Раннее обнаружение эпидемий сетевых червей в высокоскоростных каналах передачи данных

Денис Гамаюнов

лаборатория вычислительных комплексов ВМК МГУ имени М. В. Ломоносова

Ботнеты — технологическая основа современной киберпреступности

- Распространение через уязвимости в Adobe Flash, MS IE, etc
- Скрытность управления - Fast flux / double flux, p2p / www
- Разнообразие задач
  - DDoS, распространение ВПО
  - Money-mule, Scam, Spam
  - Многое другое

Жизненный цикл

- Первичное заражение и распространение с помощью сетевых червей
- Подключение к ботнету
- Подгрузка функциональных модулей
- Использование (рассылка спама, DDoS, ...)
- Уничтожение
  - На каком этапе можно эффективно блокировать ботнет?

Обнаружение шеллкода

- Типичная структура шеллкода

Activator	Shellcode payload	Return address zone

Activator	Decryption engine	Encrypted shellcode payload	Return address zone

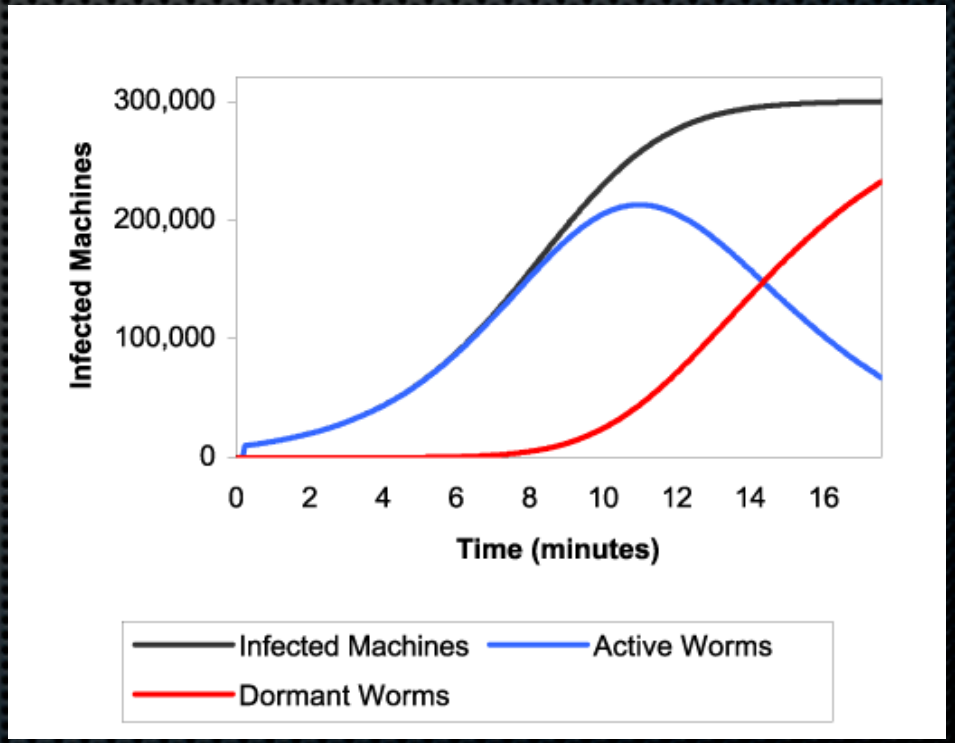
- Activator — NOP-sled для атак классов stack overflow, heap spraying, JIT spraying

Простая эпидемическая модель

- количество заражённых узлов *i* изменяется со временем по следующему закону:

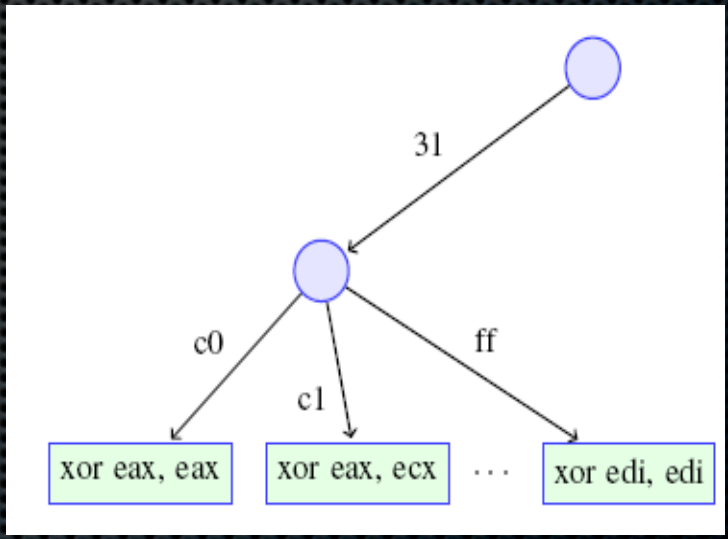
$$i(t) = \frac{1}{1 + (\frac{1}{i_0} - 1)^{-\beta \cdot t}}$$

