РОССИЙСКАЯ ФЕДЕРАЦИЯ

ПАТЕНТ

НА ИЗОБРЕТЕНИЕ

№ 2485705

СПОСОБ И СИСТЕМА ИДЕНТИФИКАЦИИ СЕТЕВЫХ ПРОТОКОЛОВ НА ОСНОВАНИИ ОПИСАНИЯ КЛИЕНТ-СЕРВЕРНОГО ВЗАИМОДЕЙСТВИЯ

Патентообладатель(ли): ***федеральное государственное бюджетное образовательное учреждение высшего профессионального образования "Санкт-Петербургский государственный политехнический университет" (ФГБОУ ВПО "СПбГПУ") (RU)***

Автор(ы): ***см. на обороте***

Заявка № **2012112436**

Приоритет изобретения **26 марта 2012 г.**

Зарегистрировано в Государственном реестре изобретений Российской Федерации ***20 июня 2013 г.***

Срок действия патента истекает **26 марта 2032 г.**

Руководитель Федеральной службы по интеллектуальной собственности

Б.П. Симонов

РОССИЙСКАЯ ФЕДЕРАЦИЯ

(19) **RU** (11) **2 485 705** (13) **C1**

(51) МПК
H04L 12/70 (2013.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: **2012112436/08, 26.03.2012**

(24) Дата начала отсчета срока действия патента:
26.03.2012

Приоритет(ы):
(22) Дата подачи заявки: **26.03.2012**

(45) Опубликовано: **20.06.2013** Бюл. № **17**

(56) Список документов, цитированных в отчете о поиске: **RU 2358397 C2, 10.06.2009. RU 94785 U1, 27.05.2010. US 2009/0141634 A1, 04.06.2009. US 2007/0058668 A1, 15.03.2007.**

Адрес для переписки:
195251, Санкт-Петербург, ул. Политехническая, 29, ФГБОУ ВПО "Санкт-Петербургский государственный политехнический университет" (ФГБОУ ВПО "СПбГПУ"), отдел интеллектуальной собственности

(72) Автор(ы):
Зегжда Петр Дмитриевич (RU),
Корт Семен Станиславович (RU),
Рудина Екатерина Александровна (RU)

(73) Патентообладатель(и):
федеральное государственное бюджетное образовательное учреждение высшего профессионального образования "Санкт-Петербургский государственный политехнический университет" (ФГБОУ ВПО "СПбГПУ") (RU)

