



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК  
G06F 21/00 (2022.08); G06Q 50/00 (2022.08)

(21)(22) Заявка: 2021125359, 27.08.2021

(24) Дата начала отсчета срока действия патента:  
27.08.2021

Дата регистрации:  
06.02.2023

Приоритет(ы):

(22) Дата подачи заявки: 27.08.2021

(45) Опубликовано: 06.02.2023 Бюл. № 4

Адрес для переписки:

115088, Москва, ул. Шарикоподшипниковская,  
1, офис, эт. 9, ком. 17, ООО "Траст", для Марей  
С.В.

(72) Автор(ы):

Нежданов Игорь Юрьевич (RU)

(73) Патентообладатель(и):

Общество с ограниченной ответственностью  
"Траст" (RU)

(56) Список документов, цитированных в отчете  
о поиске: RU 2573760 C2, 27.01.2016. RU  
2637477 C1, 04.12.2017. US 20130086677 A1,  
04.04.2013. RU 2740635 C1, 18.01.2021. CN  
105324786 A, 10.02.2016.

(54) Система и способ выявления информационной атаки

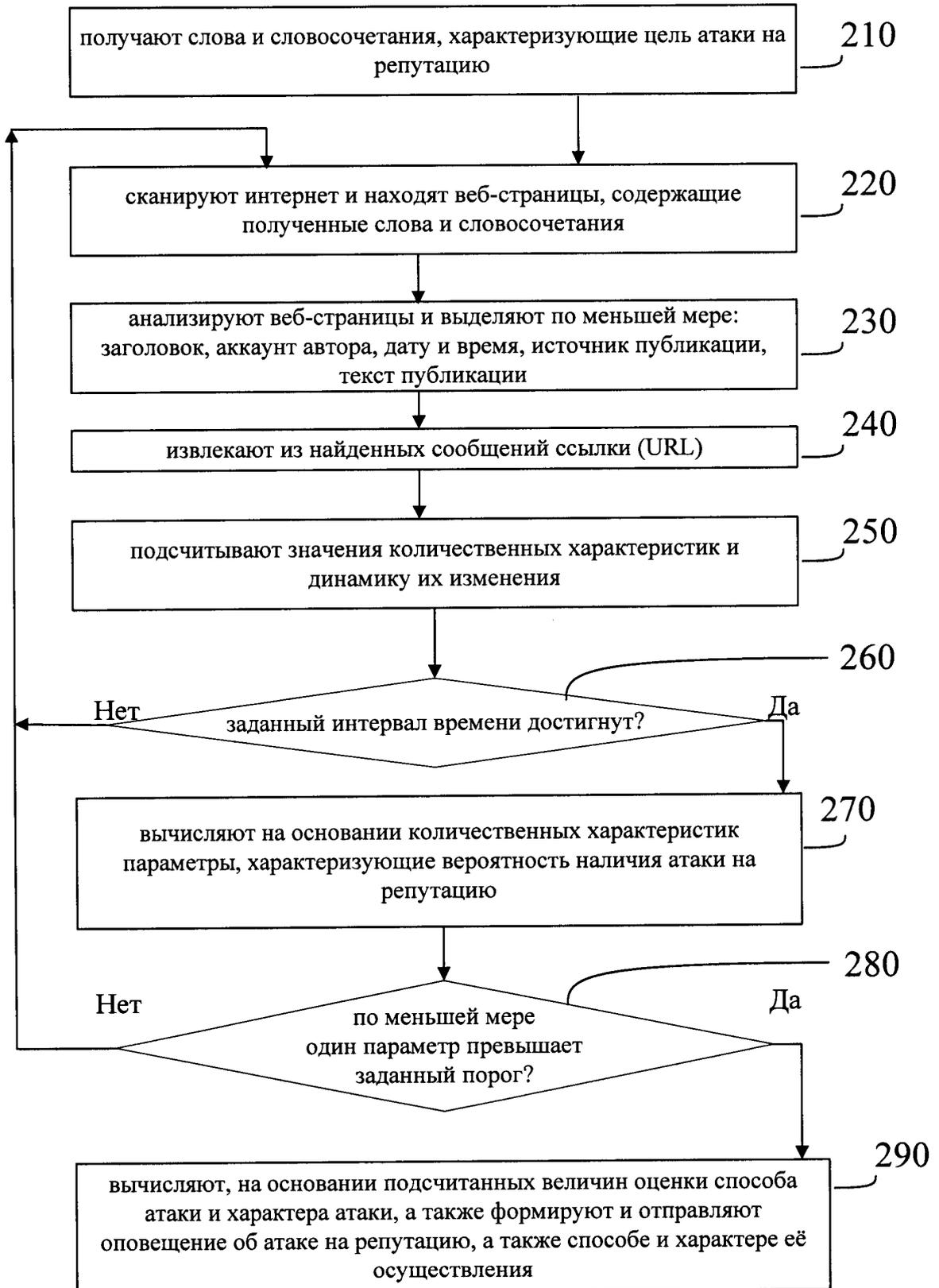
(57) Реферат:

Использование: изобретение относится к области информационной безопасности. Технический результат: обеспечение автоматизированного выявления факта информационной атаки, а также своевременное информирование ответственных лиц об обнаружении атаки. Сущность: способ выявления информационной атаки, выполняемый вычислительным устройством и содержащий шаги, на которых: на предварительном этапе сканируют сеть Интернет и находят источники публикаций, выявляют в составе найденных источников публикаций источники, используемые для информационных атак, находят аккаунты, с которых размещались записи в выявленных источниках публикаций, используемых для информационных атак, выявляют среди найденных аккаунтов те, которые управляются ботами, сохраняют полученные сведения об источниках, используемых для информационных атак, и управляемых ботами аккаунтах в базе данных, затем на рабочем этапе получают слова

и словосочетания, характеризующие цель информационной атаки, сканируют интернет и находят публикации, содержащие слова и словосочетания, характеризующие цель информационной атаки, извлекают из найденных публикаций гиперссылки, подсчитывают, используя сведения об источниках, используемых для информационной атаки, и управляемых ботами аккаунтах, количественные характеристики публикаций и динамику их изменения, вычисляют на основании подсчитанных количественных характеристик параметры, характеризующие вероятность наличия информационной атаки, и в ответ на превышение по меньшей мере одним вычисленным параметром заранее заданного порогового значения определяют, на основании вычисленных параметров, тип атаки и уровень атаки, формируют и отправляют оповещение об информационной атаке, а также о типе атаки и уровне атаки. 2 н. и 10 з.п. ф-лы, 9 ил.

