при финансовой поддержке РФФИ в рамках научного проекта № 14-41-01564 «Разработка моделей и методов вероятностного анализа сетевого трафика в условиях гетерогенных потоков данных»

Ключевые слова: потоки данных, сетевые анализаторы, анализ трафика

Milyukov V.V., Zuev S.A., Pokrova S.V. Systemic-functional analysis of the data flow across the Internet service providers in the Crimea.

The system approach of network traffic analysis are discussed in this article. Obtained statistical data to optimize and improve the quality of service in multiservice networks. Tested routes and obtained comparative characteristics data of time delays in the Crimea providers. The reported study was partially supported by RFBR, research project No. 14-41-01564.

Keywords: Data streams, network analyzers, traffic analysis

УДК 621.38

Сосновский Ю.В., доц., канд. техн. наук
Таврический национальный университет им. В.И. Вернадского, Республика Крым г.Симферополь, пр. Вернадского 4.
e-mail: mailsender256@mail.ru

ПРОБЛЕМАТИКА ИДЕНТИФИКАЦИИ ПОТОКОВ ГЕТЕРОГЕННОГО ТРАФИКА В УСЛОВИЯХ НЕСТАЦИОНАРНОСТИ[18]

Статья содержит обзор фундаментальных исследований в области анализа сетевого гетерогенного трафика в условиях нестационарности.

Ключевые слова: *анализ сетевого трафика, гетерогенность, нестационарность*

Проблема идентификации потоков гетерогенного трафика сама по себе является достаточно сложной научной задачей. Гетерогенность сетевого трафика проявляется в наличии в составе передаваемых блоков данных различных компонент трафика, например HTTP-трафика, потокового видео и аудио, а также шифрованного трафика, который имеет измененный профиль за счет перепаковки исходных пакетов того или иного протокола. Гетерогенность трафика отмечена достаточно давно, например в [1].

В отчете 2014 года, опубликованным Sandvine (http://www.sandvine.com/trends/global-internet-phenomena), указывается по сравнению с прошлым годом, в 2014 году отмечено увеличение доли зашифрованного трафика. Лидером роста является Европа (процент зашифрованного интернет-трафика в периоды пикового времени увеличился за год с 1,47% до 6,10%). Средний показатель роста абсолютного интернет-трафика увеличивается на 20%-40% за год, в тоже время пропускная способность каналов, потребляемая зашифрованным трафиком, удваивается в этот период. Аналогичная ситуация наблюдается также в сетях мобильной связи, где зашифрованный трафик также растет. Важно отметить, что значительный рост зашифрованного трафика приходится на SSL-шифрование, что с позиции системного администратора не является неразрешимой проблемой, хотя и привносит существенные трудности в сигнатурный анализ передаваемых данных и значительно снижает эффективность работы сетевых сканеров, работающих по принципам сигнатурного анализа потоков данных.

Ряд источников, как независимые эксперты, так и группы аналитиков [2] отмечают, что около 40% клиентов пиринговых одноранговых сетей (P2P) по тем или иным соображениям используют шифрование трафика. Самые известные представители P2P сетей – торренты, и шифрование не скрывает адрес источника и получателя, но скрывает сам факт передачи именно P2P трафика.

В тоже время отмечен рост популярности VPN-сетей, кроме того ведущие ученые предлагают расширить это понятие и начать переход на новый уровень виртуализации компьютерных сетей [3]. Можно привести множество ссылок на самые современные публикациипо теме роста объемов трафика, генерируемых VPN-сетями, повышением требований к пропускной способности каналов и защищенности такого трафика, однако объем статьи не позволяет этого сделать.

Несколько обособленно находится т.н. анонимная сеть виртуальных туннелей Tor, осуществляющая передачу данных в зашифрованном виде и позволяющая пользователю сохранять анонимность в Интернет. Ввиду высокой степени анонимности существует значительное число свидетельств использования данной технологии для распространения вредоносных программ, вирусных атак и т.п.[4,5]. При этом отмечается, что сама по себе идентификация подобного шифрованного трафика уже представляет собой серьёзную проблему, особенно в условиях общего гетерогенного потока, генерируемого различными системами, приложениями и сервисами. Еще боле сложной задачей в этих

[18] Работа выполнена при финансовой поддержке РФФИ (грант № 14-41-01564).