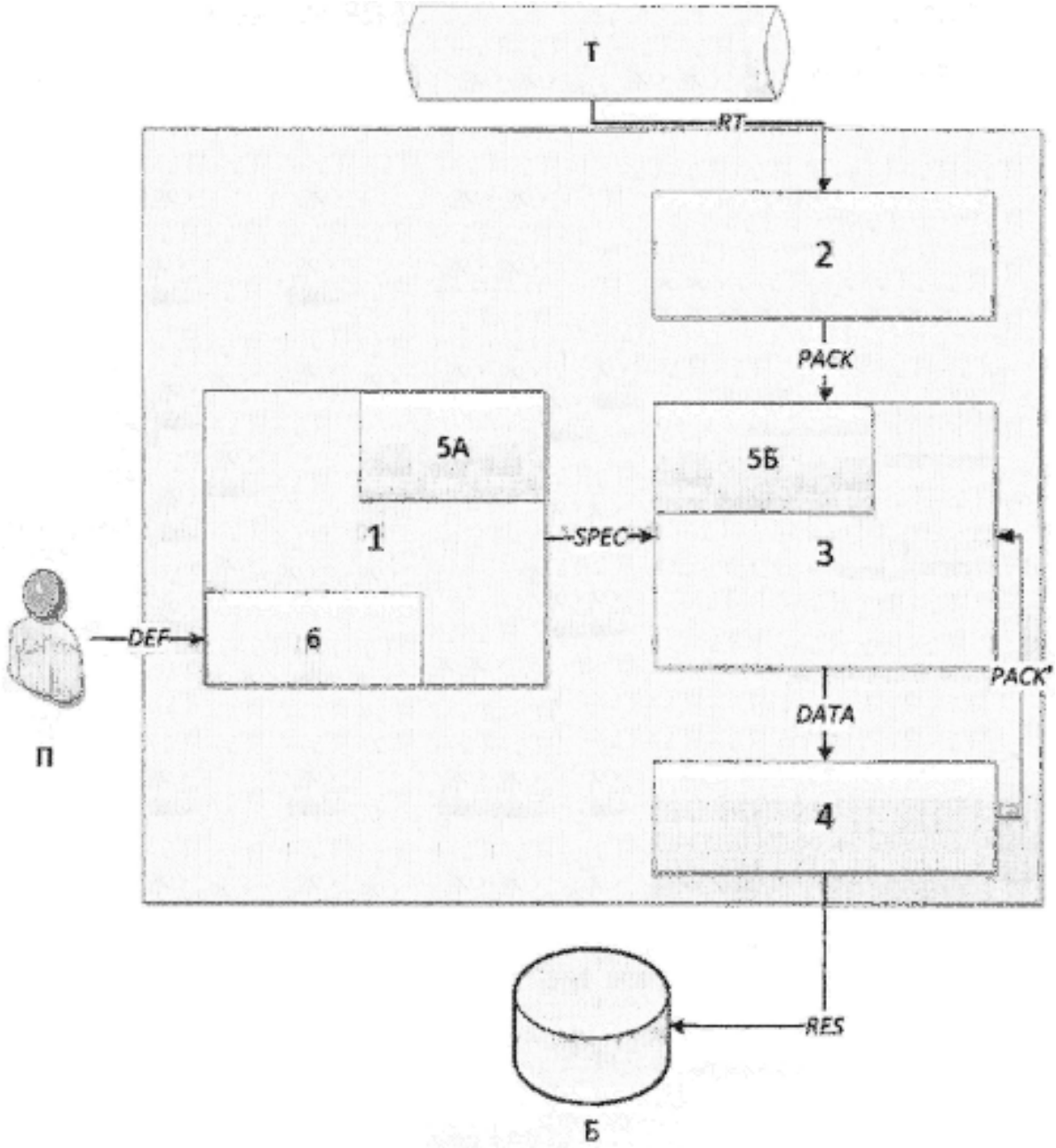
(54) СПОСОБ И СИСТЕМА ИДЕНТИФИКАЦИИ СЕТЕВЫХ ПРОТОКОЛОВ НА ОСНОВАНИИ ОПИСАНИЯ КЛИЕНТ-СЕРВЕРНОГО ВЗАИМОДЕЙСТВИЯ

(57) Реферат:

Изобретение относится к области компьютерных систем, а именно к описанию клиент-серверного взаимодействия, анализа протоколов и автоматизированного анализа сетевого трафика с целью идентификации сетевых протоколов. Техническим результатом является повышение эффективности идентификации сетевых протоколов. Способ идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия содержит описание известных протоколов клиент-серверного взаимодействия, сбор данных о двунаправленном сетевом взаимодействии между клиентом и сервером, одновременную идентификацию пакетов, относящихся к множеству сеансов двунаправленного взаимодействия между клиентом и сервером по идентифицируемым протоколам, анализ параметров и последовательности взаимодействия клиент-серверных компонент, идентификацию сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие. 2 н.п. ф-лы, 1 ил.

RU 2485705 C1

RU 2485705 C1



Фиг.1

RU 2485705 C1

RUSSIAN FEDERATION

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(19) **RU** (11) **2 485 705** (13) **C1**

(51) Int. Cl.
H04L 12/70 (2013.01)

(12) ABSTRACT OF INVENTION

(21)(22) Application: **2012112436/08, 26.03.2012**

(24) Effective date for property rights:
26.03.2012

Priority:
(22) Date of filing: **26.03.2012**

(45) Date of publication: **20.06.2013 Bull. 17**

Mail address:
195251, Sankt-Peterburg, ul. Politekhnicheskaja, 29, FGBOU VPO "Sankt-Peterburgskij gosudarstvennyj politekhnicheskij universitet" (FGBOU VPO "SPbGPU"), otdel intellektual'noj sobstvennosti

(72) Inventor(s):
Zegzhda Petr Dmitrievich (RU),
Kort Semen Stanislavovich (RU),
Rudina Ekaterina Aleksandrovna (RU)

(73) Proprietor(s):
federal'noe gosudarstvennoe bjudzhetnoe obrazovatel'noe uchrezhdenie vysshego professional'nogo obrazovanija "Sankt-Peterburgskij gosudarstvennyj politekhnicheskij universitet" (FGBOU VPO "SPbGPU") (RU)