RU 2 789 629 C1

RU 2 789 629 C1



Фиг.2А



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.  
*G06F 21/00* (2013.01)  
*G06Q 50/00* (2012.01)

(12) **ABSTRACT OF INVENTION**

(52) CPC  
*G06F 21/00* (2022.08); *G06Q 50/00* (2022.08)

(21)(22) Application: **2021125359, 27.08.2021**

(24) Effective date for property rights:  
**27.08.2021**

Registration date:  
**06.02.2023**

Priority:  
(22) Date of filing: **27.08.2021**

(45) Date of publication: **06.02.2023** Bull. № 4

Mail address:  
**115088, Moskva, ul. Sharikopodshipnikovskaya, 1,  
ofis, et. 9, kom. 17, OOO "Trast", dlya Mareya S.V.**

(72) Inventor(s):  
**Nezhdanov Igor Yurevich (RU)**

(73) Proprietor(s):  
**Obshchestvo s ogranichennoj otvetstvennostyu  
"Trast" (RU)**

(54) **SYSTEM AND METHOD FOR DETECTION OF INFORMATION ATTACK**

(57) Abstract:

FIELD: information security.

SUBSTANCE: method for detection of an information attack, performed by a computing device, contains steps, at which: at the preliminary stage, the Internet is scanned, and publication sources are found, sources used for information attacks are identified in found publication sources, accounts are found, which published posts in detected publication sources used for information attacks, among found accounts those are identified, which are controlled by bots, obtained data on sources used for information attacks and accounts controlled by bots is stored in a database. Then, at the working stage, words and phrases are obtained, characterizing a target of an information attack, the Internet is scanned, and publications are found, containing words and phrases characterizing the target of the information attack, hyperlinks are extracted

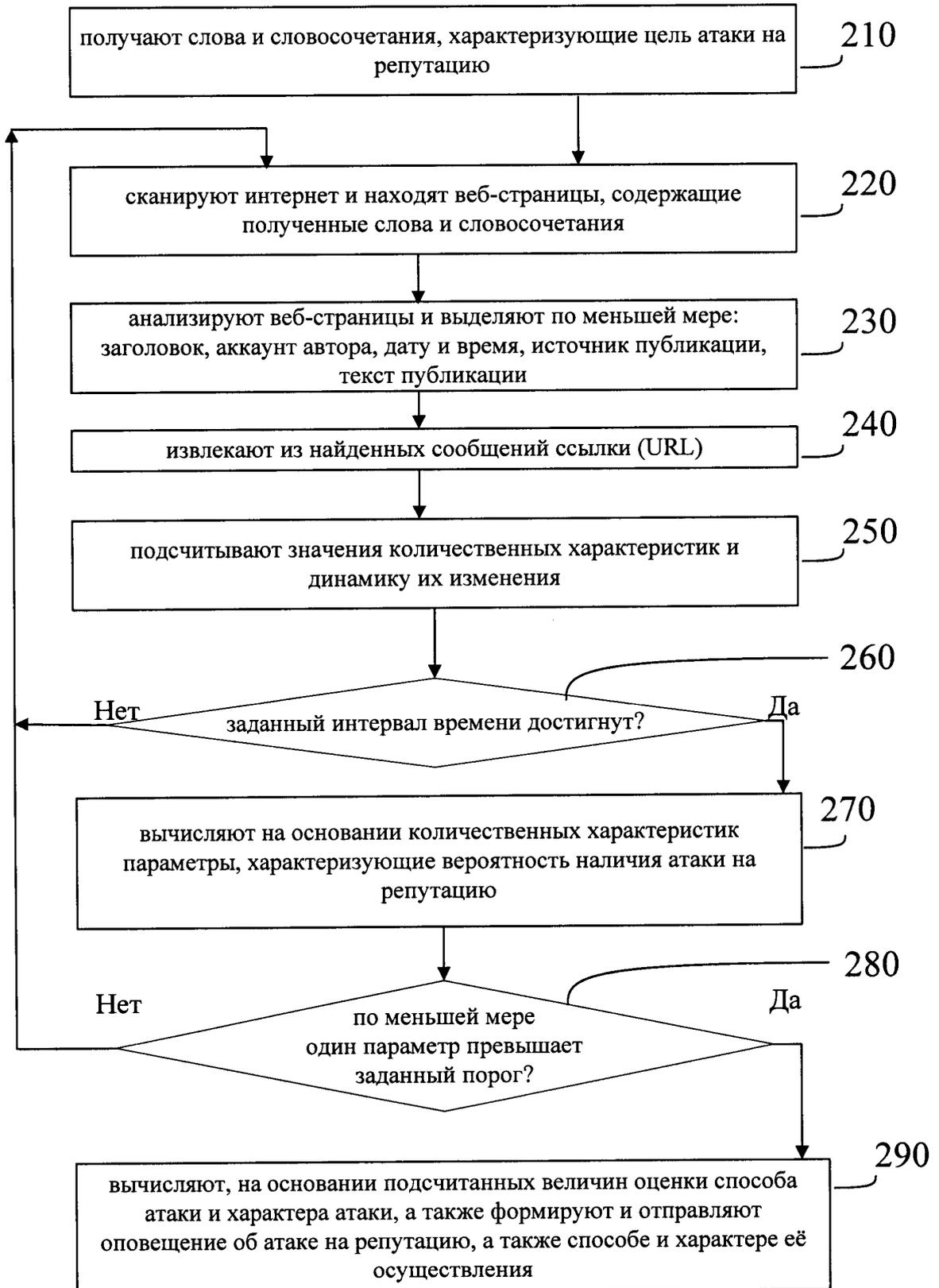
from found publications, using data on sources used for the information attack and accounts controlled by bots, quantitative characteristics of publications and dynamics of their change are counted, based on counted quantitative characteristics, parameters are calculated, characterizing a probability of the presence of an information attack, and, in response to exceeding with at least one calculated parameter a preset threshold value, based on calculated parameters, a type and a level of the attack are determined, a notification about the information attack, as well as about the type and the level of the attack is generated and sent.

