РОССИЙСКАЯ ФЕДЕРАЦИЯ

(19) **RU** (11) **2 701 040**(13) **C1**

(51) МПК
G06F 21/56 (2013.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
G06F 21/50 (2019.05); G06F 21/56 (2019.05); H04L 63/145 (2019.05); H04L 67/02 (2019.05)

(21)(22) Заявка: **2018147431, 28.12.2018**

(24) Дата начала отсчета срока действия патента:
28.12.2018

Дата регистрации:
24.09.2019

Приоритет(ы):
(22) Дата подачи заявки: **28.12.2018**

(45) Опубликовано: **24.09.2019** Бюл. № **27**

Адрес для переписки:
196066, Санкт-Петербург, А/Я 34, Пронину В.О.

(72) Автор(ы):
Калинин Александр Сергеевич (RU)

(73) Патентообладатель(и):
Общество с ограниченной ответственностью "Траст" (RU)

(56) Список документов, цитированных в отчете о поиске: **RU 2622870 C2, 20.06.2017. RU 2446459 C1, 27.03.2012. KR 10-2007-0049514 A, 11.05.2007. KR 10-1514984 B1, 24.04.2015.**

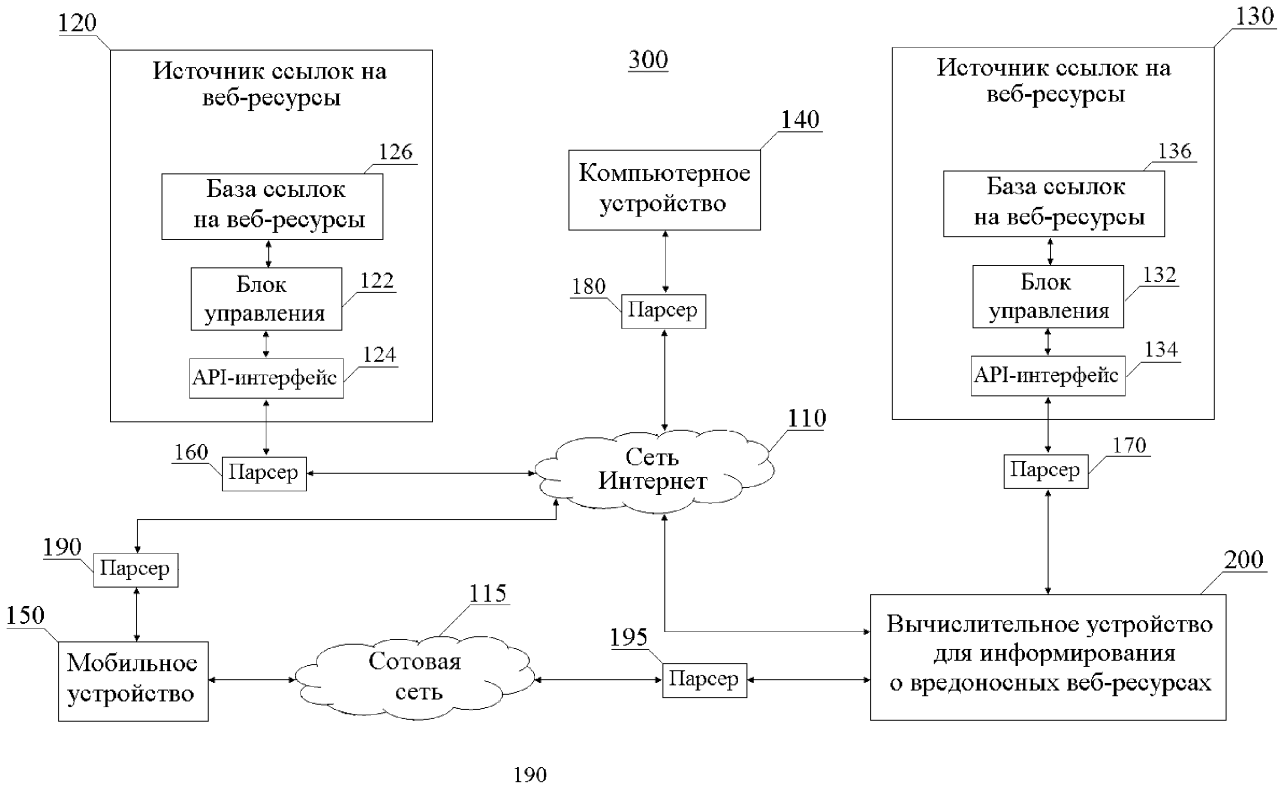
(54) СПОСОБ И ВЫЧИСЛИТЕЛЬНОЕ УСТРОЙСТВО ДЛЯ ИНФОРМИРОВАНИЯ О ВРЕДОНОСНЫХ ВЕБ-РЕСУРСАХ

(57) Реферат:

Изобретение относится к способу и вычислительному устройству для информирования о вредоносных веб-ресурсах. Техническим результатом является повышение эффективности информирования уполномоченных субъектов о выявленных веб-ресурсах с вредоносным и/или незаконным контентом. Предложенный способ включает выполнение операций, согласно которым: получают ссылки на множество веб-ресурсов; выявляют вредоносные веб-ресурсы в указанном множестве веб-ресурсов; устанавливают веб-ресурсы, связанные с каждым из выявленных вредоносных веб-ресурсов; выявляют вредоносные веб-ресурсы в установленных связанных веб-ресурсах; устанавливают по меньшей мере один уполномоченный субъект, связанный с каждым из выявленных вредоносных веб-ресурсов; формируют по меньшей мере один отчет по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом; отправляют каждый сформированный отчет в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. 2 н. и 12 з.п. ф-лы, 3 ил.

RU 2701040 C1

RU 2701040 C1



Фиг. 1

RU 2701040 C1

RU 2701040 C1

RUSSIAN FEDERATION

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(19) **RU** (11) **2 701 040**(13) **C1**

(51) Int. Cl.
G06F 21/56 (2013.01)

(12) ABSTRACT OF INVENTION

(52) CPC
G06F 21/50 (2019.05); G06F 21/56 (2019.05); H04L 63/145 (2019.05); H04L 67/02 (2019.05)

(21)(22) Application: **2018147431, 28.12.2018**

(24) Effective date for property rights:
28.12.2018

Registration date:
24.09.2019

Priority:
(22) Date of filing: **28.12.2018**

(45) Date of publication: **24.09.2019** Bull. № **27**

Mail address:
196066, Sankt-Peterburg, A/YA 34, Proninu V.O.

(72) Inventor(s):
Kalinin Alexander Sergeevich (RU)

(73) Proprietor(s):
Trust Ltd. (RU)

(54) METHOD AND A COMPUTER FOR INFORMING ON MALICIOUS WEB RESOURCES

(57) Abstract:

FIELD: information technology.

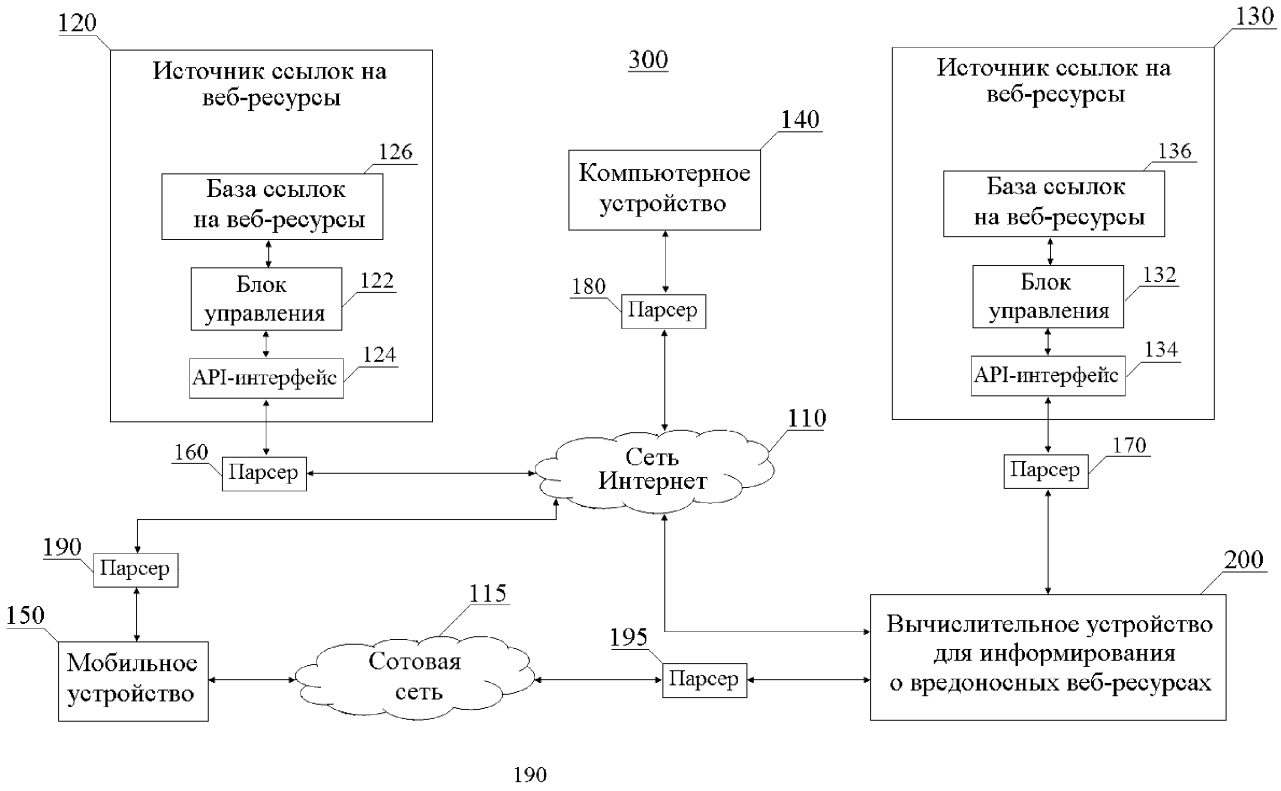
SUBSTANCE: invention relates to a method and a computer for informing on malicious web resources. Proposed method comprises the following operations: obtaining links to a plurality of web resources; detecting malicious web resources in said plurality of web resources; establishing web resources associated with each of detected malicious web resources; detecting malicious web resources in installed linked web resources; establishing at least one authorized person associated with each of the malicious web resources detected; generating at least one report for at least one of the established authorized entities based on data on detected malicious web resources associated with that authorized entity; each generated report is sent to the corresponding authorized entity on the basis of contact data of this authorized subject.

EFFECT: high efficiency of informing authorized entities on detected web resources with malicious and/or illegal content.

14 cl, 3 dwg

RU 2701040 C1

RU 2701040 C1



Фиг. 1

RU 2701040 C1

RU 2701040 C1

ОБЛАСТЬ ТЕХНИКИ

Настоящее изобретение относится к области информационной безопасности, в частности к способу и вычислительному устройству для информирования о вредоносных веб-ресурсах.

УРОВЕНЬ ТЕХНИКИ

Для размещения веб-ресурса в сети Интернет необходимо загрузить его файлы на веб-сервер хостинг-провайдера, который постоянно подключен к сети Интернет и на котором запущено специальное программное обеспечение, необходимое для обработки запросов к веб-ресурсу. При обращении к хостинг-провайдеру собственник веб-ресурса получает персональную учётную запись, а веб-ресурс получает IP-адрес, выданный этим хостинг-провайдером, при этом выданный IP-адрес поставлен в соответствие с учетной записью, выданной собственнику веб-ресурса. Таким образом, на основании IP-адреса веб-ресурса можно по меньшей мере определить хостинг-провайдера, выдавшего учетную запись, использующую этот IP-адрес. Следует отметить, что хостинг-провайдеры обычно предоставляют свои услуги на определенных условиях, согласно которым хостинг-провайдер может, помимо прочего, приостановить оказание своих услуг в случае размещения на своем веб-сервере веб-ресурса с вредоносным и/или незаконным контентом, что подразумевает блокировку такого вредоносного веб-ресурса хостинг-провайдером по его IP-адресу, в результате чего этот веб-ресурс перестает быть доступным для пользователей сети Интернет.

Для удобства запоминания адресного пространства веб-ресурса и обеспечения возможности перехода от одного хостинг-провайдера к другому хостинг-провайдеру без необходимости в изменении единого указателя веб-ресурса («URL»), при вводе которого в адресную строку веб-браузера пользователь может обратиться к указанному веб-ресурсу, владелец веб-ресурса может воспользоваться возможностями системы доменных имен, согласно которой такому веб-ресурсу может быть присвоено доменное имя, регистрируемое у регистратора доменных имён, при этом в качестве регистрируемого доменного имени может быть выбрано любое сочетание букв и цифр, которое не нарушает правил выбранной доменной зоны. Для автоматического преобразования зарегистрированного доменного имени веб-ресурса в его IP-адрес, обычно указываемый при регистрации доменного имени, используют DNS-сервера, на которых хранится информация о соответствии тех или иных доменных имен с IP-адресами веб-ресурсов, выданными хостинг-провайдерами. Следует отметить, что регистраторы доменных имён аналогично хостинг-провайдерам также обычно предоставляют свои услуги на определенных условиях, согласно которым регистратор доменных имён может, помимо прочего, заблокировать доменное имя, зарегистрированное этим регистратором доменных имен, в случае, если он, например, узнает о том, что данное доменное имя принадлежит веб-ресурсу с вредоносным и/или незаконным контентом. Таким образом, в случае блокировки регистратором доменных имён конкретного доменного имени через некоторый период времени доменное имя веб-ресурса, введенное пользователем в адресную строку браузера, не будет преобразовываться в IP-адрес, в результате чего подключения к запрашиваемому веб-ресурсу не произойдет (т.е. пользователь не сможет зайти на веб-ресурс), а браузер выдаст пользователю сообщение об ошибке, такое как, например, сообщение «Не удалось найти IP-адрес сервера».

Таким образом, одним из наиболее существенных поводов приостановления оказания вышеописанных услуг хостинг-провайдером и/или регистратором доменных имен является получение ими сведений о том, что связанный с ними веб-ресурс имеет

вредоносный характер, то есть содержит вредоносный и/или незаконный контент.

Для выявления вредоносных веб-ресурсов и направления уполномоченным субъектам уведомлений о выявленных вредоносных веб-ресурсах для их последующего блокирования используют различные интеллектуальные системы.

Один из иллюстративных примеров такой интеллектуальной системы описан в патенте KR 101514984 B1 (опубл. 24.04.2015; G06F 21/56).

В частности, в патенте KR 101514984 раскрыта система для выявления вредоносного кода, распространяемого веб-страницами. Система по KR 101514984 выполнена с возможностью подключения к веб-страницам различных веб-ресурсов для реализации различных действий пользователей, возможностью выявления любой поведенческой модели, связанной с распространением вредоносного кода, и возможностью направления уведомления в хостинг-сервер, на котором размещен такой вредоносной код, для обеспечения возможности принятия необходимых мер до распространения этого вредоносного кода в соответствии с выявленной поведенческой моделью.

Еще в одном патенте KR20070049514 (опубл. 11.05.2007; G06F 11/00) раскрыта система для выявления вредоносного кода, содержащая блок для получения ссылок на множество веб-ресурсов; базу данных для сохранения сведений об известном вредоносном коде; поисковый блок для поиска вредоносного кода среди полученных ссылок путем выявления, соответствует ли подозрительный код вредоносному коду, сведения о котором сохранены в базе данных; и уведомительный блок для направления уведомления о нахождении вредоносного кода на веб-ресурс, на котором этот вредоносный код был найден поисковым блоком, для последующего удаления исходного кода для генерирования html-документов, программы, изображения, всплывающего окна и т.п., внедренных в подозрительный код, или блокирования домена, через который происходит распространение вредоносного кода.

Следует отметить, что известные системы информирования позволяют лишь направлять отдельное уведомление об одном вредоносном веб-ресурсе, выявленном при осуществлении последовательной проверки анализируемых веб-ресурсов на вредоносность, в один уполномоченный субъект, связанный с этим вредоносным веб-ресурсом, при этом существует вероятность того, что такое уведомление будет проигнорировано уполномоченным субъектом, в результате чего такой веб-ресурс будет продолжать работать на злоумышленников, распространяя в сети вредоносный и/или незаконный контент. Следует отметить, что известные системы информирования не и используют средства и механизмы, позволяющие единовременно информировать широкий круг уполномоченных субъектов, которые могут повлиять на принятие решения о блокировке веб-ресурса с вредоносным и/или незаконным контентом или которые могут принять такое решение, о вредоносных веб-ресурсах, имеющих схожие признаки подозрительности, имеющих схожую вредоносную активность и/или принадлежащих одному и тому же злоумышленнику или одной и той же группе злоумышленников.

Таким образом, очевидна потребность в дальнейшем совершенствовании средств для информирования о вредоносных веб-ресурсах, в частности для улучшения эффективности информирования уполномоченных субъектов о выявленных веб-ресурсах с вредоносным и/или незаконным контентом.

Следовательно, техническая проблема, решаемая настоящим изобретением, состоит в создании усовершенствованных средств для информирования о вредоносных веб-ресурсах, в которых по меньшей мере частично устранён обозначенный выше недостаток известных средств информирования, заключающийся низкой эффективности

информирования уполномоченных субъектов в выявленных веб-ресурсах с вредоносным и/или незаконным контентом.

РАСКРЫТИЕ СУЩНОСТИ

Вышеупомянутая техническая проблема решена в одном из аспектов настоящего изобретения, согласно которому предложен способ информирования о вредоносном характере веб-ресурсов согласно настоящему изобретению, выполняемый на вычислительном устройстве, при этом согласно указанному способу: получают ссылки на множество веб-ресурсов; выявляют вредоносные веб-ресурсы в указанном множестве веб-ресурсов; устанавливают веб-ресурсы, связанные с каждым из выявленных вредоносных веб-ресурсов; выявляют вредоносные веб-ресурсы среди множества установленных связанных веб-ресурсов; устанавливают по меньшей мере один уполномоченный субъект, связанный с каждым из выявленных вредоносных веб-ресурсов; формируют по меньшей мере один отчет по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом; отправляют каждый сформированный отчет в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта.