(54) METHOD AND SYSTEM TO IDENTIFY NETWORK PROTOCOLS BASED ON DESCRIPTION OF CLIENT-SERVER INTERACTION

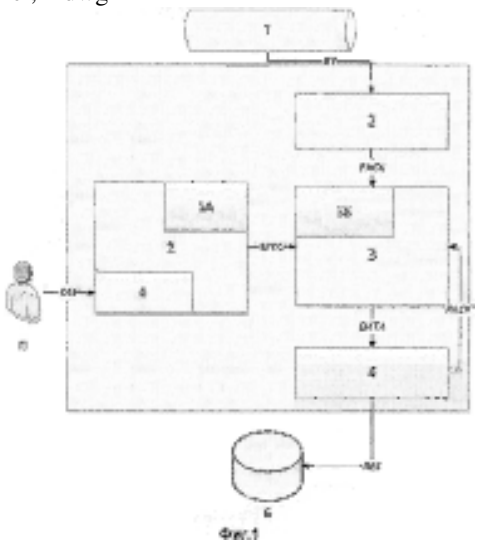
(57) Abstract:

FIELD: radio engineering, communications.

SUBSTANCE: method of network protocols identification on the basis of a description of client-server interaction contains a description of available protocols of client-server interaction, collection of data on a bidirectional network interaction between the client and the server, simultaneous identification of packets related to multiple sessions of bidirectional interaction between the client and the server by identified protocols, analysis of parameters and the sequence of interaction of client-server components, identification of network protocols, according to which the client-server interaction is carried out.

EFFECT: increased efficiency of network protocols identification.

2 cl, 1 dwg



RU 2485705 C1

RU 2485705 C1

Изобретение относится к области компьютерных систем, а именно к описанию клиент-серверного взаимодействия, анализу протоколов и автоматизированному анализу сетевого графика, в том числе с целью идентификации сетевых протоколов.

Wireshark, ранее известный как Ethereal, наиболее известный в мире анализатор сетевых протоколов, используемый как в индустрии информационных технологий, так и для образовательных целях. Wireshark выполняет сбор сетевого трафика непосредственно путем прослушивания сетевого интерфейса (посредством функций библиотеки libpcap) или путем чтения ранее сохраненного файла дампа трафика.

Wireshark, выполняющий анализ сетевых пакетов и идентификацию множества известных протоколов, предоставляет развитый пользовательский интерфейс, реализующий функции выбора источника сетевого трафика (один из установленных сетевых интерфейсов или указанный файл), просмотра суммарных данных прочитанных сетевых пакетов (протокол верхнего уровня, длина пакета), выделенных в соответствии со спецификацией протокола полей пакета, а также шестнадцатеричного дампа пакета. Пользовательский интерфейс также реализует функции фильтрации сетевых протоколов по значениям их отдельных полей. От других программ анализа сетевого трафика Wireshark отличают дополнительные функции: сбор пакетов, относящихся к одному соединению TCP (Transmission Control Protocol, протокол управления передачей), и выделение строковых данных, передаваемых посредством этого соединения.

На основе консольной версии Wireshark, программе TShark, построен аппаратно-программный анализатор трафика Cisco NX-OS Ethanalyzer. Устройство серии Cisco Nexus 7000 предназначено для выполнения наиболее полного контроля трафика, передаваемого в сети. Это аппаратный комплекс анализа сетевого трафика, базирующийся на современной модульной ОС Cisco NX-OS, построенной на ядре ОС Linux. Cisco Nexus 7000 реализует следующие ключевые возможности: сбор и анализ сетевого трафика в сети в режиме реального времени, фильтрация пакетов, получаемых при сборе трафика, по заданным выражениям формата BPF (Berkeley Packet Filter), анализ и идентификация сетевых протоколов, вывод подробной информации о проанализированных сетевых пакетах в консоль, анализ предварительно сохраненного в файл дампа сетевого трафика, подробный анализ отдельных пакетов сетевого трафика, отвечающих критериям, задаваемым специальными правилами.

Несмотря на то что среди функциональных возможностей описанного аппаратно-программного средства анализа сетевого трафика содержится возможность восстановления сеансов взаимодействия клиент-серверных компонент по протоколу TCP, а также возможность задания такого описания пакетов, которое может служить для идентификации сетевых протоколов, в целом отсутствует возможность для использования наблюдаемого клиент-серверного взаимодействия с целью идентификации сетевого протокола прикладного уровня. Это обусловлено отсутствием в указанных средствах универсального представления клиент-серверного взаимодействия и возможностей описания последовательностей групп пакетов, которые могут быть отнесены к такому взаимодействию (Cisco Nexus 7000 Series Architecture: Built-in Wireshark Capability for Network Visibility and Control, URL: http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/ps9512/white_paper_c11-554444.html).