EFFECT: provision of automated detection of an information attack, as well as timely informing of responsible persons about attack detection.

12 cl, 9 dwg

RU 2 789 629 C1

RU 2 789 629 C1



Фиг.2А

## ОБЛАСТЬ ТЕХНИКИ

[0001] Изобретение относится к области вычислительной техники, а именно к системам и способам выявления информационных атак, в частности атак на репутацию.

[0002] Атакой на репутацию в рамках настоящего описания называется способ воздействия на общественное мнение, осуществляемого посредством размещения в открытых интернет-источниках сведений, в частности текстов, порочащих репутацию объекта атаки. Иными словами, целью атаки на репутацию является формирование, посредством размещения в сети Интернет определенных публикаций, негативного отношения аудитории к объекту атаки. Термины атака и информационная атака имеют более широкие значения, но в настоящем описании употребляются наряду с термином атака на репутацию в том же самом смысле.

[0003] Объектом атаки может, в качестве неограничивающего примера, являться персона, т.е. конкретный человек; организация; проект, например, такой, как строительство Крымского моста; бренд, такой как "Адидас" или "Пятерочка"; территория или страна; событие или мероприятие, например, такое как праздник выпускников "Алые паруса"; технология или изделие, например, вакцина "Спутник V" или космическая ракета "Ангара".

## УРОВЕНЬ ТЕХНИКИ

[0004] Способы влияния на общественное мнение, в частности способы ухудшать или улучшать чью-либо репутацию, известны человечеству издревле. Однако, появление и развитие глобальной сети Интернет в роли средства массовых коммуникаций породило целый пласт новых способов, методик и техник манипуляции общественным мнением. Преследуя цели, древние как само человеческое общество, эти способы манипуляции, тем не менее, зачастую являются технически новыми. Что, в свою очередь, порождает потребность в использовании технически новых средств и способов для по меньшей мере выявления таких манипуляций.

[0005] Из уровня техники известна публикация "ОТРАЖЕНИЕ ИНФОРМАЦИОННОЙ АТАКИ: АЛГОРИТМ ДЕЙСТВИЙ" (Д. Шубенок, И. Ашманов, опубл. 28 мая 2018 года), размещенная на момент регистрации настоящей заявки по адресу <https://www.ashmanov.com/education/articles/otrazhenie-informatsionnoy-ataki-algoritm-deystviy/>.

[0006] Данная публикация носит скорее описательный характер; она указывает, какие именно современные инструменты в принципе могут использоваться для атак на репутацию, но не раскрывает конкретные подходы к выявлению таких атак. Кроме того, данная публикация содержит описание лишь одного из возможных сценариев атаки, тогда как подобных сценариев достаточно много, и выявление атак, отличающихся от описанного сценария, зачастую требует учета иных факторов, нежели те, что указаны авторами.

[0007] Тем не менее, данная публикация раскрывает достаточно важный в контексте настоящей заявки факт, а именно то, что атака на репутацию в подавляющем большинстве случаев выполняется не одиночным актором, а массированно, с использованием значительного количества аккаунтов, зачастую управляемых автоматизированно, специальными программами (ботами).

[0008] Также из уровня техники известен патент RU2656583C1, "СИСТЕМА АВТОМАТИЗИРОВАННОГО АНАЛИЗА ФАКТОВ" (АО "Крибрум", опубл. 05.06.2018), раскрывающий систему проверки и анализа поведенческих действий пользователей в социальных медиа. Технический результат соответствующего способа заключается в повышении эффективности автоматизированного выявления

поведенческих рисков пользователей социальных медиа.

[0009] Иными словами, хотя данная система и относится к системам, нацеленным на выявление способов влияния на общественное мнение, основная ее функция заключается в анализе публикаций пользователей социальных сетей и определении их навыков, а также уровня угрозы, которую может представлять конкретный пользователь. Способы выявления атак на репутацию указанным патентом не раскрываются, в отличие от описанного ниже способа.

[0010] Кроме того, из уровня техники известна публикация US20110113096A1, "System and method for monitoring activity of a specified user on internet-based social networks" (Profile Protector LLC, опубл. 12.05.2011), где раскрывается система и способ мониторинга активности в социальной сети. Критерии мониторинга заранее устанавливаются клиентом для мониторинга активности на странице определенного пользователя в социальной сети. Доступ для мониторинга активности к странице указанного пользователя в социальной сети устанавливается через интерфейс прикладного программирования социальной сети на основе заранее установленной идентификационной информации, которая идентифицирует указанного пользователя в социальной сети. Клиент получает уведомление, когда отслеживаемая активность удовлетворяет хотя бы одному из предварительно установленных критериев мониторинга.

[0011] Несложно видеть, что данная публикация также посвящена анализу активности заранее заданного аккаунта (пользователя социальной сети), и не раскрывает, в отличие от описанного ниже способа, выявление факта атаки на репутацию.

[0012] На основании результатов исследования уровня техники можно сделать вывод, что существует потребность в техническом решении, устраняющим недостатки описанных выше подходов. Описываемое ниже решение создано для решения по меньшей мере части проблем, выявленных при анализе предшествующего уровня техники.