В одном из вариантов реализации настоящего изобретения для получения ссылок на множество веб-ресурсов осуществляют по меньшей мере одну из следующих операций, согласно которым: направляют запрос по меньшей мере в один источник ссылок для получения из него по меньшей мере одной ссылки на веб-ресурс; принимают сообщения по меньшей мере от одного вычислительного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки веб-ресурс; принимают сообщения по меньшей мере от одного мобильного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки на веб-ресурс; и вводят поисковые запросы по меньшей мере в одну поисковую систему с использованием конкретного перечня ключевых слов для выявления контекстной рекламы в результатах поиска, полученных в ответ на каждый поисковых запрос в каждой из этих поисковых систем, с обеспечением извлечения по меньшей мере одной ссылки на веб-ресурс из выявленной контекстной рекламы.

Еще в одном варианте реализации настоящего изобретения для установления связанных веб-ресурсов определяют по меньшей мере одно из следующего: имеют ли доменные имена веб-ресурсов схожее написание; зарегистрированы ли доменные имена на одно и то же лицо; указаны ли для зарегистрированных доменных имен веб-ресурсов одни и те же персональные данные регистранта; находятся ли доменные имена веб-ресурсов по одному и тому же IP-адресу; и имеют ли ссылки, соответствующие веб-ресурсам, один и тот же или похожий единый указатель веб-ресурса “URL”.

В другом варианте реализации настоящего изобретения для установления связи веб-ресурсов осуществляют по меньшей мере следующие операции, согласно которым: создают математическую модель в виде графа, согласно которой вершины создаваемого графа соответствуют по меньшей мере первому веб-ресурсу и по меньшей мере второму веб-ресурсу, а ребра графа представляют собой связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом по меньшей мере по одному параметру веб-ресурса, общему по меньшей мере для первого веб-ресурса и по меньшей мере для второго веб-ресурса, при этом количество связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами ограничено заданным пороговым значением; присваивают, посредством известного алгоритма машинного обучения, веса связям по меньшей мере между первым веб-ресурсом и

вторым веб-ресурсом на основании параметра первого веб-ресурса и второго веб-ресурса; определяют коэффициент связи как отношение количества связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами и веса каждой связи по одному параметру веб-ресурса между первым веб-ресурсом и вторыми веб-ресурсами; и удаляют связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом в случае, если значение определенного коэффициента связи меньше заданного порогового значения.

В некоторых вариантах реализации настоящего изобретения для выявления вредоносных веб-ресурсов устанавливают, совпадает ли каждая полученная ссылка по меньшей мере частично с одной из известных вредоносных ссылок.

В других вариантах реализации настоящего изобретения для выявления вредоносных веб-ресурсов в дополнение к операции, согласно которой устанавливают, совпадает ли каждая полученная ссылка по меньшей мере частично с одной из известных вредоносных ссылок, осуществляют по меньшей мере одну из следующих операций, согласно которым: анализируют доменное имя веб-ресурса на вредоносность с использованием по меньшей мере одной методики анализа доменных имен; получают с веб-ресурса по меньшей мере один файл для его анализа на вредоносность с использованием по меньшей мере одной методики анализа файлов; и получают html-код веб-ресурса для его анализа на вредоносность с использованием по меньшей мере одной методики анализа html-кода.

В некоторых других вариантах реализации настоящего изобретения при анализе доменного имени веб-ресурса на вредоносность дополнительно устанавливают, совпадает ли этом анализируемое доменное имя с одним из известных вредоносных доменных имен.

В иных вариантах реализации настоящего изобретения при анализе файла, полученного с веб-ресурса, дополнительно вычисляют хеш-сумму анализируемого файла, полученного с веб-ресурса, и устанавливают, совпадает ли вычисленная хеш-сумма анализируемого файла с хеш-суммой одного из известных вредоносных файлов.

Еще в других вариантах реализации настоящего изобретения при анализе полученного html-кода веб-ресурса осуществляют поиск в указанном html-коде конкретных ключевых слов, указывающих на вредоносный характер веб-ресурса.

Согласно одному из вариантов реализации настоящего изобретения, при установлении уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, определяют владельца, администратора, хостинг-провайдера и/или регистратора доменных имён, связанных с этим вредоносным веб-ресурсом.

Согласно другому варианту реализации настоящего изобретения предложенный способ может включать дополнительный этап, согласно которому устанавливают тип угрозы из заданного набора типов угроз для каждого выявленного вредоносного веб-ресурса, а при формировании каждого отчета используют шаблон из заданного набора шаблонов отчетов, при этом каждый шаблон соответствует одному из установленных типов угроз и одному из установленных уполномоченных субъектов.

В другом варианте реализации настоящего изобретения количество отчетов, сформированных для каждого уполномоченного субъекта, может соответствовать количеству установленных типов угроз.

Еще в одном варианте реализации настоящего изобретения в каждый сформированный отчет могут быть дополнительно добавлены доказательства вредоносности каждого веб-ресурса, сведения о котором содержатся в этом отчете.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Прилагаемые чертежи, которые приведены для лучшего понимания сущности настоящего изобретения, включены в данный документ для иллюстрации нижеописанных вариантов реализации настоящего изобретения. Прилагаемые чертежи в совокупности с приведенным ниже описанием служат для пояснения сущности настоящего изобретения. На чертежах:

на фиг. 1 схематически показана система для информирования о вредоносных веб-ресурсах;

на фиг. 2 показан один из вариантов реализации устройства для информирования о вредоносных веб-ресурсах;

на фиг. 3 показана блок-схема способа информирования о вредоносных веб-ресурсах.

ОСУЩЕСТВЛЕНИЕ

Ниже описаны некоторые примеры возможных вариантов реализации настоящего изобретения, при этом не следует считать, что приведенное ниже описание определяет или ограничивает объем настоящего изобретения.

Система для информирования о вредоносных веб-ресурсах

На фиг. 1 схематически показана система 300 для информирования о вредоносных веб-ресурсах, содержащая вычислительное устройство 200 для информирования о вредоносных веб-ресурсах, источник 120 ссылок на веб-ресурсы, содержащий ссылки на потенциально вредоносные веб-ресурсы, источник 130 ссылок на веб-ресурсы, содержащий ссылки на потенциально вредоносные веб-ресурсы, а также компьютерное устройство 140, мобильное устройство 150 и сеть 110 Интернет.

В одном из вариантов реализации настоящего изобретения источник 120 ссылок на веб-ресурсы может представлять собой сайт «antiphishing.org» со ссылками на известные вредоносные веб-ресурсы, а источник 130 ссылок на веб-ресурсы может представлять собой сайт «antifraud.org» со ссылками на известные вредоносные веб-ресурсы. В таком варианте реализации настоящего изобретения все потоки данных, передаваемые от источника 120 ссылок, и все потоки данных, передаваемые от источника 130, должны быть ассоциированы соответственно с уникальным идентификатором, присвоенным источнику 120 ссылок, и уникальным идентификатором, присвоенным источнику 130 ссылок, при этом нижеописанное вычислительное устройство 200 должно быть предварительно запрограммировано или настроено на идентификацию потоков данных от таких источников ссылок, в частности от источников 120, 130 ссылок, на основании их уникальных идентификаторов, содержащихся в этих потоках данных и предварительно известных вычислительному устройству 200.

Источник 120 ссылок содержит, помимо прочего, блок 122 управления, API-интерфейс 124, обеспечивающий возможность взаимодействия с блоком 122 управления, и базу 126 ссылок на веб-ресурсы, хранящую, например, собранные из сторонних источников ссылки на веб-ресурсы, содержащие потенциально вредоносный и/или незаконный контент, и вспомогательную информацию, которая эти ссылки атрибутирует.

Источник 130 ссылок также содержит, помимо прочего, блок 132 управления, API-интерфейс 134, обеспечивающий возможность взаимодействия с блоком 132 управления, и базу 136 ссылок на веб-ресурсы, хранящую, например, собранные из сторонних источников ссылки на веб-ресурсы, содержащие потенциально вредоносный и/или незаконный контент, и вспомогательную информацию, которая эти ссылки атрибутирует.

Вычислительное устройство 200 согласно настоящему изобретению подключено к источнику 120 ссылок и источнику 130 ссылок посредством соответственно парсера 160, выполненного с возможностью подключения к API-интерфейсу 124 источника 120 ссылок и предварительно настроенного на работу с ним, и парсера 170, выполненного

с возможностью подключения к API-интерфейсу 134 источника 130 ссылок и предварительно настроенного на работу с ним, при этом парсер 160 соединен с возможностью обмена данными с вычислительным устройством 200 с использованием сети 110 Интернет, а парсер 170 соединен проводным способом с возможностью обмена данными непосредственно с самим вычислительным устройством 200. Следует отметить, что каждый из API-интерфейса 124 и API-интерфейса 134 может иметь свой собственный синтаксис команд, так что парсер 160, работающий с API-интерфейсом 124, должен быть предварительно запрограммирован понимать синтаксис команд этого API-интерфейса 124, а парсер 170, работающий с API-интерфейсом 134, должен быть аналогичным образом предварительно запрограммирован понимать синтаксис команд API-интерфейса 134, при этом настройка парсера 160 и парсера 170 на работу соответственно с API-интерфейсом 124 и API-интерфейсом 134 происходит при первоначальном подключении вычислительного устройства 200 к источникам 120, 130 ссылок. Следует отметить, что парсеры 160, 170 могут быть реализованы каждый в виде отдельного сервера или иного известного вычислительного устройства.

Вычислительное устройство 200 согласно настоящему изобретению выполнено с возможностью направления запросов в каждый из источников 120, 130 ссылок, например запросов на передачу в вычислительное устройство 200 по меньшей мере части ссылок на потенциально вредоносные веб-ресурсы, все ссылки на потенциально вредоносные веб-ресурсы или только ссылки на потенциально вредоносные веб-ресурсы, сохраненные соответственно в базе 126 ссылок или базе 136 ссылок за заданный период времени. Благодаря использованию парсеров 160, 170, предварительно настроенных на работу соответственно с API-интерфейсами 124, 134, запросы, направленные вычислительным устройством 200 в источники 120, 130 ссылок, будут содержать команды, понятные соответственно блокам 122, 132 управления, в результате чего эти блоки 122, 132 управления смогут надлежащим образом обработать указанные запросы и отреагировать на них, в частности передать запрашиваемые ссылки на потенциально вредоносные веб-ресурсы в вычислительное устройство 200, от которого эти запросы были приняты.

В частности, в ответ на принятые запросы блоки 122, 132 управления получают доступ соответственно к базе 126 ссылок и базе 136 ссылок с обеспечением извлечения из них запрашиваемых ссылок на потенциально вредоносные веб-ресурсы (также называемых в данном документе потенциально вредоносными ссылками) и передачи, посредством API-интерфейсов 124, 134, извлеченных потенциально вредоносных ссылок соответственно в парсеры 160, 170, при этом парсер 160 обеспечивает возможность обработки выходного потока данных от API-интерфейса 124 для извлечения из него потенциально вредоносных ссылок, запрошенных из источника 120 ссылок, а парсер 170 обеспечивает возможность обработки выходного потока данных от API-интерфейса 134 для извлечения из него потенциально вредоносных ссылок, запрошенных из источника 130 ссылок. Следует отметить, что для извлечения необходимых ссылок на веб-ресурсы из обрабатываемого потока данных парсеры 160, 170 используют каждый соответствующее регулярное выражение из заданного набора известных ему регулярных выражений. В частности, выходной поток данных от любого из API-интерфейсов 124, 134 содержит как сами потенциально вредоносные ссылки, так и идентификационные данные, описывающие передаваемые потенциально вредоносные ссылки, например дата и время поступления ссылок в базу ссылок на веб-ресурсы, идентификационные данные источника ссылок и/или прочие необходимые атрибуты этих передаваемых ссылок. Выходной поток данных от любого из API-интерфейсов 124, 134 обычно

представляет собой совокупность символов в виде строки символов с заданным форматом описания, которая разделена на структурные элементы с использованием некоторого предварительно заданного символа, например символа «#» (решетка), при этом формат записи такой строки символов известен парсерам 160, 170, благодаря тому, что они предварительно запрограммированы или настроены на работу с соответствующим одним из API-интерфейсов 124, 134. В частности, парсерам 160, 170 должно быть известно ключевое слово, ключевой символ или ключевая метка, указывающие на наличие следующей за ними ссылки, и иные ключевые слова/символы/метки, обычно используемые в получаемых строках символов для указания на наличие тех или иных идентификационных сведений, следующих за такими ключевыми словами. При получении подобных строк символов от API-интерфейсов 124, 134 парсеры 160, 170 соответственно извлекают из этих полученных строк, каждая из которых разделена на известную совокупность структурных элементов, потенциально вредоносные ссылки на веб-ресурсы и по меньшей мере некоторые из идентификационных данных, описывающих эти потенциально вредоносные ссылки, с обеспечением передачи извлеченных потенциально вредоносных ссылок на веб-ресурсы в вычислительное устройство 200 для их последующего анализа, особенности которого будут описаны ниже.

В случае направления в один из источников 120, 130 ссылок запроса на передачу в вычислительное устройство 200 потенциально вредоносных ссылок, сохраненных соответственно в базе 126 ссылок или базе 136 ссылок за заданный период времени, например все последние сохраненные потенциально вредоносные ссылки, начиная с определенного момента времени, например за последние несколько минут, часов, дней, недель, месяцев и т.п. в зависимости от поставленных задач, при этом такие запрашиваемые потенциально вредоносные ссылки устанавливают, например, с использованием показаний системных часов соответствующего источника ссылок, в соответствие с которыми, помимо прочего, поставлена каждая из сохраненных потенциально вредоносных ссылок.

В одном из вариантов реализации настоящего изобретения вычислительное устройство 200 может быть выполнено с возможностью подключения непосредственно к каждому из источников 120, 130 ссылок с обеспечением получения прямого доступа соответственно к их базам 126, 136 ссылок для извлечения из них потенциально вредоносных ссылок для их последующей обработки вычислительным устройством 20, особенности которой описаны ниже.

Еще в одном варианте реализации настоящего изобретения парсеры 160, 170 могут быть оба соединены с возможностью обмена данными с вычислительным устройством 200 с использованием сети 110 Интернет.

В другом варианте реализации настоящего изобретения парсеры 160, 170 могут быть оба соединены проводным способом с возможностью обмена данными непосредственно с самим вычислительным устройством 200.

В некоторых вариантах реализации настоящего изобретения источники 120, 130 ссылок могут быть выполнены каждый с возможностью обмена данными соответственно с парсерами 160, 170 с использованием сети 110 Интернет, а сами парсеры 160, 170 могут быть соединены оба проводным способом непосредственно с вычислительным устройством 200.

Компьютерное устройство 140, которое может представлять собой, помимо прочего, стационарный компьютер, ноутбук, сервер и т.п., выполнено с возможностью обмена данными с вычислительным устройством 200 посредством парсера 180, при этом

компьютерное устройство 140 подключено к парсеру 180 проводным способом с возможностью передачи ему электронных сообщений, например по адресу электронной почты, привязанной к этому парсеру 180, при этом передаваемые электронные сообщения имеют заданный формат описания аналогично вышеописанным выходным потокам API-интерфейсов 124, 134. Парсер 180 предварительно запрограммирован или настроен на работу с компьютерным устройством 140, благодаря чему этот парсер 180 принимает электронные сообщения от компьютерного устройства 140, при этом парсеру 180 известен формат записи принимаемых электронных сообщений. Аналогично вышеописанному процессу работы парсера 160 или парсера 170, парсер 180 обрабатывает каждое принятое электронное сообщение и извлекает из его текста необходимые ссылки на веб-ресурсы (при этом такие ссылки имеют свой конкретный формат записи) и по меньшей мере некоторые из идентификационных данных, описывающих эти извлеченные ссылки, с обеспечением передачи, посредством сети 110 Интернет, извлеченных ссылок, поставленных в соответствие с извлеченными идентификационными данными, в вычислительное устройство 200 для их последующего анализа, особенности которого будут описаны ниже. Следует отметить, что для извлечения необходимых ссылок на веб-ресурсы из текста обрабатываемых электронных сообщений парсер 180 использует соответствующее регулярное выражение из заданного набора известных ему регулярных выражений. Парсер 180 может быть реализован в виде отдельного сервера или иного известного вычислительного устройства.

В одном из вариантов реализации настоящего изобретения компьютерное устройство 140 может быть выполнено с возможностью передачи сообщений в парсер 180 с использованием сети 110 Интернет, а парсер 180 может быть соединен проводным способом непосредственно с вычислительным устройством 200.