Устройство, выбранное за прототип, описывает варианты реализации системы и способа адаптивной классификации сетевого трафика с использованием

исторического контекста.

Используется понятие «устройство мониторинга сети» (NMD). Это устройство может отслеживать сетевые соединения от клиента к серверу и обратно. Устройство может также извлекать информацию о пакетах различных уровней стека протоколов. Устройство может объединять, восстанавливать потоки данных, обмен которыми происходит на каждом уровне стека протоколов, а также выполнять при необходимости дешифрование данных. Устройство может пассивно прослушивать сетевой трафик или разделять и ретранслировать поток трафика как посредник.

Устройство при выполнении классификации трафика одновременно обрабатывает пакеты, относящиеся к множеству сетевых соединений. В одном из вариантов устройство выполняет классификацию сетевого трафика согласно коммуникационным сетевым протоколам. В другом варианте устройство категорирует трафик по типу передаваемой полезной нагрузки - файлы, потоковая аудиовидеоинформация, доступ к базам данных, интерактивное взаимодействие, онлайн-игры и пр. Наконец, классификатор может определять, когда трафик отвечает известным сетевым протоколам, таким как HTTP, SMTP, RTP, TDS и пр.

В статической системе, если трафик был однажды классифицирован, то впоследствии решения классификатора не пересматриваются. Адаптивная система классификации может пересматривать уже принятые решения. Система, которая не учитывает исторический контекст, может долгое время классифицировать трафик некорректно. Система, которая задействует исторический контекст, может применять уже принятые решения для одинаковых или похожих соединений. В одном из вариантов соединения являются похожими, если они имеют одни и те же идентифицирующие характеристики. Такими характеристиками могут быть аппаратный адрес, сетевой адрес или диапазон сетевых адресов, порты протоколов транспортного уровня или диапазоны портов и т.п. В другом из вариантов похожими являются соединения, установленные одновременно или в заданный промежуток времени.

В одном из вариантов метод исследует пакеты, относящиеся к множеству схожих соединений, при выполнении идентификации протоколов в сетевом трафике. Этот метод может также периодически исследовать и переоценивать уже выполненную идентификацию протоколов для таких соединений (US 2009141634, H04L 12/56).

Недостатками приведенного выше решения является то, что исторический контекст клиент-серверного взаимодействия используется для идентификации протоколов на основании эвристического положения о схожести сетевых пакетов и/или соединений и экстраполяции уже принятых решений на такие пакеты и соединения, а собственно решение задачи идентификации выполняется на основании уже известных подходов: соотнесении стандартных номеров портов и используемых сервисами, задействующими эти порты, протоколов; исследовании полезной нагрузки, в том числе с использованием сигнатурного поиска или языков описания протоколов; методах, основанных на обучении и т.п. Подход, основанный на использовании стандартных номеров портов, прост и эффективен, но не является достаточно надежным; исследование полезной нагрузки сообщений накладывает существенные вычислительные и временные ограничения на механизм идентификации. При этом методы, основанные на обучении и распознавании трафика с использованием кластеризации, непригодны для задачи идентификации известных протоколов, а языки, используемые для описания сетевых протоколов в коммерческих средствах распознавания, формируют аппарат, пригодный для сигнатурного распознавания

отдельных сообщений протоколов (часто даже на основе анализа портов), но не позволяющий проследить корректность взаимодействия клиента и сервера согласно протоколу. Наконец, формальные языки описания формируют подходящие для анализа описания протоколов, но такие описания, как правило, неприменимы к задаче идентификации протоколов сообщений в дампах сетевого трафика. В основу изобретения положена задача создания способа идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия и системы для осуществления этого способа, в которых вследствие использования формальной модели клиент-серверного взаимодействия, включающей механизм анализа такого взаимодействия, для создания описания сетевых протоколов прикладного уровня используется унифицированное представление сетевого протокола, а идентификация протоколов основана на комплексном анализе взаимодействия клиента и сервера согласно этому описанию, что позволяет улучшить распознавание шаблонов сетевого трафика и идентификацию Интернет-протоколов, а также позволяет упростить процесс выполнения описания протокола прикладного уровня (используемого при дальнейшем анализе трафика и идентификации протоколов) и свести к минимальному количество ошибок в таком описании.

