

Логико-динамические аспекты моделирования процессов контентной фильтрации прикладных протоколов

В.С.Заборовский, А.В.Силиненко

Государственный политехнический университет, ЦНИИ РТК, С.-Петербург

1. Введение. При решении многих задач защиты информации компьютерную сеть можно рассматривать как сложную логико-динамическую систему, для характеристики свойств которой используются различные средства моделирования. Так, совокупность виртуальных соединений может рассматриваться как система массового обслуживания пакетов, управляемых протоколом IP. С другой стороны, описание трафика на уровне транспортных соединений опирается на модели, характеризующие динамические свойства протокола TCP, функционирующего в среде с ярко выраженными стохастическими или даже хаотическими свойствами. Широкий класс задач защиты информации может моделироваться с использованием методов предикатного или сигнатурного задания правил фильтрации на допустимых множествах сетевых индикаторов и/или латентов. Уточним, что в этих задачах индикаторами являются параметры трафика, непосредственно доступные для прямых измерений, например заголовки пакетов, а латентами – скрытые переменные или их совокупности, для определения числовых характеристик которых используются различные алгоритмы и/или характеристические функции. Целью работы является исследование логико-динамических аспектов моделирования процессов контентной фильтрацией трафика, порождаемого различными прикладными протоколами. Исследование основывается на модели TCP соединения, используемого для организации виртуальных соединений. При этом фильтрация рассматривается как двухэтапный процесс идентификации контентных признаков, в котором на первом этапе выделяются последовательности пакетов, принадлежащие отдельным TCP сессиям, а на втором определяется набор данных, подлежащих контролю, для чего используется механизм сигнатурного анализа в области данных, ограниченных рамками выделенной TCP сессии. Работа состоит из трех разделов. В первом разделе рассматриваются особенности моделирования сетевых процессов, связанные с динамическими свойствами TCP протокола и вероятностным характером среды передачи пакетов. Во втором разделе анализируются возможности реализации процессов контентной фильтрации при решении задачи информационной безопасности. В третьем разделе проводится формализация задачи контентной фильтрации сетевых приложений, в рамках введенной алгебры правил фильтрации.

2. Модель TCP сессии. TCP соединение как объект моделирования характеризуется параметрами и индикаторами - пропускная способность, размер входного и выходного

буфера, окно перегрузки, окно приемника, размер порога медленного старта, RTT, номера последовательности и подтверждения, номера портов, размеры сегментов и др. В качестве латентных признаков или латентов рассматривается дисперсия пропускной способности, корреляционные и спектральные характеристики трафика, свойства персистентности (параметр Херста) и др. Динамика процессов передачи пакетов, связанная адаптацией TCP сессий к параметрам виртуального соединения показана на Рис. 1.

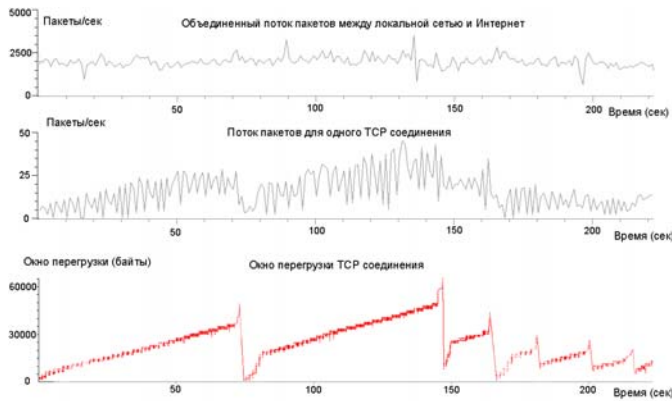


Рис. 1. Динамика сетевых процессов.