### РАСКРЫТИЕ (СУЩНОСТЬ) ИЗОБРЕТЕНИЯ

[0013] Задача предполагаемого изобретения заключается в разработке способа и системы выявления информационных атак.

[0014] Техническим результатом заявленной технологии является автоматизированное выявление факта информационной атаки, а также своевременное информирование ответственных лиц об обнаружении атаки.

[0015] Данный технический результат достигается за счет того, что способ выявления информационных атак, выполняемый вычислительным устройством, содержит шаги, на которых на предварительном этапе сканируют сеть Интернет и находят источники публикаций, выявляют в составе найденных источников публикаций источники, используемые для атак, находят аккаунты, с которых размещались записи в выявленных источниках публикаций, выявляют среди найденных аккаунтов те, которые управляются ботами, сохраняют полученные сведения об источниках, используемых для атак, и управляемых ботами аккаунтах в базе данных; затем на рабочем этапе получают слова и словосочетания, характеризующие цель атаки, сканируют интернет и находят публикации, содержащие слова и словосочетания, характеризующие цель атаки, извлекают из найденных публикаций гиперссылки, подсчитывают, используя сведения об источниках, используемых для атак, и управляемых ботами аккаунтах, количественные характеристики публикаций и динамику их изменения, вычисляют на основании подсчитанных количественных характеристик параметры, характеризующие вероятность наличия атаки, и в ответ на превышение по меньшей мере одним

вычисленным параметром заранее заданного порогового значения определяют, на основании вычисленных параметров, тип атаки и уровень атаки, формируют и отправляют оповещение об атаке, а также о типе атаки и уровне атаки.

5 [0016] Технический результат также достигается за счет того, что система выявления информационных атак, выполненная с возможностью сканирования сети Интернет, содержит, по меньшей мере, процессор, а также запоминающее устройство, содержащее по меньшей мере одну базу данных, а также машиночитаемые инструкции, которые при исполнении их процессором обеспечивают выполнение описанного способа.

10 [0017] В частном варианте реализации способ отличается тем, что к источникам публикаций, используемым для атак, относятся по меньшей мере следующие:

- агрегаторы компромата,
- социальные сети,
- агрегаторы утечек данных,
- рекламные площадки,
- 15 - группы связанных источников,
- агрегаторы отзывов пользователей,
- площадки для найма сотрудников на удаленную работу.

[0018] В другом частном варианте реализации способ отличается тем, что к группам связанных источников относят группы источников, не менее заданного количества раз  
20 разместивших идентичные публикации с разницей во времени публикации, не превышающей заранее заданное пороговое значение.

[0019] Еще в одном возможном варианте реализации способ отличается тем, что к аккаунтам, которые управляются ботами, относят аккаунты, сделавшие за заранее заданный промежуток времени не менее заранее заданного количества публикаций.

25 [0020] Еще в одном возможном варианте реализации способ отличается тем, что к количественным характеристикам публикаций относят по меньшей мере следующие величины:

- общее количество публикаций,
- количество публикаций, сделанных ботами,
- 30 - количество публикаций, сделанных на агрегаторах компромата,
- количество публикаций, сделанных группами связанных источников публикаций, количество публикаций, сделанных группами связанных источников, которые также являются агрегаторами компромата,
- количество публикаций, сделанных на рекламных площадках,
- 35 - количество публикаций, сделанных на рекламных площадках, входящих в группу связанных источников,
- количество публикаций, сделанных на агрегаторах отзывов пользователей,
- количество публикаций, сделанных на агрегаторах утечек,
- количество публикаций, сделанных на площадках для найма сотрудников на  
40 удаленную работу,
- общее количество публикаций, являющихся дублями друг друга,
- общее количество публикаций на агрегаторах компромата, являющихся дублями друг друга,
- общее количество публикаций на агрегаторах компромата, являющихся дублями  
45 друг друга и сделанных ботами,
- общее количество ссылок, являющихся дублями друг друга,
- количество аккаунтов, с которых были размещены найденные публикации,
- количество аккаунтов, управляемых ботами, с которых были размещены найденные

публикации,

- количество аккаунтов, с которых были размещены публикации, найденные на агрегаторах компромата,

5 - количество аккаунтов, управляемых ботами, с которых были размещены публикации на агрегаторах компромата,

- количество аккаунтов, с которых были размещены публикации, найденные на рекламных площадках.

[0021] Еще в одном возможном варианте реализации способ отличается тем, что динамику изменения количественных характеристик вычисляют на основании значения  
10 этих характеристик, вычисленных на протяжении заранее заданного интервала времени с заранее заданным шагом.

[0022] Еще в одном возможном варианте реализации способ отличается тем, что параметры, характеризующие вероятность наличия атаки, для каждой количественной характеристики вычисляют как абсолютную, выраженную в единицах, и относительную,  
15 выраженную в процентах, разность между соседними значениями данной характеристики.

[0023] Еще в одном возможном варианте реализации способ отличается тем, что оповещение об атаке может иметь численное выражение, характеризующее уровень интенсивности атаки.

20 [0024] Еще в одном возможном варианте реализации способ отличается тем, что оповещение об атаке может иметь один из трех уровней: "Предупреждение", "Угроза", "Атака".

[0025] Еще в одном возможном варианте реализации способ отличается тем, что по меньшей мере одно сформированное предупреждение об атаке передают посредством  
25 по меньшей мере одного из следующих способов коммуникации:

- электронной почты (e-mail),

- SMS,

- MMS,

- push-уведомления,

30 - сообщения в программе обмена мгновенными сообщениями,

- события API

#### КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[0026] Сопровождающие чертежи, которые включены для обеспечения  
35 дополнительного понимания изобретения и составляют часть этого описания, показывают варианты осуществления изобретения и совместно с описанием служат для объяснения принципов изобретения.

