютеры через эксплуатацию той или иной уязвимости в прикладном программном обеспечении. Обычно это полностью автоматическая процедура, в ходе которой так называемый сетевой червь распространяется на все доступные компьютеры, в которых есть соответствующая уязвимость. Ранее было показано, что распространение сетевых червей можно описать известными из области биологии моделями распространения болезней в ходе эпидемий [2].

Следует отметить, что раннее обнаружение эпидемии сетевых червей является весьма актуальной задачей, так как борьба с уже развёрнутыми ботнетами в процессе их эксплуатации значительно затруднена. Значительные ресурсы тратятся на фильтрацию DDoS-атак, спама, выявление управляющих узлов и их блокирование. Возможность выявления «тела» червя на этапе распространения и автоматической его фильтрации на магистральных каналах сильно упростило бы задачу в целом.

Существующие методы обнаружения эпидемий сетевых червей основаны на выявлении характерных особенностей протекания сетевых эпидемий [3-9]:

- рост количества идентичных пакетов в трафике;
- большое количество пакетов, отправленных на недоступные IP адреса – характерная особенность этапа «сканирования» и поиска жертв;
- скачкообразный рост числа входящих и исходящих соединений для некоторого номера порта;

- изменение поведения хоста при заражении сетевым червем, либо изменение поведения множества хостов.

Наибольшую точность показывают методы из первой группы, но для них характерны высокая вычислительная сложность и чувствительность к полиморфизму сетевых червей.

В данной работе предлагается искать в сетевом трафике неотъемлемую часть «тела» любого сетевого червя – его шеллкод. В состав любого шеллкода входит участок машинного кода, который отвечает за активацию шеллкода на атакованном компьютере. Одним из наиболее распространённых активаторов шеллкода являются NOP-зоны – последовательности инструкций, единственная задача которых заключается в том, чтобы привести исполнение в нужную точку шеллкода, независимо от того, с какого байта началось исполнение. NOP-зоны используются как при переполнении стека, так и при атаках переполнения «кучи» (heap overflow).

Ранее нами был предложен алгоритм Racewalk, позволяющий обнаруживать NOP-зоны с высокой скоростью и точностью. В данной работе алгоритм Racewalk используется для первичного выявления участка NOP-эквивалентного кода, после чего вычисляется частота встречаемости этого участка в наблюдаемом трафике [10].

В соответствии с простой эпидемической моделью распространения сетевых червей количество заражённых узлов i изменяется со временем по следующему закону:

$$i(t) = \frac{1}{1 + (\frac{1}{i_0} - 1)^{-\beta \cdot t}},$$

где i_0 – начальное число заражённых узлов, а β – скорость распространения червя.

Будем считать, что частота встречаемости «тела» червя в трафике линейно зависит от числа заражённых узлов. Тогда число вхождений NOP-зоны шеллкода данного червя будет расти со временем по эпидемическому закону с точности до константы-множителя. Для обнаружения такой NOP-зоны будем вычислять первую производную числа вхождений всех встреченных NOP-зон, т.е. частоту встречаемости. Для каждого обнаруженного участка будем анализировать динамику изменения частоты его встречаемости – если для некоторого участка частота монотонно возрастает со временем, можно сделать вывод о наличии эпидемии.

Таким образом, алгоритм раннего обнаружения эпидемии сетевых червей выглядит следующим образом:

1. Анализ очередной сетевой сессии алгоритмом Racewalk.
2. Если обнаружена NOP-зона, выполняем поиск найденного участка в хэш-таблице ранее встреченных NOP-зон.
    a. Если строка найдена в таблице, вычисляем новое значение частоты встречаемости для данной строки.
    b. Если полученное значение превышает установленный порог, генерируется сообщение о возможном распространении сетевого червя.
3. Если NOP-зона встречена в трафике впервые, для неё заводится ячейка в хэш-таблице и значение частоты выставляется в 0.
4. Вычисляется новое значение общей частоты встречаемости NOP-зон. Если частота превышает установленный порог, генерируется сообщение о возможном распространении сетевого червя.
5. Переход к шагу 1.

Данный алгоритм аналогичен алгоритму системы EarlyBird [4], но в нём анализируется частота встречаемости не всего поля данных каждого сетевого пакета, а лишь NOP-эквивалентных последовательностей инструкций x86, которые являются «сильным» признаком наличия вредоносного кода в передаваемых данных. При этом алгоритм устойчив к полиморфизму кода сетевых червей, так как позволяет обнаружить эпидемический рост общего числа NOP-зон в трафике, независимо от их схожести между собой. Реализация алгоритма Racewalk на экспериментах показала пропускную способность порядка

650Мбит/сек для одного процессорного ядра, что позволяет использовать предложенный метод на высокоскоростных каналах передачи данных.

В ближайшее время планируется экспериментально исследовать предложенный метод обнаружения сетевых червей на реальных каналах в Интернет.

Литература

1. Team Cymru. Developing Botnets … an analysis of recent activity. [PDF])http://www.team-cymru.org/ReadingRoom/Whitepapers/2010/developing-botnets.pdf)
2. Kristopher Joseph Hall. Thwarting Network Stealth Worms in Computer Networks through Biological Epidemiology. Ph.D. Thesis, Virginia Polytechnic Institute and State University, 2006.
3. Cliff C. Zou, Lixin Gao, Weibo Gong, Don Towsley. Monitoring and Early Warning for Internet Worms. In Proceedings of the 10th ACM International Conference on Computer and Communications Security (CCS), pages 190-199, October 2003.
4. Sumeet Singh, Cristian Estan, George Varghese, Stefan Savage. The EarlyBird System for Realtime Detection of Unknown Worms. Technical Report CS2003-0761, CSE Department, UCSD, Aug. 2003.
5. Sumeet Singh, Cristian Estan, George Varghese, Stefan Savage. Automated Worm Fingerprinting. . In Proceedings of the OSDI'04, 2004.
6. Xuan Chen, John Heidemann. Detecting Early Worm Propagation through Packet Matching. Technical Report ISI-TR-2004-585, USC/ISSI 2004.
7. D. Ellis, J. Aiken, K. Attwood, and S. Tenaglia. A Behavioral Approach to Worm Detection. In Proceedings of WORM'04, 2004.
8. Hyang-Ah Kim, Brad Karp. Autograph: Toward Automated, Distributed Worm Signature Detection. In Proceedings of the USENIX Security Symposium, 2004.
9. Ke Wang, Gabriela Cretu, Salvatore J. Stolfo. Anomalous Payload-based Worm Detection and Signature Generation. In Proceedings of the International Symposium on Recent Advances in Intrusion Detection (RAID), 2005.
10. Dennis Gamayunov, Nguyen Thoi Minh Quan, Fedor Sakharov, Edward Toroshchin. Racewalk: fast instruction frequency analysis and classification for shellcode detection in network flow. In Proceedings of 5th European Conference on Computer Network Defense (EC2ND 2009), 2009.