Мобильное устройство 150, которое может представлять собой, помимо прочего, смартфон, сотовый телефон, планшет и т.п., выполнено с возможностью обмена данными с вычислительным устройством 200 с использованием двух каналов связи. В частности, для обмена данными между мобильным устройством 150 и вычислительным устройством 200 по одному из этих каналов связи мобильное устройство 150 подключено проводным способом к парсеру 190 с возможностью передачи ему электронных сообщений, содержащих, помимо прочего, ссылки на потенциально вредоносные веб-ресурсы, по адресу электронной почты, привязанному к этому парсеру 190, при этом передаваемые электронные сообщения имеют заданный формат описания аналогично вышеописанным выходным потокам API-интерфейсов 124, 134. Парсер 190 предварительно запрограммирован или настроен на работу с мобильным устройством 150, благодаря чему этот парсер 190 принимает электронные сообщения от мобильного устройства 150, при этом парсеру 190 известен формат записи принимаемых электронных сообщений. Аналогично вышеописанному процессу работы парсера 160 или парсера 170, парсер 190 извлекает из каждого принятого электронного сообщения ссылки на веб-ресурсы (при этом такие ссылки имеют свой конкретный формат записи) и по меньшей мере некоторые из идентификационных данных, описывающих эти извлеченные ссылки, с обеспечением передачи, посредством сети 110 Интернет, извлеченных ссылок на веб-ресурсы, поставленных в соответствие с некоторыми извлеченными идентификационными данными, в вычислительное устройство 200 для их последующего анализа, особенности которого будут описаны ниже. Следует отметить, что парсер 190 может быть реализован в виде отдельного сервера или иного известного вычислительного устройства.

Кроме того, для обмена данными между мобильным устройством 150 и

вычислительным устройством 200 по другому каналу связи мобильное устройство 150 подключено, посредством сотовой сети 115, к парсеру 195 с возможностью передачи ему, например, SMS-сообщений и/или MMS-сообщений, содержащих, помимо прочего, ссылки на веб-ресурсы, по контактному номеру, привязанному к этому парсеру 195, при этом передаваемые SMS-сообщения и/или MMS-сообщения имеют заданный формат описания аналогично вышеописанным выходным потокам API-интерфейсов 124, 134. Парсер 195 предварительно запрограммирован или настроен на работу с мобильным устройством 150, благодаря чему этот парсер 195 принимает SMS-сообщения и/или MMS-сообщения от мобильного устройства 150, при этом парсеру 195 известен формат записи принимаемых SMS-сообщений и/или MMS-сообщений. Для приема SMS-сообщений и MMS-сообщений, передаваемых от мобильного устройства 150 к парсеру 195 посредством сотовой сети 115, парсер 195 подключен к внешнему модему, снабженному SIM-картой. Аналогично вышеописанному процессу работы парсера 160 или парсера 170, парсер 195 извлекает из каждого принятого SMS-сообщения или MMS-сообщения ссылки на веб-ресурсы (при этом такие ссылки имеют свой конкретный формат записи) и по меньшей мере некоторые из идентификационных данных, описывающих эти извлеченные ссылки, например контактный номер отправителя, с обеспечением передачи этих извлеченных ссылок, поставленных в соответствие с некоторыми извлеченными идентификационными данными, в вычислительное устройство 200, соединенное с парсером 195 проводным способом, для их последующего анализа, особенности которого будут описаны ниже. Следует отметить, что для извлечения необходимых ссылок на веб-ресурсы из текста принятых сообщений парсер 195 использует соответствующее регулярное выражение из заданного набора известных ему регулярных выражений. Парсер 195 может быть выполнен в виде отдельного сервера или иного известного вычислительного устройства.

В некоторых вариантах реализации настоящего изобретения преобразующий модуль, подключенный к парсеру 180, и преобразующий модуль, подключенный к парсеру 190, могут быть реализованы в виде одиночного преобразующего модуля, подключенного проводным и/или беспроводным способом с возможностью обмена данными к компьютерному устройству 140 и мобильному устройству 150, и имеющего функциональные возможности, аналогичные функциональным возможностям этих подключенных преобразующих модулей.

В одном из вариантов реализации настоящего изобретения вычислительное устройство 200 может быть выполнено с возможностью подключения непосредственно к каждому из компьютерного устройства 140 и мобильного устройства 150 с обеспечением получения прямого доступа к их внутренним базам данных, размещенных в памяти этих устройств, для получения из них сообщений, например SMS-сообщений, MMS-сообщений, электронных сообщений и т.п. (при этом на каждом из компьютерного устройства 140 и мобильного устройства 150 должна быть установлена, например, специальная клиентская программа). Вычислительное устройство 200 может обрабатывать каждое полученное сообщение для извлечения из него ссылок для их последующей обработки вычислительным устройством 200, особенности которой описаны ниже.

Еще в одном варианте реализации настоящего изобретения парсеры 180, 190, 195 могут быть соединены каждый с возможностью обмена данными с вычислительным устройством 200 с использованием сети 110 Интернет.

В другом варианте реализации настоящего изобретения парсеры 180, 190, 195 могут быть соединены каждый проводным способом с возможностью обмена данными

непосредственно с самим вычислительным устройством 200.

В некоторых вариантах реализации настоящего изобретения компьютерное устройство 140 и мобильное устройство 150 могут быть выполнены каждый с возможностью обмена данными соответственно с парсерами 180, 190 с использованием сети 110 Интернет, а сами парсеры 180, 190 могут быть соединены оба проводным способом непосредственно с вычислительным устройством 200.

В других вариантах реализации настоящего изобретения парсер 195 может быть соединен с возможностью обмена данными с вычислительным устройством 200 с использованием сети 110 Интернет.

Следует отметить, что источник 120 ссылок, источник 130 ссылок, компьютерное устройство 140 и мобильное устройство 150 показаны на фиг. 1 исключительно в качестве примера, то есть не следует считать, что возможная реализация системы 300 для информирования о вредоносных веб-ресурсах ограничена примером, показанным на фиг. 1, при этом специалисту в данной области техники должно быть ясно, что система 300 может содержать два и более источников ссылок, подобных каждый вышеописанному источнику 120 ссылок, два и более источников ссылок, подобных каждый вышеописанному источнику 130 ссылок, два или более компьютерных устройств, подобных каждый вышеописанному компьютерному устройству 140, и/или два или более мобильных устройств, подобных каждый вышеописанному мобильному устройству 150.

В одном из вариантов реализации настоящего изобретения каждый из источников ссылок, подобных каждый вышеописанному источнику 120 ссылок, может быть подключен к вычислительному устройству 200 посредством отдельного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 160, при каждый такой отдельный парсер будет предварительно запрограммирован или настроен на работу с соответствующим источником ссылок для понимания синтаксиса команд API-интерфейса этого источника ссылок.

Еще в одном варианте реализации настоящего изобретения все источники ссылок в системе 300, подобные каждый вышеописанному источнику 120 ссылок, могут быть подключены к вычислительному устройству 200 посредством одного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 160, при этом такой общий парсер должен быть предварительно запрограммирован или настроен на работу с каждым из этих подключенных источников ссылок для понимания синтаксиса команд его API-интерфейса.

В некоторых вариантах реализации настоящего изобретения каждый из источников ссылок, подобных каждый вышеописанному источнику 130 ссылок, может быть подключен к вычислительному устройству 200 посредством отдельного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 170, при этом каждый такой отдельный парсер будет предварительно запрограммирован или настроен на работу с соответствующим источником ссылок для понимания синтаксиса команд API-интерфейса этого источника ссылок.

В других вариантах реализации настоящего изобретения все источники ссылок в системе 300, подобные каждый вышеописанному источнику 130 ссылок, могут быть подключены к вычислительному устройству 200 посредством одного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 170, при этом такой общий парсер должен быть предварительно запрограммирован или настроен на работу с каждым из этих подключенных источников ссылок для понимания синтаксиса команд его API-интерфейса.

В иных вариантах реализации настоящего изобретения каждое из компьютерных устройств, подобных каждое вышеописанному компьютерному устройству 140, может быть подключено к вычислительному устройству 200 посредством отдельного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 180, при каждый такой отдельный парсер будет предварительно запрограммирован или настроен на работу с соответствующим компьютерным устройством для понимания формата записи электронных сообщений, принимаемых от этого компьютерного устройства.

Еще в одних вариантах реализации настоящего изобретения все компьютерные устройства в системе 300, подобные каждое вышеописанному компьютерному устройству 140, могут быть подключены к вычислительному устройству 200 посредством одного парсера с функциональными возможностями, аналогичными вышеописанному парсеру 170, при этом такой общий парсер должен быть предварительно запрограммирован или настроен на работу с каждым из этих подключенных компьютерных устройств для понимания формата записи электронных сообщений, принимаемых от этого компьютерного устройства.

В некотором варианте реализации настоящего изобретения каждое из мобильных устройств, подобных каждое вышеописанному мобильному устройству 150, может быть подключено к вычислительному устройству 200 посредством отдельного парсера с функциональными возможностями, аналогичными вышеописанным парсерам 190, 195, при каждый такой отдельный парсер будет предварительно запрограммирован или настроен на работу с соответствующим мобильным устройством для понимания формата записи сообщений, принимаемых от этого мобильного устройства, в частности электронных сообщений, SMS-сообщений и/или MMS-сообщений.

В некотором другом варианте реализации настоящего изобретения все мобильные устройства в системе 300, подобные каждое вышеописанному мобильному устройству 150, могут быть подключены к вычислительному устройству 200 посредством одного парсера с функциональными возможностями, аналогичными вышеописанным парсерам 190, 195, при этом такой общий парсер должен быть предварительно запрограммирован или настроен на работу с каждым из этих подключенных мобильных устройств для понимания формата записи сообщений, принимаемых от этого мобильного устройства, в частности электронных сообщений, SMS-сообщений и/или MMS-сообщений.

Согласно одному из вариантов реализации настоящего изобретения по меньшей часть из источников ссылок, подобных каждый источнику 120 ссылок, источников ссылок, подобных каждый источнику 130 ссылок, компьютерных устройств, подобных каждое компьютерному устройству 140, и мобильных устройств, подобных каждое мобильному устройству 150, могут быть подключены к вычислительному устройству 200 посредством одного парсера с функциональными возможностями, аналогичными вышеописанным парсерам 160, 170, 180, 190 и 195, при этом такой общий парсер должен быть предварительно запрограммирован или настроен на работу с каждым из подключенных источников ссылок для понимания синтаксиса команд его API-интерфейса, каждым из подключенных компьютерных устройств для понимания формата записи электронных сообщений, принимаемых от этого компьютерного устройства, и каждым из подключенных мобильных устройств для понимания формата записи сообщений вышеописанных типов, принимаемых от этого мобильного устройства.

Согласно еще одному варианту реализации настоящего изобретения вычислительное устройство может быть подписано на RSS-рассылку по меньшей мере одного из

источников ссылок, подобных каждый вышеописанному источнику 120 ссылок, и/или RSS-рассылку по меньшей мере одного из источников ссылок, подобных каждый источнику 130 ссылок, для получения от указанных источников ссылок по меньшей мере одного отчета, указывающего, например, на появление в соответствующем источнике ссылок по меньшей мере одной новой ссылки на веб-ресурс.

Согласно некоторым вариантам реализации настоящего изобретения система 300 может дополнительно содержать отдельную базу ссылок, являющуюся внешней или удаленной по отношению к вычислительному устройству 200, при этом парсеры 160, 170, 180, 190 и 195 могут быть выполнены каждый с возможностью получения доступа к этой внешней базе ссылок с обеспечением возможности записи в неё ссылок, извлеченных надлежащим образом в соответствии с приведенным выше описанием, в результате чего эта внешняя база ссылок содержит множество ссылок на потенциально вредоносные веб-ресурсы, поставленные каждая в соответствие со вспомогательными идентификационными данными, описывающими эту ссылку, например датой и временем сохранения и/или по меньшей мере одним иным идентификатором. Вычислительное устройство 200 выполнено с возможностью получения доступа к такой базе ссылок с обеспечением возможности извлечения из неё необходимых ссылок для их последующей обработки, особенности которой описаны ниже. В качестве дополнения или альтернативы в данном варианте реализации вышеописанная внешняя база ссылок может также содержать множество ссылок на известные вредоносные веб-ресурсы.

Согласно другим вариантам реализации настоящего изобретения система 300 может содержать только вычислительное устройство 200 и структурированную базу ссылок, являющуюся внешней или удаленной по отношению к вычислительному устройству 200. В данном варианте реализации внешняя база ссылок содержит ссылки на потенциально вредоносные веб-ресурсы, сохраненные из множества различных источников, при этом каждая ссылка в этой внешней базе ссылок поставлена в соответствие со вспомогательными идентификационными данными, описывающими эту ссылку, например, датой и временем сохранения и/или по меньшей мере одним иным идентификатором. Вычислительное устройство 200 выполнено с возможностью получения доступа к такой внешней базе ссылок с обеспечением возможности извлечения из неё необходимых ссылок для их последующей обработки, особенности которой описаны ниже. В качестве дополнения или альтернативы в данном варианте реализации вышеописанная внешняя база ссылок может также содержать множество ссылок на известные вредоносные веб-ресурсы.

Вычислительное устройство для информирования о вредоносных веб-ресурсах

Вычислительное устройство 200, показанное на фиг. 2, предназначено для информирования уполномоченных субъектов о выявленных вредоносных веб-ресурсах и по существу представляет собой программно-аппаратный комплекс, реализованный в виде компьютера общего назначения, имеющего описанную ниже структуру, хорошо известную специалистам в данной области техники.

Следует отметить, что в данном документе под уполномоченным субъектом следует понимать физическое лицо, способное заблокировать работу веб-ресурса или повлиять на решение о блокировке вредоносного веб-ресурса или приостановлении его функционирования, например администратор веб-ресурса, владелец веб-ресурса и т.п., или юридическое лицо, способное заблокировать работу веб-ресурса или повлиять на решение о блокировке вредоносного веб-ресурса или приостановлении его функционирования, например регистратор доменных имен, хостинг-провайдер и т.п.

В частности, компьютер общего назначения обычно содержит центральный

процессор, системную память и системную шину, которая в свою очередь содержит разные системные компоненты, в том числе память, связанную с центральным процессором. Системная шина в таком компьютере общего назначения содержит шину памяти и контроллер шины памяти, периферийную шину и локальную шину, выполненную с возможностью взаимодействия с любой другой шинной архитектурой. Системная память содержит постоянное запоминающее устройство (ПЗУ) и память с произвольным доступом (ОЗУ). Основная система ввода/вывода (BIOS) содержит основные процедуры, которые обеспечивают передачу информации между элементами такого компьютера общего назначения, например в момент загрузки операционной системы с использованием ПЗУ. Кроме того, компьютер общего назначения содержит жесткий диск для чтения и записи данных, привод магнитных дисков для чтения и записи на сменные магнитные диски и оптический привод для чтения и записи на сменные оптические диски, такие как CD-ROM, DVD-ROM и иные оптические носители информации, однако могут быть использованы компьютерные носители иных типов, выполненные с возможностью хранения данных в машиночитаемой форме, например твердотельные накопители, флеш-карты, цифровые диски и т.п., и подключенные к системной шине через контроллер. В компьютере общего назначения жесткий диск, привод магнитных дисков и оптический привод соединены соответственно с системной шиной через интерфейс жесткого диска, интерфейс магнитных дисков и интерфейс оптического привода. Приводы и соответствующие компьютерные носители информации представляют собой энергонезависимые средства хранения компьютерных инструкций, структур данных, программных модулей и прочих данных компьютера общего назначения. Компьютер общего назначения имеет файловую систему, в которой хранится записанная операционная система, а также дополнительные программные приложения, прочие программные модули и данные программ. Пользователь имеет возможность вводить команды и информацию в компьютер общего назначения с использованием известных устройств ввода, например клавиатуры, манипулятора типа «мышь», микрофона, джойстика, игровой консоли, сканера и т.п., при этом эти устройства ввода обычно подключают доступ к компьютеру общего назначения через последовательный порт, который в свою очередь подсоединен к системной шине, однако они могут быть подключены и иным способом, например с помощью параллельного порта, игрового порта или универсальной последовательной шины (USB). Монитор или иной тип устройства отображения также подсоединен к системной шине через интерфейс, такой как видеоадаптер. В дополнение к монитору персональный компьютер может быть снабжен другими периферийными устройствами вывода, например колонками, принтером и т.п. Компьютер общего назначения способен работать в сетевом окружении, при этом для соединения с одним или несколькими удаленными компьютерами может быть использовано сетевое соединение. Сетевые соединения могут образовывать локальную вычислительную сеть (LAN) и глобальную вычислительную сеть (WAN). Такие сети обычно применяют в корпоративных компьютерных сетях и внутренних сетях компаний, при этом они имеют доступ к сети Интернет. В LAN-сетях или WAN-сетях компьютер общего назначения подключают к локальной сети через сетевой адаптер или сетевой интерфейс. При использовании сетей компьютер общего назначения может использовать модем, сетевую карту, адаптер или иные средства обеспечения связи с глобальной вычислительной сетью, такой как сеть Интернет, при этом эти средства связи подключают к системной шине посредством последовательного порта. Следует отметить, что в ПЗУ компьютера общего назначения или по меньшей мере на любом одном из вышеописанных машиночитаемых носителей,

которые могут быть использованы в компьютере общего назначения, могут быть сохранены машиночитаемые инструкции, к которым может иметь доступ центральный процессор этого компьютера общего назначения, при этом выполнение этих машиночитаемых инструкций на компьютере общего назначения может вызывать исполнение его центральным процессором различных процедур или операций, описанных ниже в данном документе.

В одном из вариантов реализации настоящего изобретения вычислительное устройство 200 может быть выполнено в виде одиночного компьютерного сервера, например сервера «Dell™ PowerEdge™», использующего операционную систему «Ubuntu Server 18.04». Кроме того, в иных вариантах реализации настоящего изобретения вычислительное устройство 200 может быть выполнено в виде настольного персонального компьютера, ноутбука, нетбука, смартфона, планшета и иного электронно-вычислительного устройства, подходящего для решения поставленных задач.

В других вариантах реализации вычислительное устройство 200 может быть выполнено в виде любой другой совокупности аппаратных средств, программного обеспечения или программно-аппаратного комплекса, подходящих для решения поставленных задач.

В некоторых вариантах реализации настоящего изобретения система 300 может содержать по меньшей мере два вычислительных устройства, подобных каждое вычислительному устройству 200, при этом нижеописанные функциональные возможности вычислительного устройства 200 могут быть любым необходимым образом разделены между указанными по меньшей мере двумя вычислительными устройствами, каждый из которых, например, может быть выполнен в виде отдельного компьютерного сервера.