Из рисунка видно, что TCP обеспечивает гарантированную доставку данных за счет механизма повторной передачи потерянных пакетов и устранения дублирования пакетов при получении нескольких копий. В каждый дискретный момент времени поток TCP сегментов, которые обозначим переменной y_k , изменяется согласно принятым спецификациям (1) протокола и значению индикаторной функции $\psi(\xi)$. Значение этой функции выбирается из множества $\{0,1\}$ в зависимости от того, произошло ли событие, связанное с потерей или повторной передачей пакета.

$$y_{k+1} = y_k + 1; \quad \text{if } \psi(\xi) = 0$$

$$y_{k+1} = y_k / 2; \quad \text{if } \psi(\xi) = 1, \quad (1)$$

В этих формулах (1) $\psi(\xi)$ – индикаторная функция, ξ – стохастическая переменная момента перегрузки, имеющая равномерное распределение на интервале текущего динамического диапазона изменения окна перегрузки (Рис. 2).

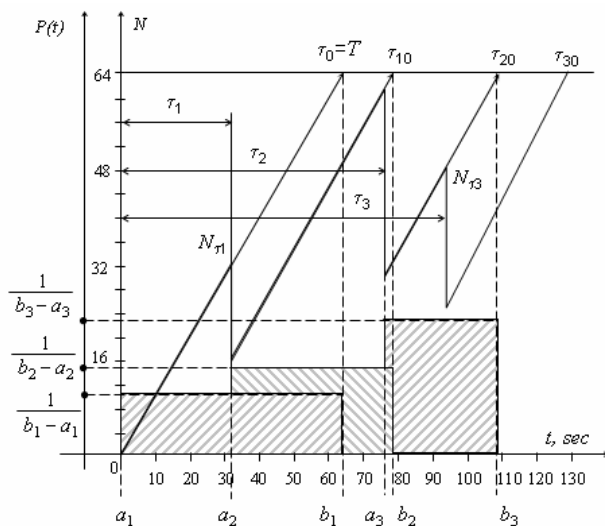


Рис. 2 Окно пререгрузки и крос-интервальное равномерное распределение переменной ξ
 Модель (1) позволяет определить статистики, характеризующие свойства TCP сессии на различных интервалах наблюдения, например:

$$M_1(\tau_{10}) = \int_0^T \frac{\partial \tau_1}{T} \int_{-\infty}^{\infty} \tau \delta(\tau - \tau_{10}) \partial \tau;$$

$$D(\tau_{10}) = M_2(\tau_{10}) - M_1^2(\tau_{10});$$

$$M_1(\tau_{20}) = \int_0^T \frac{\partial \tau_1}{T} \int_{\tau_1}^{\tau_{10}} \frac{\partial \tau_2}{\tau_{10} - \tau_1} \int_{-\infty}^{\infty} \tau \delta(\tau - \tau_{20}) \partial \tau; \quad D(\tau_{20}) = M_2(\tau_{20}) - M_1^2(\tau_{20}); \quad (2)$$

$$M_1(\tau_{30}) = \int_0^T \frac{\partial \tau_1}{T} \int_{\tau_1}^{\tau_{10}} \frac{\partial \tau_2}{\tau_{10} - \tau_1} \int_{\tau_2}^{\tau_{20}} \frac{\partial \tau_3}{\tau_{20} - \tau_2} \int_{-\infty}^{\infty} \tau \delta(\tau - \tau_{30}) \partial \tau; \quad D(\tau_{30}) = M_2(\tau_{30}) - M_1^2(\tau_{30}).$$

Эти статистики позволяют получить выражение для дисперсии потока пакетов, что важно для решения задачи прогнозирования пропускной способности: $D(t) = K(t - 64)^{1+\alpha}$..,

где $K = 1,17$; $\alpha = 0,82$. Характер этой зависимости говорит о положительной персистентности или фрактальности процессов, что подтверждается данными реальных измерений трафика для различных приложений и режимов работы сети (Рис.1).

3. Анализ возможности реализации контентной фильтрации. Рассматривая персистентность как характеристику устойчивости состояния TCP соединения, сформулируем задачу контентной фильтрацией пакетов, порождаемых различными прикладными протоколами как задачу определения множества латентов, к которым относятся различные параметры TCP сессии, среди которых имеется и такие параметры, точная структура которых неизвестна. Для моделирования процесса контентной фильтрации с использованием межсетевого экрана (МЭ) будем учитывать контекст TCP-соединения, отражающий состояния, в котором находится конкретное виртуальное соединение. При этом будем опираться на автоматную модель TCP-соединения, позволяющую описать последовательности смены состояний TCP-соединения в результате прохождения пакетов через МЭ. В данной модели предлагается расширенное описание состояния ESTABLISHED (соединение установлено), что позволяет учесть особенности функционирования прикладных протоколов, использующих TCP как протокол транспортного уровня (Рис. 3).