[0027] Заявленное изобретение поясняется следующими чертежами, на которых:

[0028] Фиг. 1А показывает блок-схему алгоритма предварительного этапа описываемого способа.

40 [0029] Фиг. 1Б показывает блок-схему алгоритма одного из шагов предварительного этапа описываемого способа.

[0030] Фиг. 1В показывает блок-схему алгоритма еще одного из шагов предварительного этапа описываемого способа.

[0031] Фиг. 1Г показывает блок-схему алгоритма еще одного из шагов  
45 предварительного этапа описываемого способа.

[0032] Фиг. 2А показывает блок-схему алгоритма рабочего этапа описываемого способа.

[0033] Фиг. 2Б показывает блок-схему алгоритма одного из шагов рабочего этапа

описываемого способа.

[0034] Фиг. 2Б показывает блок-схему алгоритма еще одного из шагов рабочего этапа описываемого способа.

5 [0035] Фиг. 3 показывает блок-схему одного из возможных алгоритмов вычисления оценок способа атаки и характера атаки.

[0036] Фиг. 4 иллюстрирует неограничивающий пример общей схемы вычислительного устройства..

## ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ

10 [0037] Ниже будет приведено описание примерных вариантов осуществления заявленного изобретения.

[0038] Объекты и признаки настоящего изобретения, способы для достижения этих объектов и признаков станут очевидными посредством отсылки к примерным вариантам осуществления. Однако настоящее изобретение не ограничивается примерными вариантами осуществления, раскрытыми ниже, оно может воплощаться в различных видах. Сущность, приведенная в описании, является ничем иным, как конкретными 15 деталями, обеспеченными для помощи специалисту в области техники в исчерпывающем понимании изобретения, и настоящее изобретение определяется только в объеме приложенной формулы.

[0039] При последующем описании способа и системы выявления атак на репутацию 20 используются следующие основные термины и определения:

[0040] Аккаунт - уникальная учетная запись, создание которой является необходимым и достаточным условием участия конкретного пользователя в коммуникациях посредством данного веб-сайта или данной социальной сети. Характеризуется наличием уникального в рамках данного веб-сайта или социальной сети идентификатора 25 пользователя: имени пользователя, его порядкового номера или иного сочетания символов.

[0041] Социальная сеть - интернет-площадка, которая позволяет зарегистрированным (имеющим аккаунт данной сети) пользователям коммуницировать между собой. Контент на такой площадке создается самими пользователями. С точки зрения интерфейса 30 пользователя социальная сеть может представлять собой как веб-сайт, например, такой, как vk.com, facebook.com, так и программу обмена мгновенными сообщениями, интернет-мессенджер, такой как Telegram или Discord.

[0042] Источник или источник публикаций (в данном случае) - веб-сайт или сообщество (канал, группа, сервер) в социальной сети, специализирующееся на 35 размещении текстов. В рамках данного описания к источникам относят:

- СМИ, на веб-сайтах которых могут находиться как собственно публикации, так и комментарии под публикациями;
- форумы;
- блоги журналистов, политиков и общественных деятелей;
- 40 • сообщества (группы, паблики) в социальных сетях;
- видеохостинги и стрим-серверы;
- сервисы вопросов и ответов;
- сервисы сбора подписей под петициями и обращениями;
- сервисы краудфандинга;
- 45 • веб-сайты, выполняющие функции:
  - a. агрегаторов отзывов пользователей,
  - b. рейтинговых агентств,
  - c. "досок объявлений", в том числе:

- i. бирж аккаунтов,
- ii. площадок для найма сотрудников на удаленную работу.

[0043] "Доска объявлений" (в данном случае) - веб-сайт, предоставляющий услуги размещения объявлений определенной или произвольной тематики.

5 [0044] Биржа аккаунтов (в данном случае) - разновидность "доски объявлений", на которой размещают предложения о продаже или сдаче в аренду аккаунтов, принадлежащих людям или ботам, а также сообщения о желании приобрести или взять в аренду такие аккаунты.

10 [0045] Бот (в данном случае) - аккаунт, управляемый программой, которая выполнена с возможностью оставлять сообщения от имени одного из пользователей заданной социальной сети. Как правило, после первичной настройки бот действует автономно, и без участия оператора размещает в заданной социальной сети сообщения заданного содержания.

15 [0046] Группа связанных источников (в данном случае) - группа источников, размещение текстов на которых выполняются одним человеком или одной организованной группой лиц.

[0047] Агрегатор компромата (в данном случае) - источник, размещающий только тексты, имеющие характер компрометирующих материалов. Примером такого источника является вебсайт [compromat.ru](http://compromat.ru).

20 [0048] Агрегатор утечек данных (в данном случае) - источник, размещающий только тексты, имеющие характер утечек данных (инсайдов). Примером такого источника является вебсайт [WikiLeaks](http://WikiLeaks).

[0049] Рейтинговое агентство (в данном случае) - веб-сайт, основная функциональность которого заключается в формировании и показе рейтинга веб-сайтов определенной специализации. Например, рейтинг самых влиятельных отзывиков (агрегаторов отзывов пользователей), рейтинг бирж аккаунтов, рейтинг бирж SMM-услуг, и так далее.

30 [0050] Рекламная площадка (в данном случае) - источник, представляющий собой средство массовой информации, размещающее новости, но отличающийся тем, что допускает размещение на правах рекламы текста произвольного содержания под видом очередной новости.

[0051] Следует также отметить, что контексте настоящего описания, если конкретно не указано иное, слова «первый» и «второй» используются исключительно для того, чтобы отличать существительные, к которым они относятся, друг от друга, а не для целей описания какой-либо конкретной взаимосвязи между этими существительными.

[0052] Для реализации описываемого способа выявления атак на репутацию сначала выполняют предварительный этап (100), как это описано ниже со ссылкой на Фиг. 1А.