Вычислительное устройство 200, показанное на фиг. 2, содержит модуль 10 связи, анализирующий модуль 100 и локальное хранилище 20 данных, каждый из которых соединен с шиной 30 связи, при этом каждый из модуля 10 связи и анализирующего модуля 100 выполнен с возможностью обмена данными, посредством шины 30 связи, с локальным хранилищем 20 данных, а модуль 10 связи также выполнен с возможностью обмена данными с анализирующим модулем 100.

В одном из вариантов реализации настоящего изобретения вышеописанные парсеры 160, 170, 180, 190 и 195 могут быть выполнены каждый в виде обособленного модуля предварительной обработки данных, встроенного в вычислительное устройство 200 (т.е. входящего в состав этого вычислительного устройства 200) и имеющего вышеописанные функциональные возможности соответствующего одного из парсеров 160, 170, 180, 190 и 195, в частности функциональные возможности по обеспечению взаимодействия или обмена данными между вычислительным устройством 200 и соответствующим одним из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150 (т.е. каждый из таких обособленных модулей предварительной обработки данных должен быть предварительно запрограммирован на работу с соответствующим одним из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150) и по обработке входных потоков данных, поступающих от соответствующего одного из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. В одной из разновидностей этого варианта реализации модуль 10 связи вычислительного устройства 200 может быть выполнен многоканальным, например четырехканальным, при этом каждый из

каналов связи в таком модуле 10 связи может быть предварительно настроен на обмен данными, посредством шины 30 связи, с одним из вышеописанных обособленных модулей предварительной обработки данных и на обмен данными с соответствующим одним из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. В другой разновидности данного варианта реализации вычислительное устройство 200 может быть снабжено четырьмя модулями связи, подобными каждый модулю 10 связи, при этом каждый из таких модулей связи предварительно настроен на обмен данными, посредством шины 30 связи, с одним из вышеописанных обособленных модулей предварительной обработки данных и на обмен данными с соответствующим одним из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. В данном варианте реализации обособленные модули предварительной обработки данных (не показаны) также выполнены каждый с возможностью взаимодействия, посредством шины 30 связи, с анализирующим модулем 100 для обработки запросов на получение ссылок, которые могут быть сформированы этим анализирующим модулем 100, с последующим их направлением от вычислительного устройства 200 в соответствующий один из вышеописанных источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. Следует также отметить, что при обработке входных потоков данных, полученных от соответствующего одного из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150, каждый из таких обособленных модулей предварительной обработки данных (не показаны) может, помимо прочего, идентифицировать или распознавать формат описания полученного входного потока данных. Если идентифицированный формат описания данных не соответствует унифицированному формату описания данных, подходящему для вычислительного устройства 200, то каждый из обособленных модулей предварительной обработки данных может быть дополнительно выполнен с возможностью преобразования этого полученного входного потока данных в указанный унифицированный формат, при этом он может быть дополнительно выполнен с возможностью связи, посредством шины 30 связи, с локальным хранилищем 20 данных с обеспечением получения из него данных об унифицированном формате описания данных (как описано ниже), понятном вычислительному устройству 200, и с возможностью сравнения указанных идентифицированного и унифицированного форматов данных для принятия решения об их соответствии или несоответствии друг другу. Таким образом, если любой из вышеописанных обособленных модулей предварительной обработки данных выявит, что среди входных потоков данных, полученных им от соответствующего одного из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150, имеются, например, голосовые сообщения или видео-сообщения, то такой обособленный модуль предварительной обработки данных преобразует такие сообщения в текст, то есть в тот формат описания данных, который понятен вычислительному устройству 200, с последующим извлечением из него ссылок на потенциально вредоносные веб-ресурсы.

Еще в одном варианте реализации настоящего изобретения вышеописанные парсеры 150, 160, 170, 180, 190 и 195 могут быть выполнены в виде одиночного модуля предварительной обработки данных (не показан), встроенного в вычислительное устройство 200 (т.е. входящего в состав этого вычислительного устройства 200) и имеющего вышеописанные функциональные возможности всех парсеров 150, 160, 170, 180, 190 и 195, в частности функциональные возможности по обеспечению взаимодействия или обмена данными между вычислительным устройством 200 и каждым

из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150 (т.е. такой одиночный модуль предварительной обработки данных должен быть предварительно запрограммирован на работу с каждым из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150) и по обработке входных потоков данных, поступающих от каждого из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. В этом варианте реализации одиночный модуль предварительной обработки данных (не показан) также должен быть подключен в вычислительном устройстве 200 к шине 30 связи с обеспечением возможности обмена данными с модулем 10 связи, обеспечивающим взаимодействие между вычислительным устройством 200 и источником 120 ссылок, источником 130 ссылок, компьютерным устройством 140 и мобильным устройством 150, при этом модуль 10 связи вычислительного устройства 200 в таком случае может быть выполнен, например, многоканальным, а каждый из каналов связи в таком модуле 10 связи может быть предварительно настроен на обмен данными с соответствующим одним из источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150. В данном варианте реализации одиночный модуль предварительной обработки данных (не показан) также выполнен с возможностью взаимодействия, посредством шины 30 связи, с анализирующим модулем 100 для обработки запросов на получение ссылок, которые могут быть сформированы этим анализирующим модулем 100, с их последующим направлением от вычислительного устройства 200 в вышеописанные источник 120 ссылок, источник 130 ссылок, компьютерное устройство 140 и мобильное устройство 150. Следует также отметить, что при обработке входных потоков данных, полученных от источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150, одиночный модуль предварительной обработки данных (не показан) может, помимо прочего, идентифицировать или распознавать формат описания этих входных потоков данных, при этом если идентифицированный формат описания данных не соответствует унифицированному формату описания данных, подходящему для вычислительного устройства 200, то он может дополнительно выполнен с возможностью преобразования этих полученных входных потоков данных в указанный унифицированный формат, при этом этот одиночный модуль предварительной обработки данных может быть дополнительно выполнен с возможностью связи, посредством шины 30 связи, с локальным хранилищем 20 данных с обеспечением получения из него данных об унифицированном формате описания данных (как описано ниже), понятном вычислительному устройству 200, и с возможностью сравнения указанных идентифицированного и унифицированного форматов данных для принятия решения об их соответствии или несоответствии друг другу. Таким образом, если вышеописанный одиночный модуль предварительной обработки данных выявит, что среди входных потоков данных, полученных им от источника 120 ссылок, источника 130 ссылок, компьютерного устройства 140 и мобильного устройства 150, имеются, например, голосовые сообщения или видео-сообщения, то такой одиночный модуль предварительной обработки данных преобразует такие сообщения в текст, то есть в тот формат описания данных, который понятен вычислительному устройству 200, с последующим извлечением из него ссылок на потенциально вредоносные веб-ресурсы.

В некоторых вариантах реализации настоящего изобретения функциональные возможности вышеописанных парсеров 160, 170, 180, 190, 195 могут быть реализованы в качестве дополнительных функциональных возможностей анализирующего модуля

100, в частности каждый из парсеров 160, 170, 180, 190, 195 или все эти парсеры могут быть реализованы в виде отдельного программного модуля, встроенного в вычислительное устройство 200 и исполняемого анализирующим модулем 100.

В одном из вариантов реализации настоящего изобретения вычислительное устройство 200 может дополнительно содержать вспомогательный модуль сбора контекстной рекламы (не показан), выполненный с возможностью автоматического сбора контекстной рекламы, демонстрируемой или показываемой пользователям в известных поисковых системах, таких как, например, Bing, Google, Yandex и т.п., с обеспечением извлечения из контекстной рекламы, собранной по меньшей мере в одной из этих известных поисковых систем, по меньшей мере одной ссылки на веб-ресурс. Модуль сбора контекстной рекламы подключен к шине 30 связи и выполнен с возможностью обмена данными, посредством шины 30 связи, с модулем 10 связи, локальным хранилищем 20 данных и анализирующим модулем 100. Следует отметить, что в последнее время злоумышленники часто прибегают к распространению ссылок на вредоносные веб-ресурсы путем размещения этих ссылок в контекстной рекламе известных поисковых систем, при этом такая вредоносная реклама обычно таргетирована на наиболее частотные поисковые запросы пользователей в каждой из этих поисковых систем, поскольку такие списки наиболее популярных у пользователей ключевых слов находятся в свободном доступе на сайтах этих поисковых систем. В данном варианте реализации настоящего изобретения локальное хранилище 20 данных дополнительно содержит отдельную базу ключевых слов поисковых запросов, содержащую несколько разделов, в каждом из которых сохраняют ключевые слова наиболее частотных поисковых запросов соответствующей одной из известных поисковых систем, на работу с которыми предварительно настроен или запрограммирован модуль сбора контекстной рекламы, так что все ключевые слова в каждом конкретном разделе этой базы поставлены в соответствие с одной из известных поисковых систем. Модуль сбора контекстной рекламы также выполнен с возможностью по меньшей мере периодического обновления (например, ежедневно) базы ключевых слов поисковых запросов, размещенной в локальном хранилище 20 данных, по меньшей мере для одной из известных ему поисковых систем, например путем периодического автоматического получения актуального перечня ключевых слов, являющихся наиболее популярными у пользователей в конкретной поисковой системе, с использованием конкретной ссылки на веб-страницу сайта этой поисковой системы, сохраненной в локальном хранилище 20 данных и извлекаемой оттуда указанным модулем сбора контекстной рекламы при обновлении конкретного раздела базы ключевых слов поисковых запросов, соответствующего указанной поисковой системе, с последующей актуализацией имеющегося перечня ключевых слов поисковых запросов в разделе базы ключевых слов поисковых запросов, соответствующем указанной поисковой системе, на основании полученного актуального перечня ключевых слов. Модуль сбора контекстной рекламы также выполнен с возможностью формирования по меньшей мере одного поискового запроса по меньшей мере для одной из известных ему поисковых систем с использованием по меньшей мере части ключевых слов, содержащихся в одном из разделов базу ключевых слов поисковых запросов, соответствующем этой поисковой системе, и возможностью автоматической передачи этого сформированного поискового запроса в эту поисковую систему. Модуль сбора контекстной рекламы также выполнен с возможностью получения результатов поиска, выданных поисковой системой в ответ на переданный запрос, и возможностью фильтрации полученных результатов поиска для выявления среди них контекстной

рекламы в виде рекламных объявлений, на основании, например, метки «реклама», которой снабжены такие рекламные объявления, при этом каждое такое рекламное объявление содержит, помимо прочего, по меньшей мере одну ссылку на веб-ресурсы. Модуль сбора контекстной рекламы дополнительно выполнен с возможностью извлечения, посредством, например, известного ему регулярного выражения, такого как, например, (https?|ftp)://(-\.)?([^\s/?.#-]+\.?)+(/[^\s]*)?$@iS, по меньшей мере одной ссылки на веб-ресурсы из каждого выявленного рекламного объявления с обеспечением передачи, посредством шины 30 связи, каждой этой ссылки на веб-ресурс в анализирующий модуль 100 для ее последующего анализа на вредоносность для выявления или установления того, относится ли веб-ресурс, находящийся по этой ссылке, к вредоносным веб-ресурсам, как более подробно описано ниже в данном документе. Таким образом, модуль сбора контекстной рекламы может, например, последовательно формировать поисковые запросы для каждой конкретной поисковой системы с использованием некоторой комбинации ключевых слов, сформированной по меньшей мере из части ключевых слов в имеющемся перечне ключевых слов, соответствующем этой поисковой системе, до тех пор, пока не будет достигнут конец этого перечня ключевых слов. Следует отметить, что вышеописанный способ получения ссылок на веб-ресурсы вычислительным устройством 200 может являться альтернативой или дополнением к вышеописанным способам получения ссылок на веб-ресурсы, используемым в системе 300. В описанном варианте реализации вспомогательный модуль сбора контекстной рекламы может быть реализован, например, в виде отдельного процессора, встроенного в вычислительное устройство 200.

В одном из вариантов реализации настоящего изобретения функциональные возможности вышеописанного модуля сбора контекстной рекламы могут реализованы в виде дополнительных функциональных возможностей анализирующего модуля 100, в частности модуль сбора контекстной рекламы может быть реализован в виде отдельного программного модуля, входящего в состав вычислительного устройства 200 и исполняемого, например, анализирующим модулем 100. В другом варианте реализации насточщего изобретения модуль сбора контекстной рекламы может представлять собой один из функциональных подмодулей анализирующего модуля 100.

Еще в одном варианте реализации настоящего изобретения вышеописанный модуль сбора контекстной рекламы может представлять собой обособленный источник ссылок, например, отдельный сервер, являющийся внешним по отношению к вычислительному устройству 200 и подключенный к нему проводным и/или беспроводным способом с обеспечением возможности передачи ему ссылок на веб-ресурсы, при этом ссылки на веб-ресурсы, передаваемые от такого внешнего источника ссылок, могут быть приняты модулем 10 связи вычислительного устройства 200.

Локальное хранилище данных

Локальное хранилище 20 данных также предназначено для хранения исполняемых программных инструкций, которые позволяют управлять работой функциональных модулей, встроенных в вычислительное устройство 200, в частности модуля 10 связи и анализирующего модуля 100, и позволяют этим функциональным модулям реализовывать свои функциональные возможности при исполнении этих программных инструкций. Исполняемые программные инструкции, хранящиеся в локальном хранилище 20 данных, также позволяют управлять работой любых подмодулей, которые в некоторых раскрытых вариантах реализации входят в состав некоторых из функциональных модулей, например анализирующего модуля 100, и позволяют этим

подмодулям реализовывать свои функциональные возможности при исполнении этих программных инструкций.

Локальное хранилище 20 данных также может хранить исполняемые программные инструкции, которые позволяют управлять работой любых дополнительных функциональных модулей, встроенных в вычислительное устройство 200, и их подмодулей, и которые позволяют этим дополнительным функциональным модулям и их подмодулям реализовывать свои функциональные возможности при исполнении этих программных инструкций.

Кроме того, локальное хранилище 20 данных предназначено для хранения различных данных, используемых при работе вычислительного устройства 200, в частности данных об унифицированном формате описания данных, понятном вычислительному устройству 200, данных об известных вредоносных ссылках, данных об известных вредоносных доменных именах, данных о хеш-суммах известных вредоносных файлов, данных о ключевых словах, указывающих на вредоносный характер веб-ресурса, данных о хостинг провайдере, данных о регистраторе доменных имен, перечня известных уполномоченных субъектов, набора известных типов угроз вредоносных веб-ресурсов, набора шаблонов отчетов и т.п. Локальное хранилище 20 данных также может хранить и иные данные, используемые при работе различных функциональных модулей, встроенных в вычислительное устройство 200, и работе по меньшей мере некоторых подмодулей, входящих в состав некоторых из этих функциональных модулей.

Кроме того, в локальном хранилище 20 данных также могут быть сохранены вспомогательные данные, используемые в работе анализирующего модуля 100, например данные о языковых словарях и заданное пороговое значение, используемые в методике анализа доменных имен на основании правильности их написания; файлы образов виртуальных машин и набор правил анализа изменений параметров состояния виртуальной машины, используемые в методике анализа файлов на подозрительность на основании изменения параметров состояния виртуальных машин, набор регулярных выражений, используемых для извлечения ссылок на веб-ресурсы из входных потоков данных, анализируемых в анализирующем модуле 100, и иные вспомогательных данные.

В вычислительном устройстве 200, показанном на фиг. 2, модуль 10 связи выполнен с возможностью приема извлеченных ссылок на веб-ресурсы, передаваемых парсерами 160, 170, 180, 190 и 195 в вычислительное устройство 200, с последующем сохранением принятых ссылок на веб-ресурсы в локальном хранилище 20 данных, в которое эти принятые данные могут быть переданы посредством шины 30 связи. Таким образом, в локальном хранилище 20 данных могут быть сохранены ссылки на веб-ресурсы, извлеченные из потоков данных от источника 120 ссылок, ссылки на веб-ресурсы, извлеченные из потоков данных от источника 130 ссылок, ссылки на веб-ресурсы, извлеченные из сообщений от компьютерного устройства 140, и/или ссылки на веб-ресурсы, извлеченные из сообщений от мобильного устройства 150, и по меньшей мере некоторые из извлеченных идентификационных данных, описывающих такие сохраненные ссылки.

В некоторых вариантах реализации настоящего изобретения локальное хранилище 20 данных в вычислительном устройстве 200 может содержать одну или несколько баз данных, выполненных каждая с возможностью сохранения в ней по меньшей мере одной обособленной группы из вышеперечисленных групп данных, используемых в работе вычислительного устройства 200, и/или по меньшей мере некоторых из принятых ссылок на веб-ресурсы.

В других вариантах реализации вычислительное устройство 200 может использовать

по меньшей мере одно обособленное удаленное хранилище данных (не показано), к которому анализирующий модуль 100 вычислительного устройства 200 может получать доступ с использованием модуля 10 связи, для сохранения в нем по меньшей мере части из вышеописанных групп данных и/или по меньшей мере части из принятых ссылок на веб-ресурсы.

В некоторых других вариантах реализации настоящего изобретения вычислительное устройство 200 может содержать по меньшей мере одно локальное хранилище данных и по меньшей мере одно удаленное хранилище данных (не показано), предназначенные каждое для хранения по меньшей мере одной из вышеописанных групп данных и/или по меньшей мере части из принятых ссылок на веб-ресурсы, при этом указанные локальные хранилища данных соединены каждое с анализирующим модулем 100 посредством шины 30 связи, а указанные удаленные хранилища данных соединены каждое с анализирующим модулем 100 посредством модуля 10 связи. Таким образом, например, возможен вариант реализации настоящего изобретения, в котором вычислительное устройство 200 содержит единственное локальное хранилище 20 данных, хранящее, например, исключительно принятые ссылки на веб-ресурсы, и содержит несколько удаленных хранилищ данных, хранящих каждое по меньшей мере некоторые из вышеописанных групп данных, используемых при работе вычислительного устройства 200.