Для формирования алгоритмов контентной фильтрации на базе автоматной модели TCP соединения требуется согласованное решение следующих задач:

- формализация алгоритмов контентной фильтрации;
- параметризация требований политики безопасности;
- синтез модели безопасности и требований к МЭ с использованием средств описания правил фильтрации прикладных протоколов.

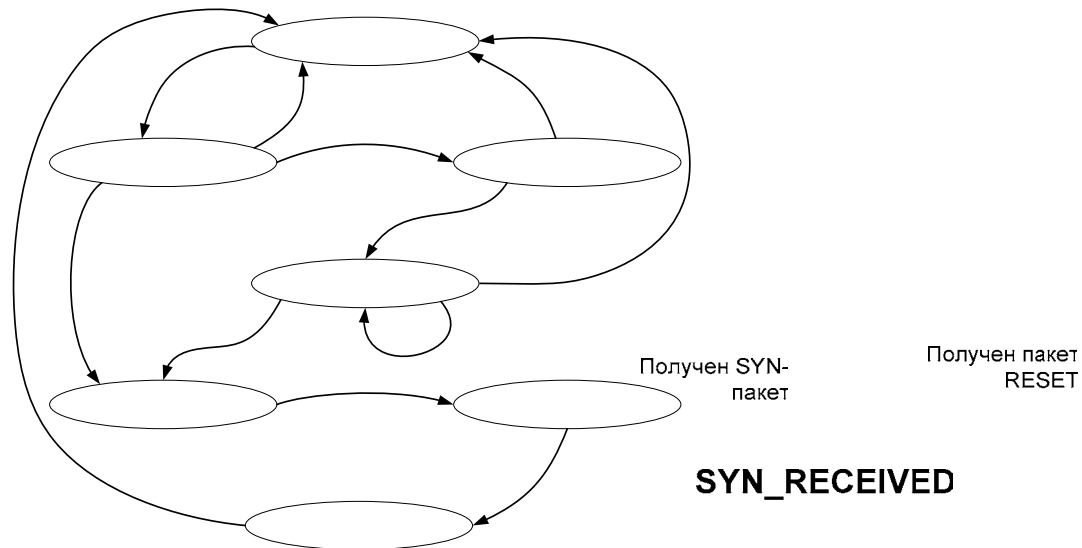


Рис. 3 Модель соединений TCP соединения.

4. Формализация методов описания правил фильтрации. Повысить эффективность синтеза средств защиты информации на базе МЭ предлагается на основе формализации используемых понятий, введения новых абстракций и формирования алгебры правил фильтрации. Для этого введем следующее описание: $\mathcal{R} = \langle R, \Sigma \rangle$, где $R = \{R_1, R_2, \dots, R_n\}$ – множество правил фильтрации, несущее множеств алгебры \mathcal{R} , $R_i = \{x_{i1}, x_{i2}, \dots, x_{ik}, a_{i1}, a_{i2}, \dots, a_{im}\}$ – правило фильтрации, состоящее из параметров $\{x_i\}$ и атрибутов $\{a_i\}$ ($i=1..n$), $\Sigma = \{\varphi_1, \varphi_2\}$ – сигнатура алгебры \mathcal{R} , φ_1 – символ операции сложения, φ_2 – символ операции умножения.