  - где i_0 – начальное число заражённых узлов, а β – скорость распространения червя
  - **Идея: наблюдать за частотой появления шеллкода в канале**



Алгоритм Racewalk*

- Основа — алгоритм Stride
- Вход — байтовая строка длины N
  - Дизассемблируем префиксным деревом
  - Проверяем основное свойство NOP-sled: исполнимость с каждого байта (и с каждого 4-го)
  - Если да, то классифицируем с помощью SVM



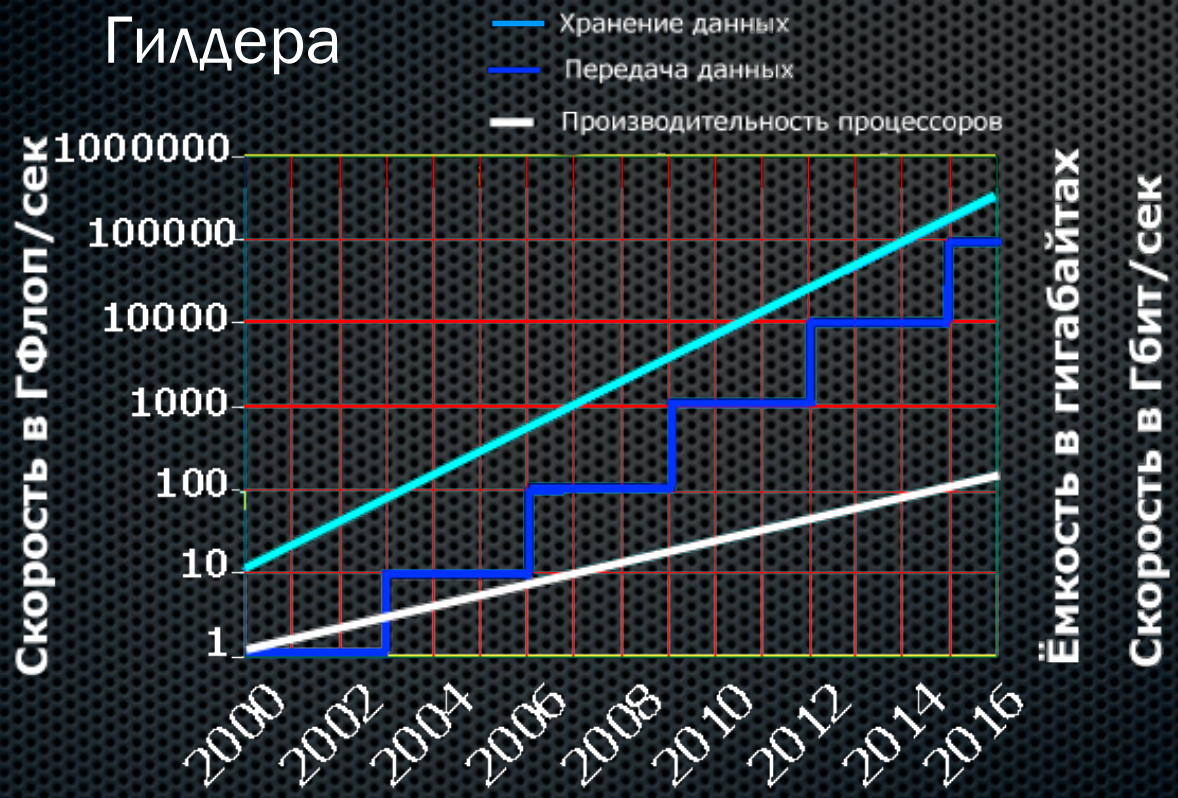
* Алгоритм описан в статье: http://redsecure.ru/papers/Racewalk-EC2ND-final.pdf

Алгоритм обнаружения эпидемии

- While new TCP session do
  - Racewalk_findNOP()
  - if NOP && lookupNOPTable (NOP) then
    - If (NOPTable[NOP].counter++ > Threshold) then Alert()
  - else if NOP then insertNOPTable(NOP)
  - If (globalNOPCounter++ > Threshold) then Alert()

Высокоскоростной канал — сложность задачи?