В одном из вариантов реализации настоящего изобретения по меньшей мере одна из вышеперечисленных групп данных и/или принятые ссылки на веб-ресурсы могут быть сохранены в соответствующем обособленном локальном хранилище данных (не показано), отличном от локального хранилища 20 данных и соединенном, посредством шины 30 связи, с анализирующим модулем 100, который в свою очередь выполнен с возможностью подключения к любому из таких обособленных локальных хранилищ данных с обеспечением извлечения из них необходимых ссылок на веб-ресурсы.

Анализирующий модуль 100 может быть реализован в виде одиночного процессора, такого как процессор общего назначения или процессор специального назначения (например, процессоры для цифровой обработки сигналов, специализированные интегральные схемы и т.п.), и выполнен с возможностью исполнения программных инструкций, хранящихся в локальном хранилище 20 данных, с обеспечением реализации нижеописанных функциональных возможностей анализирующего модуля 100.

Локальное хранилище 20 данных может быть реализовано, например, в виде одного или более известных физических машиночитаемых носителей для длительного хранения данных. В некоторых вариантах реализации настоящего изобретения локальное хранилище 20 данных может быть реализовано с использованием одиночного физического устройства (например, одного оптического запоминающего устройства, магнитного запоминающего устройства, органического запоминающего устройства, запоминающего устройства на дисках или запоминающего устройства иного типа), а в других вариантах реализации локальное хранилище 20 данных может быть реализовано с использованием двух или более известных запоминающих устройств.

Модуль связи

Модуль 10 связи, используемый в вычислительном устройстве 200, показанном на фиг. 1 и 2, имеет беспроводное соединение с вышеописанными парсерами 160, 180, 190 с возможностью обмена с ними данными, а также имеет проводное соединение с вышеописанными парсерами 170, 195 с возможностью обмена с ними данными.

В одном из вариантов реализации настоящего изобретения модуль 10 связи может быть соединен со всеми парсерами 160, 170, 180, 190, 195 проводным способом с

возможностью обмена с ними данными, например с помощью коаксиального кабеля, витой пары, оптоволоконного кабеля или другого физического соединения. В этом варианте реализации модуль 10 связи может быть реализован, например, в виде сетевого адаптера, снабженного необходимыми разъемами для подключения к ним физических кабелей необходимых типов в зависимости от типов физических соединений, использованных для обеспечения связи с парсерами 160, 170, 180, 190, 195.

Еще в одном варианте реализации настоящего изобретения модуль 10 связи может быть соединен со всеми парсерами 160, 170, 180, 190, 195 беспроводным способом с возможностью обмена с ними данными, например с помощью линии связи на основе технологии «WiFi», линии связи на основе технологии «3G», линии связи на основе технологии «LTE» и/или т.п. В этом варианте реализации модуль 10 связи может быть реализован, например, как сетевой адаптер в виде WiFi-адаптера, 3G-адаптера, LTE-адаптера или иного адаптера беспроводной связи в зависимости от типа линии беспроводной связи, использованной для обеспечения связи с парсерами 160, 170, 180, 190, 195.

В других вариантах реализации настоящего изобретения модуль 10 связи может использовать любую подходящую комбинацию из проводных и беспроводных линий связи для обмена данными по меньшей мере с частью из парсеров 160, 170, 180, 190, 195, входящих в состав системы 300.

Модуль 10 связи также может представлять собой известное устройство связи, такое как передатчик, приемник, приемопередатчик, модем и/или сетевую интерфейсную карту для обмена данными с внешними устройствами любого типа посредством проводной или беспроводной сети связи, например с помощью сетевого соединения стандарта «Ethernet», цифровой абонентской линия связи (DSL), телефонной линии, коаксиального кабеля, телефонной системы сотовой связи и т.п.

В некоторых вариантах реализации вычислительное устройство 200 может быть дополнительно снабжено модемом с SIM-картой для приема SMS-сообщений и/или MMS-сообщений от мобильных устройств, таких как мобильное устройство 150.

Анализирующий модуль

Анализирующий модуль 100, входящий в состав вычислительного устройства 200, показанного на фиг. 2, может быть реализован в виде одиночного процессора, такого как процессор общего назначения или процессор специального назначения (например, процессоры для цифровой обработки сигналов, специализированные интегральные схемы и т.п.), например в виде центрального процессора вышеописанного компьютера общего назначения, в виде которого может быть реализовано вычислительное устройство 200.

Анализирующий модуль 100 выполнен с возможностью получения доступа к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или возможностью связи с ним с использованием шины 30 связи с обеспечением извлечения из него ссылок на веб-ресурсы для их последующего анализа, как будет описано ниже.

В одном из вариантов реализации настоящего изобретения анализирующий модуль 100 может быть выполнен с возможностью связи, посредством шины 30 связи, с модулем 10 связи с обеспечением возможности получения от него ссылок на веб-ресурсы для их последующего анализа, как будет более подробно описано ниже. Таким образом, в этом варианте реализации анализирующий модуль 100 может получать ссылки на веб-ресурсы непосредственно от модуля 10 связи непосредственно после получения этих

ссылок модулем 10 связи.

В вариантах реализации настоящего изобретения, в которых полученные ссылки на веб-ресурсы хранятся в обособленном локальном хранилище, отличном от локального хранилища 20 данных, или в удаленном хранилище данных, анализирующий модуль 100 может быть выполнен с возможностью получения доступа к такому обособленному или удаленному хранилищу данных или возможностью связи с ним с использованием шины 30 связи с обеспечением извлечения из него сохраненных ссылок на веб-ресурсы для их последующего анализа, как будет более подробно описано ниже.

Анализирующий модуль 100 выполнен с возможностью анализа каждой из полученных или извлеченных ссылок на веб-ресурсы для выявления или установления веб-ресурсов с вредоносным и/или незаконным контентом, также называемых вредоносными веб-ресурсами, среди веб-ресурсов, находящиеся по анализируемым ссылкам, как будет более подробно описано ниже.

В частности, для выявления вредоносных веб-ресурсов при анализе ссылок на веб-ресурсы анализирующий модуль 100 (i) получает доступ к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или устанавливает с ним связь с использованием шины 30 связи с обеспечением получения из него данных об известных вредоносных ссылках; и (ii) устанавливает, путем посимвольного сравнения каждой анализируемой ссылки с известными вредоносными ссылками из указанных полученных данных, факт по меньшей мере частичного совпадения анализируемой ссылки по меньшей мере с одной из известных вредоносных ссылок.

Таким образом, если анализирующий модуль 100 установил или выявил, что конкретная ссылка имеет по меньшей мере частичное совпадение по меньшей мере с одной из известных вредоносных ссылок, то это свидетельствует о том, что эта ссылка относится к вредоносным ссылкам и, соответственно, веб-ресурс, находящейся по этой ссылке, относится к вредоносным веб-ресурсам.

Если же анализирующий модуль 100 установил или выявил, что анализируемая ссылка не имеет по меньшей мере частичного совпадения ни с одной из известных вредоносных ссылок, то он дополнительно выполняет по меньшей мере одну из следующих операций, согласно которым он: 1) анализирует доменное имя для анализируемой ссылки на вредоносность с использованием по меньшей мере одной известной ему методики анализа доменных имен; 2) получает или загружает по меньшей мере один файл, находящийся по анализируемой ссылке, с последующим его анализом на вредоносность с использованием по меньшей мере одной известной ему методики анализа файлов; и 3) получает html-код веб-ресурса, находящегося по анализируемой ссылке, с последующим его анализом на вредоносность с использованием по меньшей мере одной известной ему методики анализа html-кода.

При анализе доменного имени для любой анализируемой ссылки на вредоносность анализирующий модуль 100 (i) получает доступ к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или устанавливает с ним связь с использованием шины 30 связи с обеспечением получения из него данных об известных вредоносных доменных именах, (ii) устанавливает или выявляет, путем посимвольного сравнения каждого анализируемого доменного имени с известными вредоносными доменными именами из указанных полученных данных, факт по меньшей мере частичного совпадения этого анализируемого доменного имени

с одним из известных вредоносных доменных имен. Если анализирующий модуль 100 установил или выявил, что анализируемое доменное имя не имеет по меньшей мере частичного совпадения ни с одним из известных вредоносных доменных имен, то он может дополнительно применить в отношении такого анализируемого доменного имени по меньшей мере одну из известных ему методик анализа доменных имен на подозрительность, например методику анализа доменного имени на основании его длины (при этом чем длиннее доменное имя, тем оно подозрительнее), методику анализа доменного имени на основании его энтропии (при этом чем выше информационная энтропия, вычисленная для конкретного доменного имени по общеизвестной формуле Шеннона, тем подозрительнее это доменное имя), методику анализа доменного имени на основании его осмысленности и/или методику анализа доменного имени на основании правильности его написания. В качестве примера в случае, когда анализирующий модуль 100 анализирует доменные имя на вредоносность с использованием методики анализа доменных имен на основании правильности его написания, то он выполняет по меньшей мере следующие операции, согласно которым он: (i) устанавливает связь с локальным хранилищем 20 данных (обособленным локальным или удаленным хранилищем данных в зависимости от варианта реализации, как описано выше в данном документе) для получения из него данных о языковых словарях, (ii) извлекает по меньшей мере одно слово из каждого из полученных доменных имен, (iii) определяет расстояние Левенштейна между каждым из указанных извлеченных слов и соответствующим одним из слов в языковых словарях из указанных полученных данных и (iv) сравнивает определенное расстояние Левенштейна с заданным пороговым значением, в качестве которого, например, может быть использована константа, равная двум (2), с обеспечением отнесения анализируемого доменного имени к вредоносным доменным именам, если определенное расстояние Левенштейна превышает заданное пороговое значение, равное, например, двум (2).

Таким образом, если анализирующий модуль 100 установил или выявил, посредством по меньшей мере одной из вышеописанных методик анализа, что доменное имя для конкретной анализируемой ссылки относится к вредоносным доменным именам, то это свидетельствует о том, что эта ссылка относится к вредоносным ссылкам и, соответственно, веб-ресурс, находящейся по этой ссылке, относится к вредоносным веб-ресурсам.

При анализе файла, находящегося по анализируемой ссылке, на вредоносность анализирующий модуль 100 выполняет по меньшей мере следующие операции, согласно которым он: (i) получает файл, находящийся по анализируемой ссылке; (ii) вычисляет хеш-сумму полученного файла; (iii) получает доступ к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или устанавливает с ним связь с использованием шины 30 связи с обеспечением получения из него данных о хеш-суммах известных вредоносных файлов; (iv) устанавливает, путем сравнения вычисленной хеш-суммы файла с хеш-суммами известных вредоносных файлов из указанных полученных данных, факт совпадения вычисленной хеш-суммы файла с одной из хеш-сумм известных вредоносных файлов.

Таким образом, если анализирующий модуль 100 установил или выявил, что хеш-сумма конкретного файла совпадает с одной из хеш-сумм известных вредоносных файлов, то этот файл относится к вредоносным файлам, что свидетельствует о том, что эта ссылка относится к вредоносным ссылкам и, соответственно, веб-ресурс, находящейся по этой ссылке, относится к вредоносным веб-ресурсам.

Если же анализирующий модуль 100 установил или выявил, что хеш-сумма полученного файла не совпадает ни с одной из хеш-сумм известных вредоносных файлов, то он может дополнительно применить в отношении такого полученного файла по меньшей мере одну из известных ему методик анализа файлов на подозрительность, например методику анализа файлов на подозрительность на основании изменения параметров состояния виртуальных машин, согласно которой анализирующий модуль 100 осуществляет по меньшей мере следующие операции, согласно которым он: (i) запускает каждый полученный файл по меньшей мере на одной виртуальной машине, характеризующейся заданным набором параметров состояния, (ii) регистрирует изменения в заданном наборе параметров состояния указанной по меньшей мере одной виртуальной машины в течение заданного периода времени, (iii) анализирует полученные изменения параметров состояния с использованием заданного набора правил анализа с обеспечением отнесения указанного запущенного файла к вредоносным файлам, если проанализированные изменения параметров состояния характерны для вредоносных файлов.

Таким образом, если анализирующий модуль 100 установил или выявил, посредством по меньшей мере одной из вышеописанных методик анализа, что файл, находящийся по конкретной ссылке, относится к вредоносным файлам, то это свидетельствует о том, что эта ссылка относится к вредоносным ссылками и, соответственно, веб-ресурс, находящейся по этой ссылке, относится к вредоносным веб-ресурсам.

При анализе html-кода веб-ресурса, находящегося по анализируемой ссылке, на вредоносность анализирующий модуль 100 выполняет по меньшей мере следующие операции, согласно которым он: (i) загружает html-код веб-ресурса, находящегося по этой ссылке; (ii) анализирует загруженный html-код на вредоносность с использованием по меньшей мере одной из известных ему методик анализа html-кода, например методики анализа html-кода на основе ключевых слов, указывающих на вредоносный характер веб-ресурса. Кроме того, при анализе загруженного html-кода на вредоносность анализирующий модуль 100 также может загрузить все изображения и/или иные файлы, связанные с веб-ресурсом, например графические элементы оформления (*.JPG, *.PNG и т.п.), таблицы стилей (*.css), JS-скрипты и т.п., на основании списков таких изображений и/или иных файлов, полученных анализирующим модулем 100 из извлеченного html-кода, с обеспечением проверки так называемых screen-сигнатур, то есть поиска похожих изображений и анализа связанных с ними веб-ресурсов, при этом поиск похожих изображений может быть осуществлен, например, с использованием методики поиска похожих изображений на основе общеизвестного метода поиска ближайших соседей. В ходе такого поиска анализирующий модуль 100 определяет, соответствуют ли, например, изображения, размещенные на анализируемом веб-ресурсе, доменному имени и регистрационным данным веб-ресурса, при этом анализирующий модуль 100 также может дополнительно вычислять хеш-суммы всех изображений, присутствующих на анализируемом веб-ресурсе, и устанавливать, совпадает ли каждая вычисленная хеш-сумма изображения с одной из хеш-сумм известных вредоносных элементов, которые могут быть сохранены, например, в локальном хранилище 20 данных. Кроме того, анализирующий модуль 100 может дополнительно проверять так называемые resource-сигнатуры, для чего он может вычислять хеш-суммы всех ранее загруженных ресурсов анализируемого веб-ресурса, например изображений, каскадных таблиц стилей (CSS), JS-файлов, шрифтов и т.п., и устанавливать или выявлять, совпадает ли каждая вычисленная хеш-сумма ресурса с одной из хеш-сумм известных вредоносных ресурсов, которые могут быть сохранены, например, в локальном хранилище 20 данных.

Таким образом, если анализирующий модуль 100 установил или выявил, посредством по меньшей мере одной из вышеописанных известных ему методик анализа html-кода, что веб-ресурс, находящийся по конкретной ссылке, содержит вредоносный контент, то это свидетельствует о том, что эта ссылка относится к вредоносным ссылкам и, соответственно, веб-ресурс, находящейся по этой ссылке, относится к вредоносным веб-ресурсам.

Анализирующий модуль 100 также выполнен с возможностью сохранения сведений о каждом вредоносном веб-ресурсе, выявленном или установленном с использованием по меньшей мере одного из вышеописанных способов анализа веб-ресурсов на вредоносность, в базу данных вредоносных веб-ресурсов, размещенную в локальном хранилище 20 данных (обособленном локальном хранилище данных о взаимосвязанных вредоносных веб-ресурсах, к которому анализирующий модуль 100 может получать доступ или с которым он может устанавливать связь с использованием шины 30 связи, или в обособленном удаленном хранилище данных о взаимосвязанных вредоносных веб-ресурсах, к которому анализирующий модуль 100 может получать доступ или с которым он может устанавливать связь с использованием модуля 10 связи, соединенного с анализирующим модулем 100 посредством шины 30 связи, в зависимости от варианта реализации настоящего изобретения).

Анализирующий модуль 100 также выполнен с возможностью установления или выявления веб-ресурсов, связанных с каждым из вредоносных веб-ресурсов, выявленных в анализирующем модуле 100 с использованием по меньшей мере одного из вышеописанных способов анализа веб-ресурсов на вредоносность.

Для установления веб-ресурсов, связанных с каждым из выявленных вредоносных веб-ресурсов, анализирующий модуль 100 (i) получает доступ к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или устанавливает с ним связь с использованием шины 30 связи с обеспечением получения из него всех остальных сохраненных ссылок на веб-ресурсы; (ii) устанавливает возможную связь между каждой вредоносной ссылкой, по которой находится соответствующий выявленный вредоносный веб-ресурс, и каждой из полученных ссылок; и (iii) в случае установления указанной связи между ссылками объединяет веб-ресурсы, расположенные по этим связанным ссылкам, в группу взаимосвязанных веб-ресурсов. Следует отметить, что каждая такая группа взаимосвязанных веб-ресурсов образована из одного вредоносного веб-ресурса и по меньшей мере одного связанного с ним веб-ресурса, рассматриваемого в качестве потенциально вредоносного веб-ресурса.