условиях является идентификация вредоносных действий, таких как повышенной вирусной активности, сетевых атак и т.п. [6]

Таким образом, задача идентификации типа сетевого трафика в рамках гетерогенного потока данных в условиях нестационарности как степени гетерогенности, так и интенсивности является фундаментальной научной проблемой, для решения которой требуется разделение ее на подзадачи.

Целью работы является анализ фундаментальных исследований в области идентификации гетерогенного трафика и задач кластеризации подобного трафика. В той или иной степени задачу идентификации сетевого трафика с различными входными условиями решают различные коллективы авторов. В частности, в [7] выполнена значительная работа по упорядочиванию и классификации функциональных требований к системам обнаружения вторжений с учетом требований стандарта Российской Федерации ГОСТ Р ИСО/МЭК 15408 и новых требований ФСТЭК , доступных на официальном сайте организации (www.fstec.ru).

Авторский коллектив в составе Кондратьева А.А., Талалаева А.А., Тищенко И.П., Фраленко В.П., Хачумова В.М. в [8] выполнили обобщение результатов автоматного подхода к описанию и моделированию сетевых атак на основе теории распознающих автоматов, теории взаимодействующих автоматов В.М. Глушкова, теории сетей Петри и автоматного подхода к описанию сетевых атак на облачные вычисления. Предлагается на этой основе новый класс алгоритмов обнаружения сетевых атак с генетически конфигурируемыми конечными автоматами и созданием таблично-алгоритмического подхода к описанию и обнаружению сетевых атак и метода построения комитета классификаторов, учитывающих комплекс интеллектуальных методов, включая алгебраический подход к распознаванию Ю.И. Журавлева.

Группа авторов в составе М.П. Сычева, А.В. Астрахова, К.Б. Здирука, И.И. Подвойского из МГТУ им. Н.Э. Баумана и ОАО «НИИАА им. академика В.С. Семинихина» [9] предлагает подход, использующий экспертные системы с элементами нечеткой логики, а для дополнительных модулей предлагается применение методов математической статистики. Используются такие характеристики, как среднее арифметическое, математическое ожидание и среднеквадратическое отклонение.

Большой объем работ проводится иностранными коллективами авторов. В [10] отмечается необходимость предоставления различному типу трафика разных сервисов, в тоже время некоторая часть трафика может являться нежелательной. Ввиду отсутствия в рамках как аппаратных, так и программных систем (IDS, брандмауэр, router и т.п.) механизмов автоматической классификации трафика, предлагается упрощенная модель, предназначенная для классификации и идентификации трафика.

Известны попытки создания архитектуры системы идентификации, использующие наборы различных техник, а также оперирующей как реальным трафиком, так и предварительно сохраненными дампами [11].

В тоже время значительная часть иностранных авторских коллективов [12] фокусируется на отдельных частях проблемы идентификации гетерогенного трафика в условиях нестационарности или использует предположение, что трафик генерируется известными приложениями или, по крайней мере, определенным и известным набором сетевых приложений [13].

Из представленного обзора можно сделать несколько выводов:

- проблема идентификации, разбиения на классы гетерогенных потоков данных является актуальной и над решением частных задач работает значительное число авторских коллективов;
- перечень методов и предлагаемых программных реализаций значителен, в тоже время большинство решений решают ограниченный круг частных задач.
- большое число иностранных коллективов, работающих в рамках данной тематики, сосредоточено на создании коммерческих продуктов;
- несмотря на большой объем выполняемых работ можно констатировать практическое отсутствие значимых фундаментальных результатов по проблеме идентификации, кластеризации гетерогенного сетевого трафика.