40 [0053] Предварительный этап (100) начинается с шага (110), на котором сканируют интернет и находят веб-страницы, содержащие публикации. Сканирование выполняют любым общеизвестным способом, при помощи какой-либо программы, реализующей функции веб-парсера, то есть автоматического "сборщика" публикаций с различных веб-сайтов, такой, например, как [CloudScrape](http://CloudScrape) или [Scrapinghub](http://Scrapinghub). В одном варианте реализации описываемого способа перед сканированием задают язык или языки, на котором должны быть написаны публикации (например, русский, либо русский и английский). В другом возможном варианте реализации поиск ведется без ограничения по языку.

[0054] Возможен также вариант реализации описываемого способа, при котором источники публикаций дополнительно извлекают из веб-страниц, полученных в ходе

вышеупомянутого сбора публикаций с различных веб-сайтов. Автоматизированная, например, выполняемая заранее подготовленным скриптом, обработка (парсинг) таких веб-страниц может быть использована для извлечения из них ссылок на источники публикаций и пополнения извлекаемыми ссылками общего списка публикаций.

5 [0055] Возможен также вариант реализации способа, при котором дополнительно находят источники публикаций, анализируя электронные письма (email), в том числе незапрошенные рассылки (спам). Это может выполняться любым общеизвестным образом. Например, может быть заблаговременно зарегистрирован ряд аккаунтов электронной почты, адреса которых могли быть размещены в открытом доступе.  
10 Подобные адреса, как правило, вскоре попадают в списки рассылок (спам), и на эти адреса начинают поступать электронные письма, в том числе, содержащие ссылки на различные вышеперечисленные источники публикаций. Автоматизированная, например, выполняемая заранее подготовленным скриптом, обработка (парсинг) таких писем может быть использована для извлечения из них ссылок на источники публикаций и  
15 пополнения извлекаемыми ссылками общего списка публикаций.

[0056] Результатом выполнения шага (110) становится сохраненный в базе данных список обнаруженных веб-страниц.

[0057] На этом шаг (110) завершается и способ переходит к шагу (120), на котором анализируют найденные веб-страницы, при этом выделяют и сохраняют по меньшей  
20 мере: заголовок, аккаунт (автора), гиперссылку (URL) на веб-страницу, время ее появления в открытом доступе (время и дата публикации), ее источник, например, образованный усечением гиперссылки до доменного имени второго или третьего уровня, а также собственно текст публикации. Подобное выявление перечисленных типов данных на веб-странице является одной из типовых функций веб-парсеров и может  
25 выполняться средствами используемой программы. Альтернативно, выделение названных полей может выполняться предварительно изготовленным скриптом, реализующим любой общеизвестный алгоритм.

[0058] Например, в результате выполнения вышеописанного шага (120) в базе данных может быть сохранена публикация с заголовком "Внимание!", имеющая текст: "Я  
30 слышал, что скоро введут налог на домашних животных!", опубликованная с аккаунта sampleuser, с датой и время публикации 11.02.2021 17:21:35, гиперссылка на эту публикацию: <http://www.livejournal.com/sampleuser/12345678.html>, а также источник: [sampleuser.livejournal.com](http://sampleuser.livejournal.com).

[0059] Затем способ переходит к шагу (130), на котором в составе найденных  
35 источников публикаций выявляют по меньшей мере следующие типы источников: социальные сети, агрегаторы компромата, агрегаторы утечек данных, площадки сбора подписей под петициями.

[0060] Следует заметить, что для всех перечисленных источников характерно постоянное использование одного и того же доменного имени. Как правило,  
40 значительную часть бюджета таких источников составляют доходы от рекламы; нередко, привлекая новых пользователей, они ведут собственные рекламные кампании. Поэтому доменные имена таких источников годами остаются одними и теми же; что, в свою очередь, позволяет иметь постоянные списки доменных имен и проверять по ним принадлежность очередного источника к одному из названных типов.

45 [0061] Например, могут существовать отдельные списки, например, список "Социальные сети", в котором хранятся такие доменные имена, как [facebook.com](http://facebook.com), [vk.com](http://vk.com), [livejournal.com](http://livejournal.com) и т.д., список "Агрегаторы компромата", в котором хранятся доменные имена вроде [compromat.ru](http://compromat.ru) или [compromat.livejournal.com](http://compromat.livejournal.com), список "Агрегаторы утечек

данных", в котором хранятся доменные имена наподобие wikileaks.com, а также список "Площадки сбора подписей под петициями", содержащий доменные имена вроде change.org, democrator.ru, e-petition.am, и т.д.

5 [0062] На шаге (130) каждый очередной найденный источник публикаций проверяют поочередно на наличие в каждом из указанных списков. При совпадении проверяемый источник соответственно рубрицируется, то есть в базе данных для него проставляется тэг, соответствующий списку, где было обнаружено его доменное имя.

[0063] Так, в вышерассмотренном примере для публикации, найденной по адресу <http://www.livejournal.com/sampleuser/12345678.html>, а также всех остальных публикаций, найденных на домене livejournal.com, в базе будет проставлен тэг "Социальные сети", 10 поскольку доменное имя livejournal.com будет найдено в списке "Социальные сети".

[0064] Следует отметить, что одно и то же доменное имя, либо похожие доменные имена могут находиться в разных списках. Например, в списке "Социальные сети" может присутствовать доменное имя livejournal.com, а в списке "Агрегаторы компромата" 15 могут присутствовать доменные имена slivaem-kompromat.livejournal.com, compromat.livejournal.com и т.д. По окончанию шага (130) по меньшей мере часть публикаций и соответствующих им источников, которые окажутся найдены в перечисленных списках, будут рубрицированы. Технически рубрикация может представлять собой, например, проставленные в базе данных тэги, каждый из которых 20 соответствует одному из типов источников: "Социальная сеть", "Агрегатор компромата" и т.д. Как уже было отмечено, источник публикаций может одновременно относиться к разным типам источников, поэтому в результате выполнения шага (130) источнику может быть проставлено более одного тэга.