Для установления вышеупомянутой связи между ссылками анализирующий модуль 100 выполняет по меньшей мере одну из следующих операций, согласно которым он устанавливает для каждой пары сравниваемых ссылок по меньшей мере одно из следующего: (1) имеют ли доменные имена схожее написание (например, путем их посимвольного сравнения, вычисления расстояния Левенштейна между доменными именами, сравнения их хеш-сумм, вычисленных анализирующим модулем 100, и/или иной известной методики); (2) зарегистрированы ли доменные имена на одно и то же лицо; (3) указаны ли для зарегистрированных доменных имен одни и те же персональные данные регистранта, то есть физического или юридического лица, на которое зарегистрированы доменные имена, в частности телефон, фактический адрес и/или адрес электронной почты; (4) находятся ли доменные имена по одному и тому же IP-адресу; и (5) имеют ли ссылки один и тот же или похожий единый указатель веб-ресурса “URL” (например, путем их посимвольного сравнения, вычисления расстояния

Левенштейна между этими "URL", сравнения их хеш-сумм, вычисленных анализирующим модулем 100, и/или иной известной методики), например www.site.com и www.si1e.com, при этом сведения о лицах, на которые зарегистрированы доменные имена, сведения о персональных данных регистранта (входят в регистрационные данные доменного имени), указанных для зарегистрированных доменных имен, и IP-адресах, по которым находятся зарегистрированные доменные имена, могут быть автоматически получены анализирующим модулем 100 с использованием, например, онлайн-службы «Whois», в частности путем автоматического направления подходящего поискового запроса в онлайн-службу «Whois» и извлечения необходимых сведений из ответа онлайн-службы «Whois» или из веб-страницы с результатами выполнения поискового запроса путем использования, например, специального парсера, встроенного в анализирующий модуль 100 и анализирующего, например, текст ответа онлайн-службы «Whois» или html-код указанной веб-страницы.

Согласно одному из вариантов реализации настоящего изобретения, в качестве дополнения или альтернативы вышеупомянутая связь между ссылками также может быть установлена анализирующим модулем 100 путем сравнения для каждой пары сравниваемых ссылок истории изменения IP-адресов, работающих сервисов, истории доменных имен, истории DNS-серверов, истории изменения DNS-записей, SSL-ключей, SSH-отпечатков, исполняемых файлов и иных параметров веб-ресурсов. Следует отметить, что наличие связи между сравниваемыми ссылками может быть установлено или определено анализирующим модулем 100 на основании совпадения по меньшей мере одного из вышеперечисленных параметров веб-ресурсов. В частности, в одной из разновидностей данного варианта реализации настоящего изобретения связь между веб-ресурсами, находящимися по анализируемым ссылкам, может быть установлена анализирующим модулем 100 путем создания известной математической модели в виде графа, согласно которой вершины создаваемого графа соответствуют по меньшей мере первому веб-ресурсу и по меньшей мере второму веб-ресурсу, а ребра графа представляют собой связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом по меньшей мере по одному параметру из вышеперечисленных параметров, общему по меньшей мере для первого веб-ресурса и по меньшей мере для второго веб-ресурса. В такой разновидности вышеописанного варианта реализации настоящего изобретения анализирующий модуль 100 может быть выполнен с возможностью присвоения, посредством, например, известного алгоритма машинного обучения, весов связям по меньшей мере между первым веб-ресурсом и вторым веб-ресурсом на основании параметра первого веб-ресурса и второго веб-ресурса, при этом количество связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами может быть ограничено пороговым значением. Анализирующий модуль 100 дополнительно выполнен с возможностью определения коэффициента связи как отношения количества связей по одному параметру между одним первым веб-ресурсом и вторыми веб-ресурсами и веса каждой связи по одному параметру между первым веб-ресурсом и вторыми веб-ресурсами и возможностью удаления связей между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом в случае, если значение определенного коэффициента связи меньше заданного порогового значения.

Анализирующий модуль 100 дополнительно выполнен с возможностью анализа на вредоносность каждого из потенциально вредоносных веб-ресурсов в каждой образованной группе взаимосвязанных веб-ресурсов для выявления вредоносных веб-ресурсов среди этих потенциально вредоносных веб-ресурсов путем осуществления по

меньшей мере одного из вышеописанных способов анализа веб-ресурсов на вредоносность.

В случае подтверждения вредоносного характера по меньшей мере одного из вышеописанных потенциально вредоносных веб-ресурсов в конкретной группе взаимосвязанных веб-ресурсов анализирующий модуль 100 сохраняет сведения о каждом из таких взаимосвязанных вредоносных веб-ресурсов в вышеописанную базу данных вредоносных веб-ресурсов, при этом сведения, сохраненные для каждого такого вредоносного веб-ресурса, содержат, помимо прочего, данные, указывающие на связь этого вредоносного веб-ресурса по меньшей мере с одним другим вредоносным веб-ресурсом.

Следует отметить, что при получении каждой новой ссылки на веб-ресурс анализирующий модуль 100 дополнительно проверяет, относится ли веб-ресурс, находящийся по этой полученной ссылке, к вредоносным веб-ресурсам, для чего этот анализирующий модуль 100 (i) получает доступ к вышеописанной базе вредоносных веб-ресурсов для получения из неё данных о выявленных вредоносных веб-ресурсах; (ii) осуществляет поиск этого анализируемого веб-ресурса среди выявленных вредоносных веб-ресурсах полученных данных путем посимвольного сравнения ссылки, по которой находится анализируемый веб-ресурс, с каждой из ссылок, по которым находятся эти выявленные вредоносные веб-ресурсы, для установления факта их по меньшей мере частичного совпадения. Таким образом, если для полученной новой ссылки было установлено, что она по меньшей мере частично совпадает с одной из ссылок, по которым находятся ранее выявленные вредоносные веб-ресурсы, то анализирующий модуль 100 относит веб-ресурс, находящейся по этой новой ссылке, к вредоносным веб-ресурсам. В противном же случае, то есть когда полученная новая ссылка не имеет даже частичного совпадения ни с одной их ссылок, по которым находятся ранее выявленные вредоносные веб-ресурсы, в отношении веб-ресурса, находящегося по этой новой ссылке, осуществляют вышеописанный анализ на вредоносность.

Анализирующий модуль 100 также выполнен с возможностью классификации или идентификации типа угрозы, которую несет в себе каждый из выявленных вредоносных веб-ресурсов, в зависимости от вредоносного контента этого вредоносного веб-ресурса, выявленного с использованием по меньшей мере одного из вышеописанных способов анализа веб-ресурсов на вредоносность (каждый тип угрозы соответствует тем или иным характерным вредоносным элементам, например тексту, приглашающему пользователя выполнить какое-либо действие, файлу определенного формата, скриптам, подменённым логотипам и т.п.). Например, анализирующий модуль 100 может идентифицировать, что конкретный вредоносный веб-ресурс относится к угрозам типа «фишинг», «вредоносный код», «мошенничество», «бот-нет» и/или т.п. Таким образом, для каждого из выявленных вредоносных веб-ресурсов анализирующий модуль 100 дополнительно выполнен с возможностью сохранения в вышеописанную базу вредоносных веб-ресурсов данных о типе угрозы, которую несет в себе этот вредоносный веб-ресурс, при этом эти сохраненные данные о типе угрозы будут ассоциированы с конкретным вредоносным веб-ресурсом.

Кроме того, для каждого выявленного вредоносного веб-ресурса анализирующий модуль 100 выполнен с возможностью сохранения в вышеописанную базу вредоносных веб-ресурсов доказательств или оснований, полученных в результате применения по меньшей мере одного из вышеописанных способов анализа веб-ресурсов на вредоносность в отношении этого веб-ресурса при его анализе и позволивших отнести

этот анализируемый веб-ресурс к вредоносным веб-ресурсам, при этом такие сохраненные доказательства или основания вредоносности веб-ресурса будут ассоциированы с конкретным вредоносным веб-ресурсом.

Кроме того, анализирующий модуль 100 выполнен с возможностью установления или выявления для каждого из выявленных вредоносных веб-ресурсов, сведения о которых сохранены в вышеописанной базе вредоносных веб-ресурсов, по меньшей мере одного уполномоченного субъекта, связанного с этим вредоносным веб-ресурсом. В качестве уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, могут выступать администратор этого вредоносного веб-ресурса, владелец этого вредоносного веб-ресурса, регистратор доменных имен, хостинг-провайдер и/или иные известные физические и юридические лица, способные заблокировать работу этого вредоносного веб-ресурса или повлиять на решение о блокировке или приостановлении функционирования этого вредоносного веб-ресурса.

Для установления уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, анализирующий модуль 100 предварительно настроен или запрограммирован определять по меньшей мере одного из владельца, администратора, хостинг-провайдера и/или регистратора доменных имён, связанных с этим вредоносным веб-ресурсом, а также их контактные данные, такие как, например, фактический адрес, контактный телефон, адрес электронной почты и т.п.

Следует отметить, что вышеупомянутые уполномоченные субъекты, устанавливаемые или выявляемые анализирующим модулем 100, могут быть определены с использованием любого из известных онлайн-сервисов, например онлайн-службы «Whois», и/или любой из известных утилит, такой как, например, утилита «nslookup», на основании, например, доменного имени, используемого для формирования поискового запроса. Следует также отметить, что необходимые контактные данные по меньшей мере некоторых из необходимых уполномоченных субъектов также могут быть получены с использованием любого из этих известных онлайн-сервисов и/или любой из этих известных утилит, поскольку они входят в регистрационные данные доменного имени, указанные для зарегистрированных доменных имен в этих сервисах и/или утилитах. В частности, в любом из известных онлайн-сервисов и/или любой из известных утилит могут быть получены контактные данные владельца конкретного веб-ресурса, а именно его контактный телефон, фактический адрес места его проживания и/или адрес его электронной почты, а также (в случае наличия) контактные данные администратора этого веб-ресурса, а именно его контактный телефон, фактический адрес места его проживания и/или адрес его электронной почты.

Таким образом, для определения владельца, администратора, хостинг-провайдера и/или регистратора доменных имён, связанных с конкретным вредоносным веб-ресурсом, и получения контактных данных владельца и/или администратора этого вредоносного веб-ресурса анализирующий модуль 100 выполнен с возможностью автоматического направления, например, в онлайн-службу «Whois» подходящего поискового запроса, сформированного на основе доменного имени, извлеченного анализирующим модулем 100 из ссылки, по которой находится этот вредоносный веб-ресурс, и с возможностью автоматического извлечения необходимых сведений из ответа этой онлайн-службы «Whois» или из веб-страницы с результатами выполнения поискового запроса путем использования, например, специального парсера, встроенного в анализирующий модуль 100 и анализирующего, например, текст ответа онлайн-службы «Whois» или html-код указанной веб-страницы. Таким образом, из сведений, полученных от любого из известных онлайн-сервисов и/или любой из известных утилит, анализирующий модуль

100 может однозначно установить владельца и администратора доменных имен для каждого из выявленных вредоносных веб-ресурсов, а также контактные данные каждого из них, а также установить наименования регистратора доменных имен и хостинг-провайдера, связанные с этим вредоносным веб-ресурсом.

В локальном хранилище 20 данных предварительно сохранена обновляемая база данных уполномоченных субъектов, предназначенная для хранения сведений об известных уполномоченных субъектах, в частности перечня известных регистраторов доменных имен и их контактных данных, перечня известных хостинг-провайдеров и их контактных данных, а также перечня гос. учреждений, способных повлиять на решение о блокировке или приостановлении функционирования вредоносных веб-ресурсов и т.п., и их контактных данных, при этом контактные данные в этой базе данных уполномоченных субъектов поставлены в соответствие с конкретным уполномоченным субъектом, к которому они относятся. Анализирующий модуль 100 выполнен с возможностью получения доступа к локальному хранилищу 20 данных или возможностью связи с ним с использованием шины 30 связи с обеспечением извлечения из базы данных уполномоченных субъектов контактных данных по меньшей мере одного из интересуемых субъектов, связанных с конкретным вредоносным веб-ресурсом, на основании наименований этих интересуемых субъектов, ранее установленных анализирующим модулем 100 с использованием любого из известных онлайн-сервисов и/или любой из известных утилит, как более подробно описано выше в данном документе. Таким образом, анализирующий модуль 100 извлекает из базы данных уполномоченных субъектов контактные данные регистратора доменных имен и/или хостинг-провайдера, ранее установленных анализирующим модулем 100 для выявленного вредоносного веб-ресурса с использованием любого из известных онлайн-сервисов и/или любой из известных утилит.

Для каждого из выявленных вредоносных веб-ресурсов анализирующий модуль 100 дополнительно выполнен с возможностью сохранения в вышеописанную базу вредоносных веб-ресурсов наименований уполномоченных субъектов, связанных с этим выявленным вредоносным веб-ресурсом, и контактных данных этих уполномоченных субъектов. Таким образом, для каждого из выявленных вредоносных веб-ресурсов анализирующий модуль 100 сохраняет в базу вредоносных веб-ресурсов имя владельца этого веб-ресурса и его контактные данные, имя администратора этого веб-ресурса и его контактные данные, наименование регистратора доменных имен для этого веб-ресурса и его контактные данные и/или наименование хостинг-провайдера для этого веб-ресурса и его контактные данные, при этом каждые такие контактные данные в базе вредоносных веб-ресурсов ассоциированы с конкретным уполномоченным субъектом из вышеописанных уполномоченных субъектов, к которому они относятся, и с конкретным вредоносным веб-ресурсом, с которым связаны эти уполномоченные субъекты.

В одном из вариантов реализации настоящего изобретения для каждого выявленного вредоносного веб-ресурса анализирующий модуль 100 может быть дополнительно выполнен с возможностью получения доступа к локальному хранилищу 20 данных (обособленному локальному хранилищу данных или удаленному хранилищу данных в зависимости от варианта реализации, как описано выше в данном документе) или возможностью связи с ним с использованием шины 30 связи с обеспечением выявления, содержит ли база вредоносных веб-ресурсов сведения об уполномоченных субъектах, связанных с этим вредоносным веб-ресурсом, то есть, например, имя владельца этого вредоносного веб-ресурса и его контактные данные, имя администратора этого

вредоносного веб-ресурса и его контактные данные, наименование регистратора доменных имен для этого вредоносного веб-ресурса и его контактные данные и/или наименование хостинг-провайдера для этого вредоносного веб-ресурса и его контактные данные. В случае, если анализирующий модуль 100 устанавливает, что база вредоносных веб-ресурсов уже содержит все необходимые сведения об уполномоченных субъектах, связанных с этим вредоносным веб-ресурсом, или по меньшей мере часть из таких необходимых сведений, то анализирующий модуль 100 не осуществляет вышеописанные операции, связанные с направлением поисковых запросов в известные онлайн-службы и/или известные утилиты и получением доступа к базе данных уполномоченных субъектов, а сразу приступает к нижеописанному процессу формирования по меньшей мере одного отчета по меньшей мере для одного уполномоченного субъекта, связанного с этим вредоносным веб-ресурсом, на основании указанных сведений об уполномоченных субъектах из базы вредоносных веб-ресурсов.

Еще в одном варианте реализации настоящего изобретения, в котором вычислительное устройство 200 принимает, посредством модуля 10 связи, ссылки на известные вредоносные веб-ресурсы по меньшей мере от одного источника ссылок, имеющего уникальный идентификатор, по которому анализирующий модуль 100 определяет, что принятые потоки данных от указанного по меньшей мере от одного источника ссылок содержат ссылки на веб-ресурсы с вредоносным и/или незаконным контентом, анализирующий модуль 100 может не осуществлять вышеописанный анализ таких принятых ссылок на вредоносность, а сразу направлять поисковый запрос в вышеописанную базу вредоносных веб-ресурсов для выявления, содержит ли эта база сведения об уполномоченных субъектах, связанных с вредоносным веб-ресурсом, находящимся по принятой ссылке, с последующим формированием по меньшей мере одного отчета по меньшей мере для одного уполномоченного субъекта, связанного с этим вредоносным веб-ресурсом, на основании указанных сведений об уполномоченных субъектах из базы вредоносных веб-ресурсов, как более подробно описано ниже. В противном случае, то есть в случае отсутствия сведений об уполномоченных субъектах, связанных с вредоносным веб-ресурсом, в базе вредоносных веб-ресурсов, анализирующий модуль 100 осуществляет вышеописанные операции, связанные с направлением поисковых запросов в известные онлайн-службы и/или известные утилиты и получением доступа к базе данных уполномоченных субъектов, с последующим формированием по меньшей мере одного отчета по меньшей мере для одного уполномоченного субъекта, связанного с этим вредоносным веб-ресурсом, на основании указанных сведений об уполномоченных субъектах из базы вредоносных веб-ресурсов, как более подробно описано ниже.

Следует отметить, что в локальном хранилище 20 данных предварительно сохранен заданный набор шаблонов отчетов, при этом каждый шаблон отчета по сути представляет собой предварительно составленное письмо-обращение, информирующее конкретный уполномоченный субъект о вредоносном характере по меньшей мере одного конкретного веб-ресурса с просьбой принять решение о блокировке или приостановлении функционирования указанного по меньшей мере одного вредоносного веб-ресурса или повлиять на принятие такого решения, при этом каждый шаблон из этого набора шаблонов отчетов поставлен в соответствие или ассоциирован с одним из известных типов угроз, которые могут нести в себе вредоносные веб-ресурсы, и одним из уполномоченным субъектов. Таким образом, для каждого известного уполномоченного субъекта в локальном хранилище 20 данных могут храниться несколько шаблонов отчетов, каждый из которых предварительно составлен в

соответствии только с одним типом угроз из известных типов угроз.

Анализирующий модуль 100 также выполнен с возможностью формирования по меньшей мере одного отчета по меньшей мере для одного уполномоченного субъекта по истечению заданного периода времени (например, каждые 10 минут, раз в полчаса, раз в час, раз в несколько часов, раз в сутки, раз в неделю и т.п.) или по существу в режиме реального времени на основании следующей информации:

- данных по меньшей мере об одном из вредоносных веб-ресурсов, связанных с одним из указанных уполномоченных субъектов и извлекаемых анализирующим модулем 100 из вышеописанной базы вредоносных веб-ресурсов по меньшей мере на основании наименования этого уполномоченного субъекта, и

- конкретного шаблона отчета, соответствующего одному из указанных уполномоченных субъектов и одному из типов угроз, идентифицированных анализирующим модулем 100 для указанных вредоносных веб-ресурсов, и извлеченных им из базы вредоносных веб-ресурсов по меньшей мере на основании данных об указанных вредоносных веб-ресурсах, в частности уникального идентификатора каждого из указанных вредоносных веб-ресурсов.