Для выполнения дальнейших работ важно выполнить математическую формализацию сетевого потока в крупных сетях или магистральных канала где, как известно, трафик представляет собой совокупность потоков трафика различного типа, от различных источников и протоколов.

Сетевой поток данных в сетях класса А, В, тем более WAN является структурно-сложным потоком, допускающим декомпозицию на ряд взаимосвязанных между собой подпотоков, изменчивых во времени и обладающих собственными характеристиками. Т.о. сетевой потокс временным параметром $t(t \in T, T \subset R)$, допускающий различные виды описаний зависимости от поставленной задачи, может быть описан n-мерным случайным процессом $\xi(\omega,t)$, выполняющим измеримое отображение $\xi:\Omega \to R^n$, зависящее от параметра t. Точное отображение определяется рядом параметров, таких как доступные ресурсы, функциональная задача, решаемая отображением и пр.

Ввиду указанных особенностей сетевого трафика, соответствующий случайных процесс является вектором с размерностью n: $\vec{\xi}(\omega,t)=[\xi_1(\omega,t),\xi_2(\omega,t),...,\xi_n(\omega,t)]^T$, причем $n\subset Z, n\in[1,m], m\subset Z$, $m\geq 1$. Так как сетевой поток сложная и неоднородная структура, понятие «гетерогенность» для процессов, его образующих, требуют уточнения. Один из вариантов представления гетерогенности – зависимость размерности случайного процесса $\vec{\xi}$ от времени.

Базовые классы задач идентификации гетерогенных потоков трафика, которые могут быть использованы комплексом анализа [14], в соответствии с теорией случайных процессов, показаны на рисунке 1.

Для класса стационарных потоков разработан и успешно применяется аппарат аналитического моделирования. В тоже время для нестационарных потоков аналитический аппарат либо чрезмерно сложен, обладает рядом значительных допущений, либо его вообще нет.

Рисунок 1 – Базовые классы задач идентификации гетерогенных потоков трафика

Таким образом, направление дальнейших исследований автор видит в создании информационной технологии и системы поддержки принятия решений по идентификации сетевого гетерогенного трафика, для чего возможно применение методов как прямого, так и имитационного моделирования. Определенная работа в этом направлении была выполнена в ходе исследовательских работ, в частности, анализа потоков данных в сетях провайдеров Крыма (Зуев С.А., Покрова С.В.), реализации системы сбора характеристических данных сетевого трафика (Сосновский Ю.В., Зганяйко Д.О.). Кроме того, была создана функциональная схема комплекса анализа сетевого трафика, которая может быть применима, в том числе, для тестирования IDS/IPS систем [14].

Библиографический список использованной литературы

1. Голышко, А. Сетевая борьба за качество обслуживания / А. Голышко // Рынок ценных бумаг. - 2001. - № 20. - С. 46-47

2. Федеральное агентство по печати и массовым коммуникациям. Интернетв России: Состояние, тенденции и перспективы развития. Отраслевой доклад. Москва. – 2013.

3. Чл.-корр. РАН, д.ф-м.н, профессор Р.Л.Смелянский Центр прикладных исследований компьютерных сетей, ВМК МГУ им. Ломоносова Программно-конфигурируемые сети и Виртуализация сетевых сервисов – новый вызов для разработчиков ПО. Конференция SECR-2013. – Москва.

4. Aaron Johnson, Chris Wacek, Rob Jansen, Micah Sherr, PaulSyverson Users Get Routed: Traffic Correlation on Tor by Realistic Adversaries. — USA: U.S. Naval Research Laboratory (англ.), Georgetown University. — 18 с.

5. Nicholas Hopper Short Paper: Challenges in protecting Tor hidden services from botnet abuse. — USA: University of Minnesota, Minneapolis. — 10 с. Электронныйресурс. Режим доступа: [http://fc14.ifca.ai/papers/fc14_submission_152.pdf]

6. Marco Preuss. ChewBacca - a New Episode of Tor-based Malware (англ.) Электронныйресурс. Режимдоступа [https://securelist.com/blog/incidents/ 58192/chewbacca-a-new-episode-of-tor-based-malware/]

7. Половко И.Ю., Пескова О.Ю. Анализ функциональных требований к системам обнаружения вторжений / Известия южного федерального университета // №2(151), 2014г. С.86 – 91.