[0065] Затем способ переходит к шагу (140), на котором в составе найденных 25 источников публикаций выявляют группы связанных источников публикаций.

[0066] Все источники, рубрицированные на шаге (130), не исключаются из дальнейшей обработке в ходе следующего шага (140), поскольку, например, группа ("публик") социальной сети может выполнять функции, например, рекламной площадки или входить в группу связанных источников.

30 [0067] Далее со ссылкой на Фиг. 1Б описано выполнение шага (140), на котором в составе найденных источников публикаций выявляют группы связанных источников публикаций.

[0068] Шаг (140) начинается с того, что из найденных публикаций выбирают (141) очередную публикацию. Затем на шаге (142) проверяют, существуют ли среди всех 35 найденных публикаций дубликаты выбранной публикации. Под дубликатом в данном случае понимается строгое совпадение текста выбранной публикации с текстом какой-либо еще публикации.

[0069] Технически шаг (142) заключается в поиске в базе данных всех публикаций из других источников, у которых текст в поле базы данных "Текст публикации" является 40 точной копией текста, который присутствует в данном поле у выбранной публикации. Такой поиск может выполняться любым общеизвестным образом, выбранным в зависимости от архитектуры используемой базы данных.

[0070] В том случае, если на шаге (142) дубликаты не найдены, то есть нет ни одной публикации, текст которой совпадал бы с текстом выбранной публикации, способ 45 возвращается к шагу (141), на котором выбирают очередную публикацию.

[0071] В том случае, если на шаге (142) дубликаты найдены, то есть найдена по меньшей мере одна публикации, текст которой точно совпадает с текстом выбранной публикации, способ переходит к шагу (143).

[0072] На шаге (143) относят группу источников публикаций-дубликатов, найденных на шаге (142), к группе источников-кандидатов. При этом сохраняют в виде отдельного списка те источники, к которым относятся найденные публикации, и проверяют, совпадает ли время публикаций во всех найденных публикациях.

5 [0073] Совпадение времени в данном случае может быть нечетким, когда совпадающим считается время публикации, отличающееся в любую сторону от времени публикации, выбранной на шаге (141), не более чем на заранее заданную величину  $dT$ , например, не более чем на 30 секунд.

10 [0074] Источники, разместившие публикации-дубликаты с большей, нежели заранее заданная величина, разницей во времени  $dT$ , исключают из группы источников-кандидатов.

[0075] Например, если на шаге (142) были найдены следующие источники публикаций-дубликатов, сделавшие публикации в указанное время:

15	• sampleuser.livejournal.com	11.02.2021 17:21:35
	• website.com	11.02.2021 17:21:07
	• sample.newspaper.ru	11.02.2021 17:21:59
	• examplechange.org	15.02.2021 07:01:06

то в результате выполнения шага (143) при заданной величине  $dT$ , равной 30 секунд, в списке группы источников-кандидатов останутся:

20

- sampleuser.livejournal.com
- website.com
- sample.newspaper.ru

25 [0076] После окончания шага (143) способ переходит к шагу (144), на котором проверяют, является ли найденная группа нулевой (пустой). В том случае, если группа источников-кандидатов оказывается нулевой, то есть если все найденные публикации сделаны источниками-кандидатами с разницей во времени большей, чем  $dT$ , то группу источников-кандидатов не сохраняют (удаляют) и способ возвращается к шагу (141), на котором выбирают очередную публикацию.

30 [0077] В том случае, если группа источников-кандидатов оказывается ненулевой, то есть найдены по меньшей мере два источника-кандидата, разместившие публикации с совпадающим текстом с разницей во времени не большей, чем  $dT$ , то список группы источников-кандидатов сохраняют, присваивают счетной переменной  $J$ , значение которой хранится ассоциированно с каждым таким списком, начальное значение  $J=1$ , и способ переходит к шагу (145).

35 [0078] На шаге (145) проверяют, был ли по меньшей мере один из источников-кандидатов, найденных на предыдущем шаге, найден повторно. Иначе говоря, проверяют, входит ли по меньшей мере один из источников-кандидатов, найденных на шаге (144), в по меньшей мере один список групп источников-кандидатов, сохраненных ранее. Это выполняется любым общеизвестным способом, путем поочередного поиска  
40 каждого из источников-кандидатов, найденных на шаге (144), во всех ранее сохраненных списках групп источников-кандидатов.

[0079] В том случае, если все источники-кандидаты, найденные на шаге (144) отсутствуют во всех ранее сохраненных списках групп источников-кандидатов, то есть  
45 группа источников-кандидатов, найденная на шаге (144), является новой, то способ возвращается к шагу (141), на котором выбирают очередную публикацию.

[0080] В том случае, если на шаге (145) будет найден по меньшей мере один источник-кандидат, входящий в по меньшей мере один сохраненный ранее список групп источников-кандидатов, то способ переходит к шагу (146).

[0081] На шаге (146) объединяют списки групп источников-кандидатов, в которых были найдены одни и те же источники-кандидаты. Для этого выполняют следующие действия: добавляют все источники-кандидаты, имеющиеся в каждом списке, в новый объединенный список, причем если источник-кандидат встречается более чем в одном списке, повторно его не добавляют; затем сохраняют полученный объединенный список. Далее суммируют все значения счетной переменной J, ассоциированные с каждым из найденных списков, и присваивают полученное значение J объединенному списку. После чего удаляют исходные списки, оставляя только полученный объединенный список.

[0082] Например, если в ходе выполнения шага (145) применительно к ранее показанной группе источников-кандидатов, имевшему значение  $J=1$ :

- sampleuser.livejournal.com
- website.com
- sample.newspaper.ru

один из этих источников будет найден в другом, сохраненном ранее списке, имевшем значение  $J=3$ , например,

- anotherwebsite.es
- sampleuser.livejournal.com
- justasite.co.il

то на шаге (146) эти два списка будут объединены в один список следующим образом:

- sampleuser.livejournal.com
- website.com
- sample.newspaper.ru
- anotherwebsite.es
- justasite.co.il.