Решение поставленной технической задачи обеспечивается тем, что в способе идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия, включающем описание известных протоколов клиент-серверного взаимодействия, сбор данных о двунаправленном сетевом взаимодействии между клиентом и сервером, одновременную идентификацию пакетов, относящихся к множеству сеансов двунаправленного взаимодействия между клиентом и сервером по идентифицируемым протоколам, анализ параметров и последовательности взаимодействия клиент-серверных компонент, идентификацию сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие, описание производят с помощью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент; идентификацию протоколов дополнительно проводят на основе единообразных описаний клиент-серверного взаимодействия рекурсивным образом для каждого слоя семиуровневой модели протоколов OSI, при этом механизмы дефрагментации, восстановления, распаковки, дешифрования и подобных преобразований выполняют автоматически после идентификации протокола каждого уровня, после чего выделенные данные передают как пакет верхнего уровня для последующей идентификации вышележащего протокола.

В системе идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия, включающей модуль описания известных протоколов клиент-серверного взаимодействия, модуль сбора данных о двунаправленном сетевом взаимодействии между клиентом и сервером, модуль анализа, модуль идентификации сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие, в модуль описания включены блок трансляции, реализующий возможность бинарного представления описания сетевых протоколов, блок интерактивного взаимодействия, выполненный с возможностью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного

языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент.

В предложенном решении реализация заявленного технического результата обеспечивается тем, что подход к идентификации протоколов основан не на разборе отдельных сообщений в сетевом трафике, а на комплексном анализе взаимодействия клиента и сервера согласно указанному протоколу. При этом способ описания взаимодействия клиента и сервера, включая последовательность обмена сообщениями и соответствующие ограничения, основан на унифицированном представлении сетевых протоколов, что позволяет сделать реализацию заявленного способа расширяемой как в отношении множества идентифицируемых протоколов, так и в отношении диалектов этих протоколов и специфических сценариев использования сетевых протоколов различными сервисами и программами. Более того, возможность описания произвольного протокола (и последующего анализа согласно этому описанию) строго обоснована путем проведения формального доказательства существования обобщенного механизма анализа взаимодействия клиента и сервера.

Перечисленные преимущества позволят улучшить распознавание шаблонов сетевого трафика и идентификацию Интернет-протоколов. Подобные решения не применяются ни одним из известных средств анализа сетевого трафика.

Изобретение поясняется с помощью фиг.1, на которой приведена система идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия.

Система идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия включает модуль описания известных протоколов клиент-серверного взаимодействия 1, модуль сбора данных о двунаправленном сетевом взаимодействии между клиентом и сервером 2, модуль анализа 3, модуль идентификации сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие 4. В модуль описания 1 включены блок трансляции 5А, реализующий возможность бинарного представления описания сетевых протоколов, блок интерактивного взаимодействия 6, выполненный с возможностью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент. В другом варианте, блок трансляции заменяется на блок интерпретации описания 5Б, включаемый в модуль анализа. Конкретное техническое исполнение зависит от свойств языка описания взаимодействия: компилируемый, интерпретируемый.

Система выполняется как программно-аппаратный блок, который для идентификации сетевых протоколов подключается к шине передачи данных в режиме прослушивания сетевого трафика, либо в разрыв сетевого соединения с ретрансляцией трафика между передающей и принимающей сторонами.

Способ идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия осуществляют следующим образом.

Исходно составляют описание известных протоколов клиент-серверного взаимодействия с использованием блока интерактивного взаимодействия 6 (связь DEF пользователя П и модуля описания 1 на схеме системы). Описание производят с помощью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент. При наличии блока трансляции 5А выполняют проверку синтаксических ошибок и сборку бинарного представления для описания клиент-серверного взаимодействия. Полученный образ бинарного представления подают на вход модуля анализа. При отсутствии блока трансляции на вход модуля анализа подается непосредственно созданное описание (связь SPEC на схеме системы).

Затем осуществляют сбор данных о двунаправленном сетевом взаимодействии между клиентом и сервером (связь RT источника сетевого трафика Т и модуля сбора данных 2 на схеме системы). Модуль сбора данных 2 осуществляет первичное разбиение трафика на пакеты с их возможной фильтрацией и подает на вход модуля анализа пакеты в том порядке, в котором они поступили из сети либо в другом (связь PACK модуля сбора данных 2 и модуля анализа 3 на схеме системы). Таким образом, с использованием модуля анализа 3 обеспечивают одновременную обработку пакетов, относящихся к множеству сеансов двунаправленного взаимодействия между клиентом и сервером по идентифицируемым протоколам. С использованием модуля анализа 3 осуществляют разбор сетевых пакетов в соответствии с интерпретируемым с использованием блока 5Б описанием протоколов (или в соответствии с бинарным представлением этого описания, без интерпретации), выделяют существенные характеристики пакета для всех протоколов, к которым может быть отнесен этот пакет, определяют место сетевого пакета в сеансах взаимодействия клиента и сервера по этим протоколам. Характеристики пакета для всех возможных вариантов протоколов, к которым может быть отнесен пакет, а также параметры текущего контекста взаимодействия клиента и сервера по этим протоколам передают модулю идентификации 4 (связь DATA на схеме системы).