Таким образом, анализирующий модуль 100 может, например, сформировать по одному отчету для одного из известных хостинг-провайдеров и одного из известных регистраторов доменных имен, при этом в каждый такой отчет могут быть включены сведения как сразу о нескольких конкретных вредоносных веб-ресурсах (в случае, если эти веб-ресурсы несут в себе угрозу одного и того же типа, например угрозу типа «фишинг», и ассоциированы соответственно с одним и тем же хостинг-провайдером или регистратором доменных имен), так и сведения только об одном конкретном вредоносном веб-ресурсе (в случае, если он несет в себе угрозу типа, отличного от других вредоносных веб-ресурсов, и/или ассоциирован соответственно с хостинг-провайдером или регистратором доменных имен, отличным от других вредоносных веб-ресурсов). В качестве дополнения или альтернативы анализирующий модуль 100 может, например, сформировать по одному отчету для каждого из администраторов веб-ресурсов, связанных с вредоносными веб-ресурсами, сведения о которых были включены в вышеописанные отчет для хостинг-провайдера и отчет для регистратора доменных имен, при этом в каждый такой отчет могут быть включены сведения как сразу о нескольких конкретных вредоносных веб-ресурсах (в случае, если эти веб-ресурсы несут в себе угрозу одного и того же типа, например угрозу типа «мошенничество», и ассоциированы соответственно с одним и тем же администратором), так и сведения только об одном конкретном вредоносном веб-ресурсе (в случае, если он несет в себе угрозу типа, отличного от других вредоносных веб-ресурсов, и/или ассоциирован соответственно с администратором, отличным от других вредоносных веб-ресурсов).

Следует также дополнительно отметить, что количество отчетов, сформированных анализирующим модулем 100 для каждого из вышеописанных уполномоченных субъектов для вредоносных веб-ресурсов, связанных с этим уполномоченным субъектом, в количестве, полученном за заданный период времени, будет соответствовать количеству типов угроз, которые несут в себе эти вредоносные веб-ресурсы.

Возможен вариант реализации, в котором анализирующий модуль 100 по каждому из вредоносных веб-ресурсов будет формировать отчеты для каждого из уполномоченных субъектов, связанных с этим вредоносным веб-ресурсом, в режиме реального времени, то непосредственно сразу после установления того факта, что веб-ресурс, находящийся по принятой ссылке, относится к вредоносным веб-ресурсам,

несущим в себе конкретный тип угрозы, как более подробно описано выше в данном документе.

В одном из вариантов реализации настоящего изобретения анализирующий модуль 100 также может добавлять по меньшей мере в один из отчетов, формируемых анализирующим модулем 100 для уполномоченных субъектов, доказательства вредоносности каждого веб-ресурса, сведения о котором вошли в этот отчет, при этом анализирующий модуль 100 может получать все необходимые доказательства из базы вредоносных веб-ресурсов, в которой они поставлены в соответствии с конкретным вредоносным веб-ресурсом.

Анализирующий модуль 100 также выполнен с возможностью отправки каждого вышеописанного сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта, получаемых анализирующим модулем 100 из базы вредоносных веб-ресурсов, для информирования этого уполномоченного субъекта по меньшей мере об одном веб-ресурсе с вредоносным и/или незаконным контентом.

Согласно одному из вариантов реализации настоящего изобретения, по меньшей мере часть из вышеописанных функциональных возможностей анализирующего модуля 100 может быть реализована с использованием по меньшей мере одного отдельного функционального блока или модуля, которые могут быть необходимым образом соединены с возможностью обмена данными как с анализирующим модулем 100, так и друг с другом.

В качестве примера в одной из разновидностей такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть выполнен с возможностью выполнения исключительно вышеописанной операции выявления вредоносных веб-ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам. Вычислительное устройство 200 может дополнительно содержать, например, отдельный модуль выявления взаимосвязанных веб-ресурсов, соединенный с анализирующим модулем 100 с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции установления веб-ресурсов, связанных с каждым из вредоносных веб-ресурсов, выявленных анализирующим модулем 100, и отдельный модуль информирования о вредоносных веб-ресурсах, соединенный с модулем выявления взаимосвязанных веб-ресурсов и анализирующим модулем 100 с возможностью обмена с ними данными и выполненный с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных анализирующим модулем 100 и/или модулем выявления взаимосвязанных веб-ресурсов, а также вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, и вышеописанной операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения анализирующий модуль 100 может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, модуль выявления взаимосвязанных веб-ресурсов может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи и модуль информирования о вредоносных веб-ресурсах может быть выполнен с возможностью

обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

В качестве еще одного примера в другой разновидности такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть выполнен с возможностью выполнения исключительно вышеописанной операции выявления вредоносных веб-ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам. Вычислительное устройство 200 может дополнительно содержать, например, отдельный модуль выявления взаимосвязанных веб-ресурсов, соединенный с анализирующим модулем 100 с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции установления веб-ресурсов, связанных с каждым из вредоносных веб-ресурсов, выявленных анализирующим модулем 100, а также отдельный модуль установления уполномоченных субъектов, соединенный с модулем выявления взаимосвязанных веб-ресурсов и анализирующим модулем 100 с возможностью обмена с ними данными и выполненный с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных анализирующим модулем 100 и/или модулем выявления взаимосвязанных веб-ресурсов, и отдельный модуль формирования отчетов, соединенный с модулем установления уполномоченных субъектов с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, а также вышеописанной операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения анализирующий модуль 100 может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, а каждый из модуля выявления взаимосвязанных веб-ресурсов, модуля установления уполномоченных субъектов и модуля формирования отчетов может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

В качестве другого примера еще в одной разновидности такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть выполнен с возможностью осуществления вышеописанной операции выявления вредоносных веб-ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам, а также осуществления вышеописанной операции установления веб-ресурсов, связанных с каждым из выявленных вредоносных веб-ресурсов. Вычислительное устройство 200 может дополнительно содержать, например, отдельный модуль установления уполномоченных субъектов, соединенный с анализирующим модулем 100 с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных анализирующим модулем 100, и отдельный модуль формирования отчетов, соединенный с модулем установления уполномоченных субъектов с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, а также вышеописанной

операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения анализирующий модуль 100 может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, а каждый из модуля установления уполномоченных субъектов и модуля формирования отчетов может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

Согласно еще одному варианту реализации настоящего изобретения, анализирующий модуль 100 может быть образован по меньшей мере из одного подмодуля, выполненного с возможностью реализации по меньшей мере части из вышеописанных функциональных возможностей анализирующего модуля 100, при этом такие функциональные подмодули в анализирующем модуле 100 могут быть соединены друг с другом необходимым образом с возможностью обмена данными. В качестве примера в одной из разновидностей такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть образован из подмодуля выявления вредоносных веб-ресурсов, выполненного с возможностью осуществления вышеописанной операции выявления вредоносных веб-ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам, подмодуля выявления взаимосвязанных веб-ресурсов, соединенного с подмодулем выявления вредоносных веб-ресурсов с возможностью обмена данными и выполненного с возможностью выполнения вышеописанной операции установления веб-ресурсов, связанных с каждым из вредоносных веб-ресурсов, выявленных подмодулем выявления вредоносных веб-ресурсов, а также подмодуля установления уполномоченных субъектов, соединенного с подмодулем выявления взаимосвязанных веб-ресурсов и подмодулем выявления вредоносных веб-ресурсов с возможностью обмена с ними данными и выполненный с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных подмодулем выявления вредоносных веб-ресурсов и/или подмодулем выявления взаимосвязанных веб-ресурсов, и подмодуль формирования отчетов, соединенный с подмодулем установления уполномоченных субъектов с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, а также вышеописанной операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения подмодуль выявления вредоносных веб-ресурсов может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, а каждый из подмодуля выявления взаимосвязанных веб-ресурсов, подмодуля установления уполномоченных субъектов и подмодуля формирования отчетов может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

В качестве примера в одной из разновидностей такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть образован из подмодуля выявления вредоносных веб-ресурсов, выполненного с возможностью осуществления вышеописанной операции выявления вредоносных веб-

ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам, подмодуля выявления взаимосвязанных веб-ресурсов, соединенного с подмодулем выявления вредоносных веб-ресурсов с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции установления веб-ресурсов, связанных с каждым из вредоносных веб-ресурсов, выявленных подмодулем выявления вредоносных веб-ресурсов, а также подмодуль информирования о вредоносных веб-ресурсах, соединенный с подмодулем выявления взаимосвязанных веб-ресурсов и подмодулем выявления вредоносных веб-ресурсов с возможностью обмена с ними данными и выполненный с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных подмодулем выявления вредоносных веб-ресурсов и/или подмодулем выявления взаимосвязанных веб-ресурсов, а также вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, и вышеописанной операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения подмодуль выявления вредоносных веб-ресурсов может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, а каждый из подмодуля выявления взаимосвязанных веб-ресурсов и подмодуля информирования о вредоносных веб-ресурсах может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

В качестве другого примера еще в одной разновидности такого варианта реализации настоящего изобретения вышеописанный анализирующий модуль 100 может быть может быть образован из подмодуля выявления вредоносных веб-ресурсов, выполненного с возможностью осуществления вышеописанной операции выявления вредоносных веб-ресурсов во множестве веб-ресурсов, находящихся по полученным ссылкам, а также осуществления вышеописанной операции установления веб-ресурсов, связанных с каждым из выявленных вредоносных веб-ресурсов, а также подмодуля установления уполномоченных субъектов, соединенного с подмодулем выявления вредоносных веб-ресурсов с возможностью обмена данными и выполненного с возможностью выполнения вышеописанной операции установления по меньшей мере одного уполномоченного субъекта, связанного с каждым из вредоносных веб-ресурсов, выявленных подмодулем выявления вредоносных веб-ресурсов, и подмодуль формирования отчетов, соединенный с подмодулем установления уполномоченных субъектов с возможностью обмена данными и выполненный с возможностью выполнения вышеописанной операции формирования по меньшей мере одного отчета по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом, а также вышеописанной операции отправки каждого сформированного отчета в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта. Следует отметить, что в такой разновидности варианта реализации настоящего изобретения подмодуль выявления вредоносных веб-ресурсов может быть выполнен с возможностью обмена данными с модулем 10 связи и локальным хранилищем 20 данных с использованием шины 30 связи, а каждый из подмодуля установления уполномоченных субъектов и подмодуля формирования

отчетов может быть выполнен с возможностью обмена данными с локальным хранилищем 20 данных с использованием шины 30 связи.

На фиг. 3 показана блок-схема способа 400 информирования о вредоносном характере веб-ресурсов согласно настоящему изобретению. Следует отметить, что способ 400 может быть выполнен с использованием вычислительного процессора любого известного вычислительного устройства, в частности с использованием вышеописанного анализирующего модуля 100 вычислительного устройства 200 для информирования о вредоносном характере веб-ресурсов, показанного на фиг. 2.

Способ 400, показанный на фиг. 3, начинается с этапа 410, согласно которому получают ссылки на множество веб-ресурсов.

В одном из вариантов реализации настоящего изобретения для получения ссылок на множество веб-ресурсов на этапе 410 осуществляют по меньшей мере одну из следующих операций, согласно которым: (1) направляют запрос по меньшей мере в один источник ссылок для получения из него по меньшей мере одной ссылки на веб-ресурс; (2) принимают сообщения по меньшей мере от одного вычислительного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки веб-ресурс; (3) принимают сообщения по меньшей мере от одного мобильного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки на веб-ресурс; и (4) вводят поисковые запросы по меньше мере в одну поисковую систему с использованием конкретного перечня ключевых слов для выявления контекстной рекламы в результатах поиска, полученных в ответ на каждый поисковый запрос в каждой из этих поисковых систем, с обеспечением извлечения по меньшей мере одной ссылки на веб-ресурс из выявленной контекстной рекламы.

В дальнейшем способ 400 переходит к выполнению этапа 420, согласно которому выявляют вредоносные веб-ресурсы в указанном множестве веб-ресурсов, а затем к выполнению этапа 430, согласно которому устанавливают веб-ресурсы, связанные с каждым из вредоносных веб-ресурсов, выявленных на вышеописанном этапе 420.

В одном из вариантов реализации настоящего изобретения для установления связанных веб-ресурсов на этапе 430 определяют по меньшей мере одно из следующего: (i) имеют ли доменные имена веб-ресурсов схожее написание; (ii) зарегистрированы ли доменные имена на одно и то же лицо; (iii) указаны ли для зарегистрированных доменных имен веб-ресурсов одни и те же персональные данные регистранта, то есть физического или юридического лица, на которое зарегистрированы доменные имена; (iv) находятся ли доменные имена веб-ресурсов по одному и тому же IP-адресу; и (v) имеют ли ссылки, соответствующие веб-ресурсам, один и тот же или похожий единый указатель веб-ресурса “URL” (например, www.site.com и www.si1e.com).

Еще в одном варианте реализации настоящего изобретения для установления связи веб-ресурсов на этапе 430 осуществляют по меньшей мере следующие операции, согласно которым: (i) создают математическую модель в виде графа, согласно которой вершины создаваемого графа соответствуют по меньшей мере первому веб-ресурсу и по меньшей мере второму веб-ресурсу, а ребра графа представляют собой связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом по меньшей мере по одному параметру веб-ресурса, общему по меньшей мере для первого веб-ресурса и по меньшей мере для второго веб-ресурса, при этом количество связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами ограничено заданным пороговым значением; (ii) присваивают, посредством известного алгоритма машинного обучения, веса связям по меньшей мере между первым веб-ресурсом и вторым веб-ресурсом на основании параметра первого веб-ресурса и второго

веб-ресурса; (iii) определяют коэффициент связи как отношение количества связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами и веса каждой связи по одному параметру веб-ресурса между первым веб-ресурсом и вторыми веб-ресурсами; и (iv) удаляют связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом в случае, если значение определенного коэффициента связи меньше заданного порогового значения.

В дальнейшем способ 400 переходит к выполнению этапа 440, согласно которому выявляют вредоносные веб-ресурсы в связанных веб-ресурсах, установленных на этапе 430.

В некоторых вариантах реализации настоящего изобретения для выявления вредоносных веб-ресурсов на этапе 420 или этапе 440 устанавливают, совпадает ли каждая полученная ссылка по меньшей мере частично с одной из известных вредоносных ссылок.

В других вариантах реализации настоящего изобретения для выявления вредоносных веб-ресурсов на этапе 420 или этапе 440 дополнительно к операции, согласно которой устанавливают, совпадает ли каждая полученная ссылка по меньшей мере частично с одной из известных вредоносных ссылок, осуществляют по меньшей мере одну из следующих операций, согласно которым: (1) анализируют доменное имя веб-ресурса на вредоносность с использованием по меньшей мере одной методики анализа доменных имен; (2) получают с веб-ресурса по меньшей мере один файл для его анализа на вредоносность с использованием по меньшей мере одной методики анализа файлов; и (3) получают html-код веб-ресурса для его анализа на вредоносность с использованием по меньшей мере одной методики анализа html-кода.

В некоторых других вариантах реализации настоящего изобретения при анализе доменного имени веб-ресурса на вредоносность при выполнении операции (1) в рамках выполнения этапа 420 или этапа 440 дополнительно устанавливают, совпадает ли этом анализируемое доменное имя с одним из известных вредоносных доменных имен.

В иных вариантах реализации настоящего изобретения при анализе файла, полученного с веб-ресурса, при выполнении операции (2) в рамках выполнения этапа 420 или этапа 440 дополнительно вычисляют хеш-сумму анализируемого файла, полученного с веб-ресурса, и устанавливают, совпадает ли вычисленная хеш-сумма анализируемого файла с хеш-суммой одного из известных вредоносных файлов.

Еще в других вариантах реализации настоящего изобретения при анализе полученного html-кода веб-ресурса при выполнении операции (3) в рамках выполнения этапа 420 или этапа 440 дополнительно осуществляют поиск в указанном html-коде конкретных ключевых слов, указывающих на вредоносный характер веб-ресурса.

В дальнейшем способ 400 переходит к выполнению этапа 450, согласно которому устанавливают по меньшей мере один уполномоченный субъект, связанный с каждым из вредоносных веб-ресурсов, выявленных на этапе 420 и/или этапе 440.

В одном из вариантов реализации настоящего изобретения при установлении уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, на этапе 450 определяют владельца, администратора, хостинг-провайдера и/или регистратора доменных имён, связанных с этим вредоносным веб-ресурсом. В другом варианте реализации настоящего изобретения для владельца вредоносного веб-ресурса, определенного на этапе 450 при установлении уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, дополнительно направляют запрос хостинг провайдеру и/или регистратору доменных имен, также определенным на этапе 450 при установлении уполномоченных субъектов, связанных

с каждым из выявленных вредоносных веб-ресурсов, и связанным с этим вредоносным веб-ресурсом, с обеспечением получения дополнительных ссылок на веб-ресурсы, связанные с указанным владельцем.

В дальнейшем способ 400 переходит к выполнению этапа 460, согласно которому формируют по меньшей мере один отчет по меньшей мере для одного из уполномоченных субъектов, установленных на этапе 450, на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом.