Для задач фильтрации функция сложения определяется следующим образом:

$$R_3 = R_1 + R_2 = \{x_{11}, x_{12}, \dots, x_{1k}, a_{11}, a_{12}, \dots, a_{1m}\} + \{x_{21}, x_{22}, \dots, x_{2k}, a_{21}, a_{22}, \dots, a_{2m}\}$$

$$R_3 = \left\{ \begin{aligned} &\{x_{11} \vee x_{21}, x_{12} \vee x_{22}, \dots, x_{1k} \vee x_{2k}, a_{11}, a_{21} \vee a_{22}, \dots, a_{1m} \vee a_{2m}\}, \text{ где } a_{11} = a_{21} \\ &\{x_{11} \wedge x_{21}, x_{12} \wedge x_{22}, \dots, x_{1k} \wedge x_{2k}, a_{11}, a_{21} \vee a_{22}, \dots, a_{1m} \vee a_{2m}\}, \text{ где } a_{11} \neq a_{21} \end{aligned} \right. \quad (3)$$

где a_{il} – атрибут действия правила фильтрации, $a_{il} = \{0,1\}$, где $a_{il} = 0$ – правило производит удаление пакета, $a_{il} = 1$ – правило разрешает пропуск пакета. Другими словами, сумма правил R_1 и R_2 есть объединение их одноименных параметров при одинаковом действии правил R_1 и R_2 (например, пропуск) и разность их одноименных параметров при различных действиях правила R_1 и R_2 (например, правило R_1 на пропуск, правило R_2 на удаление).

Функция умножения для правил фильтрации задается следующим образом:

$$R_3 = R_1 * R_2 = \{x_{11}, x_{12}, \dots, x_{1k}, a_{11}, a_{12}, \dots, a_{1m}\} * \{x_{21}, x_{22}, \dots, x_{2k}, a_{21}, a_{22}, \dots, a_{2m}\}$$

$$R_3 = \{x_{11} \wedge x_{21}, x_{12} \wedge x_{22}, \dots, x_{1k} \wedge x_{2k}, a_{11} \wedge a_{21}, a_{21} \vee a_{22}, \dots, a_{1m} \vee a_{2m}\} \quad (4)$$

Другими словами, произведение правил R_1 и R_2 есть пересечение их одноименных параметров, при этом действие правила представляет собой результат конъюнкции значений атрибутов a_{il} правил R_1 и R_2 .

В результате процесс обработки пакетов в МЭ представляется функцией $\psi(\varphi, R)$, которая определяется следующим образом:

$$\psi(\varphi, R) = \begin{cases} \{a_1, a_2, \dots, a_m\}_d \text{ для } \varphi(R, p) = 0 \\ \{a_1, a_2, \dots, a_m\}_i \text{ для } \varphi(R, p) = i (i > 0) \end{cases}, \quad (5)$$

где R – множество правил фильтрации, $\{a_1, a_2, \dots, a_m\}_d$ – вектор атрибутов по умолчанию для обработки сетевых пакетов, $\{a_1, a_2, \dots, a_m\}_i$ – вектор атрибутов правила фильтрации, p – обрабатываемый пакет, $p = \{y_1, y_2, \dots, y_k\}$, y_i – параметры принятого пакета, $i = 1..k$, $\varphi(R, p)$ – характеристическая функция правила фильтрации, которая вычисляется следующим образом:

$$\varphi(R, p) = \begin{cases} i \text{ для } \bigwedge_{j=1}^k x_{ij} \wedge y_j = 1, i = 1..n \\ 0 \text{ для } \bigvee_{i=1}^n \bigwedge_{j=1}^k x_{ij} \wedge y_j = 0 \end{cases}$$

В последней формуле в качестве параметров $\{x_j\}$, $j = 1..k$, могут выступать как индикаторные параметры пакета (такие, как IP-адреса и порты), так и латентные параметры, задаваемые регулярными выражениями и позволяющими производить контентную фильтрацию прикладных протоколов.

5. Заключение. В работе рассматривались актуальные вопросы построения моделей сетевых процессов, учитывающие логико-динамические аспекты, связанные с реализацией методов контентной фильтрации трафика. Представленные модели позволяют получить верифицируемые характеристики, определяемые динамическими свойствами TCP протокола и особенностями его использования для контроля состояний виртуальных транспортных соединений. Проведенный анализ созданных моделей показывает возможность реализации двухуровневой системы контентной фильтрации, в которой разделение трафика осуществляется с использованием средств контроля состояний TCP сессий и сигнатурного анализа выделенных последовательностей пакетов. Для формализации процесса синтеза правил фильтрации предлагается использовать элементы новой алгебры, в которой операции могут интерпретироваться в контексте теоретико-множественного описания индикаторных и латентных переменных.