С использованием модуля идентификации 4 осуществляют анализ параметров и последовательности взаимодействия клиент-серверных компонент и проводят идентификацию сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие, причем идентификацию протоколов дополнительно проводят на основе единообразных описаний клиент-серверного взаимодействия рекурсивным образом для каждого слоя семиуровневой модели протоколов OSI, при этом механизмы дефрагментации, восстановления, распаковки, дешифрования и подобных преобразований выполняют автоматически после идентификации протокола каждого уровня, после чего выделенные данные передают как пакет верхнего уровня для последующего анализа и идентификации вышележащего протокола (связь PACK' на схеме системы). Полученные результаты модуль идентификации 4 сохраняет во внешней базе данных или передает для дальнейшей обработки (связь RES модуля идентификации 4 и базы данных Б на схеме системы).

Результаты идентификации трафика сетевых протоколов наиболее часто используются в области защиты информации. В частности, задача идентификации может использоваться как одна из подзадач при обнаружении атак, являясь

компонентом автоматизированных систем обнаружения атак. Включение данного компонента в систему обнаружения атак позволит как повысить точность обнаружения методов, основанных на информации об используемом протоколе, так и повысить адаптивность системы обнаружения атак.

Кроме того, идентификация трафика прикладных протоколов может служить для идентификации и блокировки нежелательного трафика. К нежелательному трафику можно отнести трафик, генерируемый вредоносным программным обеспечением (ВПО), а также трафик, противоречащий политике безопасности, принятой в организации (например, трафик пиринговой или игровой сети). В этой связи целесообразна работа модуля идентификации трафика прикладных Интернет-протоколов с межсетевым экраном. Идентификация трафика прикладных Интернет-протоколов также может использоваться в биллинговых системах для более точного учета типов трафика, потребляемого пользователем.

Перспективы использования способа и системы идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия включают, кроме уже указанных вариантов применения, выполнение анализа безопасности сетевых протоколов на основании их описания и поиска скрытых каналов утечки информации в полях этих протоколов, выявление нецелевого использования сетевых протоколов, в том числе вредоносным программным обеспечением, а также профилирование и обнаружение аномалий в обращениях к сетевым сервисам в Интернет.

Формула изобретения

1. Способ идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия, включающий описание известных протоколов клиент-серверного взаимодействия, сбор данных о двунаправленном сетевом взаимодействии между клиентом и сервером, одновременную идентификацию пакетов, относящихся к множеству сеансов двунаправленного взаимодействия между клиентом и сервером по идентифицируемым протоколам, анализ параметров и последовательности взаимодействия клиент-серверных компонент, идентификацию сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие, отличающийся тем, что описание производят с помощью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент; идентификацию протоколов дополнительно проводят на основе единообразных описаний клиент-серверного взаимодействия рекурсивным образом для каждого слоя семиуровневой модели протоколов OSI, при этом механизмы дефрагментации, восстановления, распаковки, дешифрования и подобных преобразований выполняют автоматически после идентификации протокола каждого уровня, после чего выделенные данные передают как пакет верхнего уровня для последующей идентификации вышележащего протокола.

2. Система идентификации сетевых протоколов на основании описания клиент-серверного взаимодействия, включающая модуль описания известных протоколов клиент-серверного взаимодействия, модуль сбора данных о двунаправленном сетевом взаимодействии между клиентом и сервером, модуль анализа, модуль идентификации сетевых протоколов, по которым осуществляется клиент-серверное взаимодействие,

отличающаяся тем, что в модуль описания включены блок трансляции, реализующий возможность бинарного представления описания сетевых протоколов, блок интерактивного взаимодействия, выполненный с возможностью унифицированного представления взаимодействия клиент-серверных компонент, в том числе с использованием специализированного языка, автоматической или полуавтоматической программной генерации описаний взаимодействия клиент-серверных компонент и шаблонов или заготовок типовых сценариев клиент-серверного взаимодействия, с возможностью добавления новых, отвечающих не идентифицируемым ранее протоколам, описаний взаимодействия клиент-серверных компонент.