В одном из вариантов реализации настоящего изобретения способ 400 может включать дополнительный этап, согласно которому устанавливают тип угрозы из заданного набора типов угроз для каждого вредоносного веб-ресурса, выявленного на этапе 420 и/или этапе 440, а при формировании каждого отчета используют шаблон из заданного набора шаблонов отчетов, при этом каждый шаблон соответствует одному из установленных типов угроз и одному из установленных уполномоченных субъектов.

В другом варианте реализации настоящего изобретения количество отчетов, сформированных для каждого уполномоченного субъекта на этапе 460, может соответствовать количеству установленных типов угроз.

Еще в одном варианте реализации настоящего изобретения в каждый отчет, сформированный на этапе 460, могут быть дополнительно добавлены доказательства вредоносности каждого веб-ресурса, сведения о котором содержатся в этом отчете.

В дальнейшем способ 400 переходит к выполнению окончательного этапа 470, согласно которому отправляют каждый отчет, сформированный на этапе 460, в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта.

Следует отметить, что предложенный способ 400 улучшает эффективность информирования уполномоченных субъектов о выявленных веб-ресурсах с вредоносным и/или незаконным контентом как за счет расширения круга уполномоченных субъектов, получающих такие отчеты, так и за счет улучшенной информационной репрезентативности каждого такого отчета, который может сразу охватывать всю группу вредоносных веб-ресурсов, задействованных злоумышленниками и несущих в себе один и тот же тип угрозы.

Представленные иллюстративные варианты осуществления, примеры и описание служат лишь для обеспечения понимания заявляемого технического решения и не являются ограничивающими. Другие возможные варианты осуществления будут ясны специалисту из представленного выше описания. Объем настоящего изобретения ограничен лишь прилагаемой формулой изобретения.

(57) Формула изобретения

1. Способ информирования о вредоносных веб-ресурсах, выполняемый на вычислительном устройстве, при этом согласно указанному способу:

получают ссылки на множество веб-ресурсов,

выявляют вредоносные веб-ресурсы в указанном множестве веб-ресурсов,

устанавливают веб-ресурсы, связанные с каждым из выявленных вредоносных веб-ресурсов,

выявляют вредоносные веб-ресурсы в установленных связанных веб-ресурсах,

устанавливают по меньшей мере один уполномоченный субъект, связанный с каждым из выявленных вредоносных веб-ресурсов,

формируют по меньшей мере один отчет по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных

вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом,

отправляют каждый сформированный отчет в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта.

2. Способ по п. 1, согласно которому при установлении уполномоченных субъектов, связанных с каждым из выявленных вредоносных веб-ресурсов, определяют владельца, администратора, хостинг-провайдера и/или регистратора доменных имён, связанных с этим вредоносным веб-ресурсом.

3. Способ по п. 1, согласно которому дополнительно устанавливают тип угрозы из заданного набора типов угроз для каждого выявленного вредоносного веб-ресурса, а при формировании каждого отчета используют шаблон из заданного набора шаблонов отчетов, при этом каждый шаблон соответствует одному из установленных типов угроз и одному из установленных уполномоченных субъектов.

4. Способ по п. 3, согласно которому количество отчетов, сформированных для каждого уполномоченного субъекта, соответствует количеству установленных типов угроз.

5.Способ по любому из пп. 1-4, согласно которому в каждый отчет дополнительно добавляют доказательства вредоносности каждого веб-ресурса, сведения о котором содержатся в этом отчете.

6. Способ по п. 1, согласно которому для выявления вредоносных веб-ресурсов устанавливают, совпадает ли каждая полученная ссылка по меньшей мере частично с одной из известных вредоносных ссылок.

7. Способ по п. 6, согласно которому для выявления вредоносных веб-ресурсов дополнительно осуществляют по меньшей мере одну из следующих операций, согласно которым:

- анализируют доменное имя веб-ресурса на вредоносность с использованием по меньшей мере одной методики анализа доменных имен,

- получают с веб-ресурса по меньшей мере один файл для его анализа на вредоносность с использованием по меньшей мере одной методики анализа файлов и

- получают html-код веб-ресурса для его анализа на вредоносность с использованием по меньшей мере одной методики анализа html-кода.

8. Способ по п. 7, согласно которому при анализе доменного имени веб-ресурса на вредоносность дополнительно устанавливают, совпадает ли это анализируемое доменное имя с одним из известных вредоносных доменных имен.

9. Способ по п. 7, согласно которому при анализе файла, полученного с веб-ресурса, на вредоносность дополнительно вычисляют его хеш-сумму и устанавливают, совпадает ли вычисленная хеш-сумма анализируемого файла с хеш-суммой одного из известных вредоносных файлов.

10. Способ по п. 7, согласно которому при анализе полученного html-кода веб-ресурса осуществляют поиск в указанном html-коде конкретных ключевых слов, указывающих на вредоносный характер веб-ресурса.

11. Способ по п. 1, согласно которому для установления связанных веб-ресурсов определяют по меньшей мере одно из следующего:

- имеют ли доменные имена веб-ресурсов схожее написание;

- зарегистрированы ли доменные имена на одно и то же лицо;

- указаны ли для зарегистрированных доменных имен веб-ресурсов одни и те же персональные данные регистранта;

- находятся ли доменные имена веб-ресурсов по одному и тому же IP-адресу и

- имеют ли ссылки, соответствующие веб-ресурсам, один и тот же или похожий единый

указатель веб-ресурса “URL”.

12. Способ по п. 1, согласно которому для установления связи веб-ресурсов осуществляют по меньшей мере следующие операции, согласно которым:

- создают математическую модель в виде графа, согласно которой вершины создаваемого графа соответствуют по меньшей мере первому веб-ресурсу и по меньшей мере второму веб-ресурсу, а ребра графа представляют собой связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом по меньшей мере по одному параметру веб-ресурса, общему по меньшей мере для первого веб-ресурса и по меньшей мере для второго веб-ресурса, при этом количество связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами ограничено заданным пороговым значением;

- присваивают, посредством известного алгоритма машинного обучения, веса связям по меньшей мере между первым веб-ресурсом и вторым веб-ресурсом на основании параметра первого веб-ресурса и второго веб-ресурса;

- определяют коэффициент связи как отношение количества связей по одному параметру веб-ресурса между одним первым веб-ресурсом и вторыми веб-ресурсами и веса каждой связи по одному параметру веб-ресурса между первым веб-ресурсом и вторыми веб-ресурсами и

- удаляют связи между по меньшей мере первым веб-ресурсом и по меньшей мере вторым веб-ресурсом в случае, если значение определенного коэффициента связи меньше заданного порогового значения.

13. Способ по п. 1, согласно которому для получения ссылок на множество веб-ресурсов осуществляют по меньшей мере одну из следующих операций, согласно которым:

- направляют запрос по меньшей мере в один источник ссылок для получения из него по меньшей мере одной ссылки на веб-ресурс;

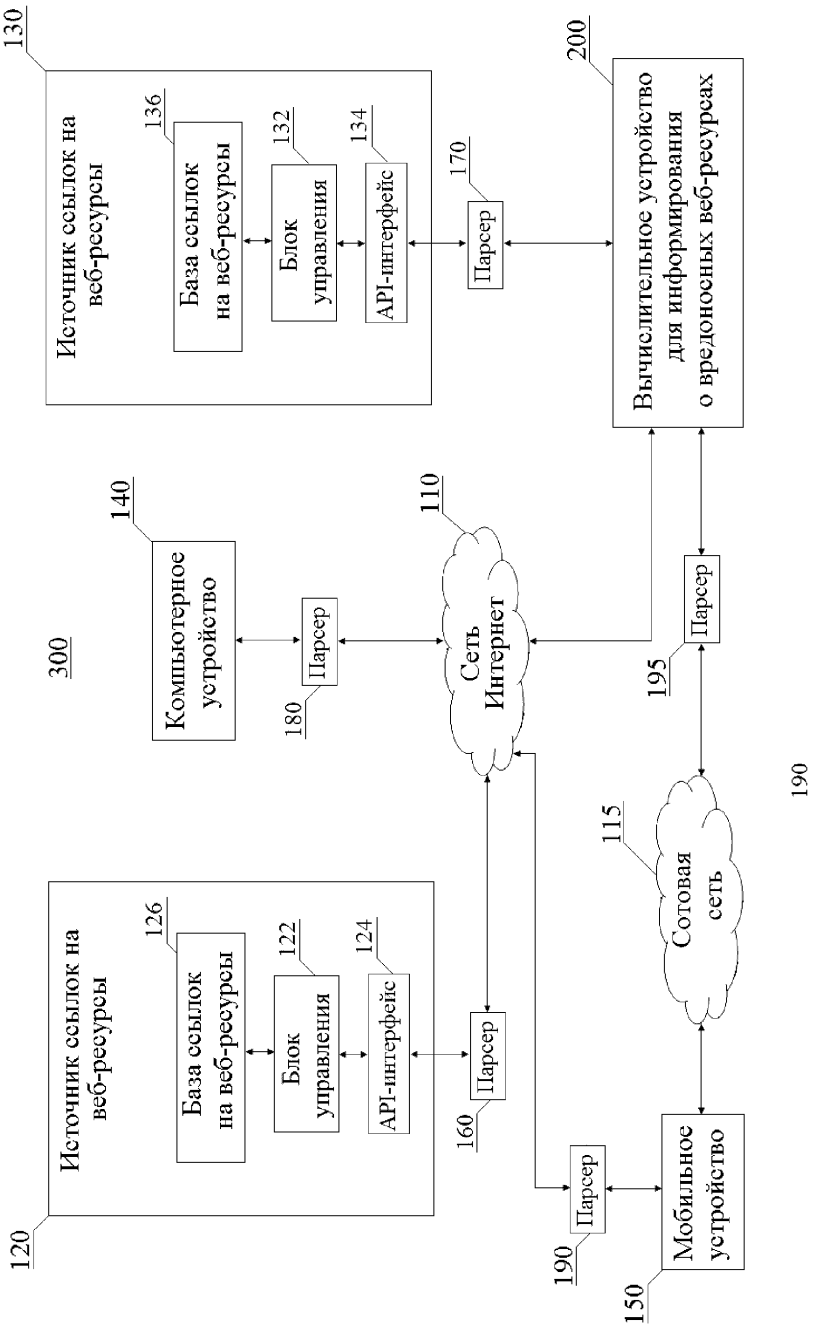
- принимают сообщения по меньшей мере от одного вычислительного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки на веб-ресурс;

- принимают сообщения по меньшей мере от одного мобильного устройства с обеспечением их обработки для извлечения по меньшей мере одной ссылки на веб-ресурс;

- вводят поисковые запросы по меньшей мере в одну поисковую систему с использованием конкретного перечня ключевых слов для выявления контекстной рекламы в результатах поиска, полученных в ответ на каждый запрос в каждой из этих поисковых систем, с обеспечением извлечения по меньшей мере одной ссылки на веб-ресурс из выявленной контекстной рекламы.

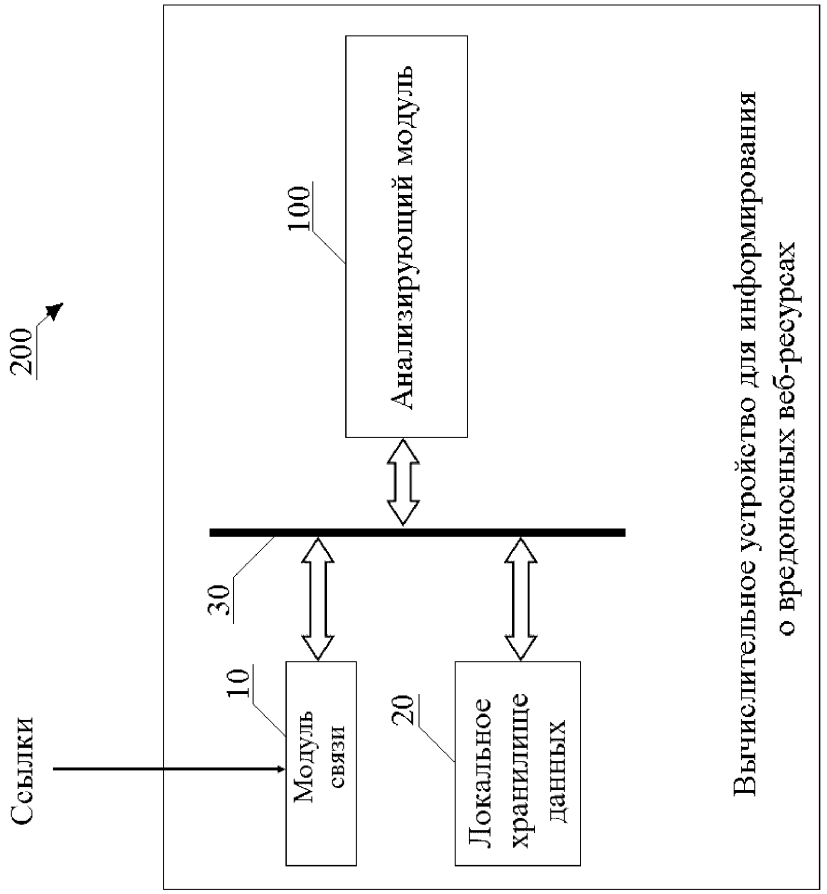
14. Вычислительное устройство для информирования о вредоносных веб-ресурсах, содержащее память для хранения машиночитаемых инструкций и по меньшей мере один вычислительный процессор, выполненный с возможностью исполнения машиночитаемых инструкций с обеспечением осуществления способа информирования о вредоносных веб-ресурсах по любому из пп. 1-13.

1


130
Источник ссылок на веб-ресурсы
136
База ссылок на веб-ресурсы
132
Блок управления
134
API-интерфейс
170
Парсер
200
Вычислительное устройство для информирования о вредоносных веб-ресурсах
140
Компьютерное устройство
180
Парсер
110
Сеть Интернет
195
Парсер
300
190
115
Сотовая сеть
120
Источник ссылок на веб-ресурсы
126
База ссылок на веб-ресурсы
122
Блок управления
124
API-интерфейс
160
Парсер
190
Парсер
150
Мобильное устройство

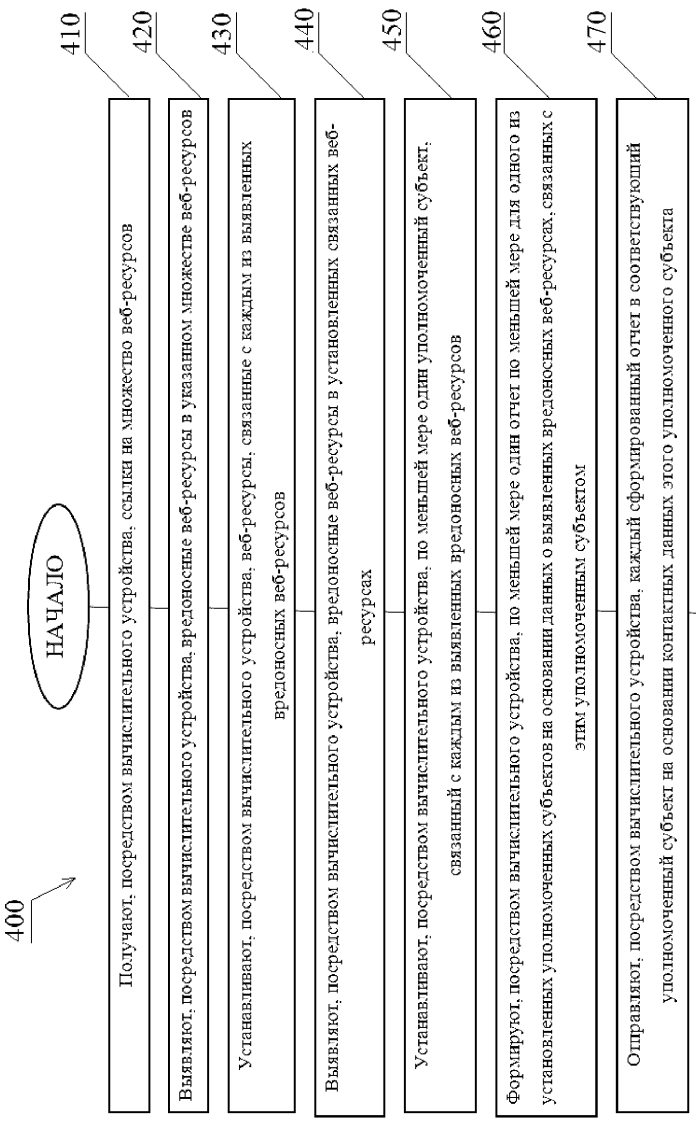
Фиг. 1

2


200
Ссылки
10
Модуль связи
20
Локальное хранилище данных
30
100
Анализирующий модуль
Вычислительное устройство для информирования о вредоносных веб-ресурсах

Фиг. 2

3


400
НАЧАЛО
410
Получают, посредством вычислительного устройства, ссылки на множество веб-ресурсов
420
Выявляют, посредством вычислительного устройства, вредоносные веб-ресурсы в указанном множестве веб-ресурсов
430
Устанавливают, посредством вычислительного устройства, веб-ресурсы, связанные с каждым из выявленных вредоносных веб-ресурсов
440
Выявляют, посредством вычислительного устройства, вредоносные веб-ресурсы в установленных связанных веб-ресурсах
450
Устанавливают, посредством вычислительного устройства, по меньшей мере один уполномоченный субъект, связанный с каждым из выявленных вредоносных веб-ресурсов
460
Формируют, посредством вычислительного устройства, по меньшей мере один отчет по меньшей мере для одного из установленных уполномоченных субъектов на основании данных о выявленных вредоносных веб-ресурсах, связанных с этим уполномоченным субъектом
470
Отправляют, посредством вычислительного устройства, каждый сформированный отчет в соответствующий уполномоченный субъект на основании контактных данных этого уполномоченного субъекта
КОНЕЦ

Фиг. 3