- Закон Мура vs закон Гилдера



1GigE:

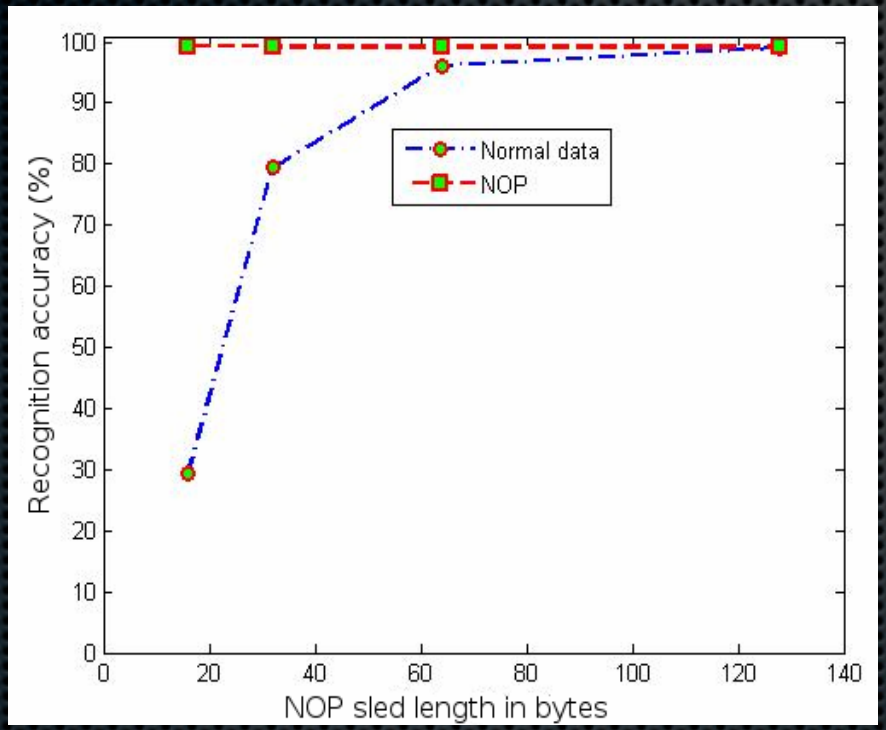
- 100-200 тыс. HTTP-запросов в секунду

Обучение Racewalk

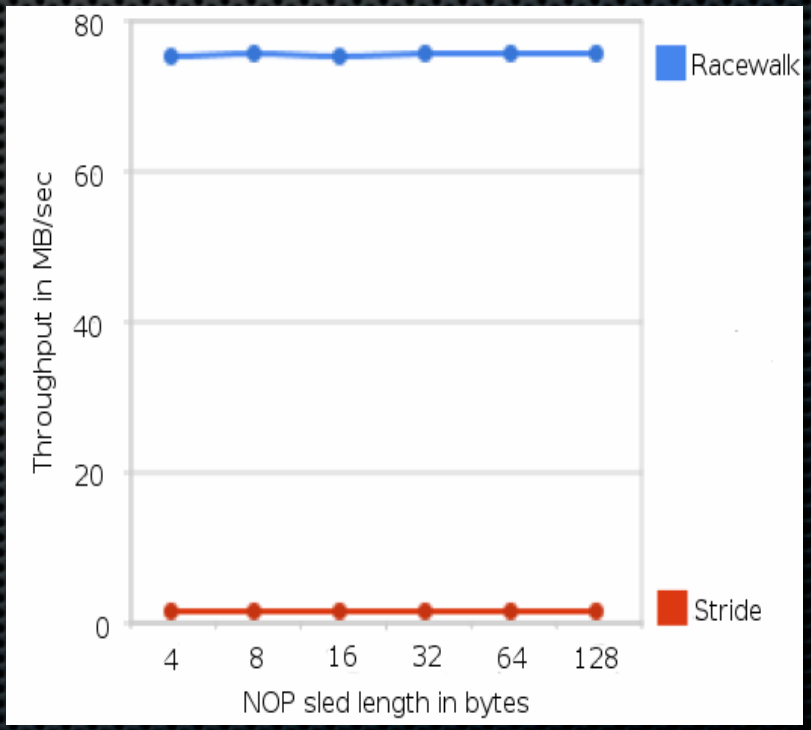
- Обучение на примерах NOP-зон, полученных реальными генераторами шеллкода:
  - CLET
  - ADMmutate
  - Ecl-Poly
  - Metasploit
- Обучающие и тестовые выборки:
  - Поиск минимальной длины эффективно обнаруживамого следа — 16, 32, 64, 128 байтов

Эксперименты

- Точность и FP rate



- Производительность



* Замеры получены на одном процессорном ядре Intel Xeon 3.16Ghz

Ближайшие перспективы

- Исследование на HoneyNet
- Создание статико-динамического метода обнаружения упакованного и шифрованного шеллкода

Спасибо за внимание

- Контактная информация:
  - Денис Гамаюнов: gamajun@lvk.cs.msu.su
  - Тел. +7 (495) 939 46 71
  - Москва, 119899 Ленинские горы вл. 1/52, факультет ВМК МГУ имени М. В. Ломоносова, к. 764