8. Кондратьев А.А. Методологическое обеспечение интеллектуальных систем защиты от сетевых атак. ФГБУН Институт программных систем им. А.К. Айламазяна Российской академии наук, Исследовательский центр мультипроцессорных систем // А.А. Кондратьев, А.А. Талалаев, И.П. Тищенко, В.П. Фраленко, В.М. Хачумов. Современные проблемы науки и образования. - №2. – 2014.

9. Сычев М.П., Астрахов А.В., Здирук К.Б., Подвойский И.И. Обоснование архитектуры перспективной системы обнаружения и предотвращения вторжений / Вестник МГТУ им. Н.Э. Баумана: электронное издание. 2013 г. Режим доступа: [http://engjournal.ru/articles/533/533.pdf]

10. Ekmanis M. Unwanted Traffic Identification Problems // Scientific Journal of RTU. 7. series., Telekomunikācijasunelektronika. - 7. vol. (2007), pp 19-22

11. W. de Donato, A. Pescapè, and A. Dainotti, "Traffic Identification Engine: An Open Platform for Traffic Classification", IEEE Network, vol. 28, no. 2, pp. 56--64, Mar 2014

12. Zhang, Jun, Chen, Chao, Xiang, Yang and Zhou, Wanlei 2013, Robust network traffic identification with unknown applications, in Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security; ASIA CCS 2013, ACM - The Association for Computing Machinery, New York, N.Y., pp. 405-414

13. Proceeding ASIA CCS '13 Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications securityP. 405-414 ACM New York, NY, USA, 2013

14. Сосновский Ю.В. Функциональная схема комплекса анализа трафика и тестирования IDS/IPS систем / Ю.В. Сосновский // Оптимизация производственных процессов. Сборник научных трудов Севастопольского национального технического университета. Вып.15. – Севастополь, 2014. – С.190-194

Sosnovskij Y.V. The identification problems of heterogeneous traffic flows under nonstationarity

This article contains an overview of basic research in the analysis of network traffic in a heterogeneous environment nonstationarity.

Key words: network traffic analysis, heterogeneity, nonstationarity.

УДК 004.32.2+004.04

Ю.В. Доронина, доц., д-р техн. наук
Т.С. Карпова
Севастопольский национальный технический университет
ул. Университетская 33, г. Севастополь, Россия, 299053
E-mail: 1-2-1-21@mail.ru

АНАЛИЗ ПРОЦЕССА УСТАРЕВАНИЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Рассматривается процесс и строятся модели изменения во времени эффективности информационной системы вследствие ее устаревания. Определяется целесообразность применения различных законов распределения в этих моделях.

***Ключевые слова**: информационная система, модель устаревания систем, реинжиниринг.*

Введение. Информационные системы (ИС) незаменимы как для организации процессов управления малыми производствами и бизнесом, так и для управления крупными организациями и государством. Любая система, в том числе информационная, нуждается в оценке качества и эффективности функционирования. Анализ эффективности ИС дает возможность оценить целесообразность использования системы, увеличить прибыль или сократить затраты предприятия на его автоматизацию и повысить отдачу от используемой ИС [1].

В существующих методах в области совершенствования ИС в большинстве случаев не учитывается изменение эффективности во времени вследствие устаревания системы. Модели устаревания систем, в основном, основываются на утверждении об экспоненциальном законе устаревания либо закон вообще не упоминается [2,3]. Для того чтобы построить адекватную модель устаревания реальной системы надо учитывать скорость изменения процесса старения системы и готовность системы к реинжинирингу [4].

В работе предложены модели устаревания ИС, основанные на различных законах распределения случайных величин, а также приведены расчеты параметров эффективности системы при ее устаревании, показана явная зависимость этих параметров от вида закона распределения.

Цель статьи. Построить модели устаревания ИС и установить ее связи с видом закона