и значение счетной переменной J, которое будет храниться ассоциированно с этим объединенным списком, будет рассчитано как сумма:

$$J=1+3=4.$$

[0083] В том частном случае, если на шаге (145) будет найден сохраненный ранее список, состоящий из тех же самых источников, что и список, созданный на шаге (144), то есть будут обнаружены два полностью идентичных списка, то суммируют значения счетной переменной J, ассоциированные с каждым из списков, один из списков удаляют, а полученное значение J присваивают оставшемуся списку.

[0084] На этом способ переходит к шагу (147), на котором сравнивают полученное на шаге (146) значение счетной переменной J с заранее заданным пороговым значением  $J_{max}$ . Это заранее заданное пороговое значение выбирают на этапе настройки системы, реализующей способ. Оно имеет смысл количества "групповых" публикаций, сделанных в разное время пересекающимися или совпадающими группами источников и может быть, например, равно 3.

[0085] В том случае, если оказывается, что значение счетной переменной J меньше или равно заданному пороговому значению  $J_{max}$ , то способ возвращается к шагу (141), на котором выбирают очередную публикацию.

[0086] Если значение счетной переменной J оказывается больше порогового значения  $J_{max}$ , то способ переходит к шагу (148), на котором относят все входящие в объединенный список источники публикаций к группе связанных источников публикаций. Иными словами, тот список, для которого выполняется  $J > J_{max}$ , считают списком, содержащим группу связанных источников публикаций. Для всех входящих

в него источников публикаций проставляют в базе данных соответствующий тэг, "Группа связанных источников", после чего способ возвращается к шагу (141), на котором выбирают очередную публикацию.

5 [0087] Список, для которого было выполнено J>Jmax, не удаляют, он по-прежнему обрабатывается в ходе шага (140) наряду со всеми остальными списками источников-кандидатов, как это было описано выше.

[0088] Шаг (140) выполняется циклически до тех пор, пока не будет достигнут конец списка публикаций, из которого выбирают публикации на шаге (141). На этом выполнение шага (140) завершается и способ переходит к шагу (150), как это было  
10 описано выше применительно к Фиг. 1А.

[0089] На шаге (150), как это будет описано далее применительно к Фиг. 1В, в составе найденных источников публикаций выявляют по меньшей мере следующие типы источников: рекламные площадки, агрегаторы отзывов пользователей (отзовики), биржи аккаунтов, биржи SMM-услуг и площадки для найма сотрудников на удаленную  
15 работу (биржи фрилансеров).

[0090] Шаг (150) начинается с поиска в сети интернет, выполняемого на шаге (151), в ходе которого находят веб-сайты, выполняющие функции рейтинговых агентств.

[0091] Этот поиск выполняется любым общеизвестным способом, посредством любой известной поисковой системы, такой как Google. В качестве ключевых слов  
20 используют заранее подготовленные наборы строк, позволяющие сформировать соответствующий поисковый запрос, например, такие как:

"рейтинг бирж SMM"

"рейтинг бирж аккаунтов"

"рейтинг бирж фрилансеров"

25 "рейтинг лучших отзовиков"

[0092] Затем, анализируя поисковую выдачу, что может быть выполнено любым общеизвестным образом, например, посредством заблаговременно подготовленного скрипта, извлекают гиперссылки (URL) на веб-сайты, выполняющие функции  
30 рейтинговых агентств, и сохраняют эти ссылки в виде списков, например, список рейтингов бирж SMM, список рейтингов бирж аккаунтов и т.д. Таким образом, в результате выполнения шага (151) получают упорядоченные по специфике деятельности веб-сайта списки ссылок на веб-сайты рейтинговых агентств.

[0093] На этом шаг (151) завершается и способ переходит к шагу (152), на котором сканируют найденные сайты рейтинговых агентств. Для этого используют списки URL,  
35 составленные на шаге (151). Сканирование выполняют любым общеизвестным способом, при помощи какой-либо программы, реализующей функции веб-парсера, то есть автоматического "сборщика" публикаций с различных веб-сайтов, такой, например, как CloudScrape или Scrapinghub.

[0094] В результате выполнения шага (152) получают и сохраняют в базе веб-страницы  
40 просканированных сайтов, которые содержат, помимо прочего, собственно рейтинги, то есть упорядоченные списки веб-сайтов, выполняющих функции бирж аккаунтов, бирж SMM-услуг, бирж фрилансеров, а также агрегаторов отзывов пользователей (отзовиков).

[0095] На этом шаг (152) завершается и способ переходит к шагу (153), на котором формируют списки веб-сайтов, выполняющих функции различных бирж, а также  
45 агрегаторов отзывов пользователей. Это выполняют любым общеизвестным способом, позволяющим извлечь из сохраненных на предыдущем шаге веб-страниц рейтинговых агентств ссылки (URL) на перечисленные в рейтингах сайты. Извлеченные ссылки

сохраняют в списках, формируя таким образом:

- список ссылок на биржи аккаунтов,
- список ссылок на биржи SMM-услуг,
- список ссылок на биржи фрилансеров,
- **список ссылок на отзывы (агрегаторы отзывов пользователей)** (1)

[0096] Сформированные таким образом списки сохраняют, и на этом способ переходит к шагу (154), на котором анализируют и очищают сформированные на предыдущем шаге списки. Для этого из вышеперечисленных списков удаляют повторные вхождения, то есть исключают повторяющиеся ссылки (URL). Кроме того, на данном шаге полученные ссылки усекают до домена второго уровня, таким образом, что URL вида

`https://example-otzovik.su/index.html`

оказывается преобразован в строку вида

`example-otzovik.su.`

[0097] Это может выполняться любым общеизвестным образом. В результате выполнения шага (154) получают и сохраняют в базе данных четыре списка источников, соответствующих списку (1).

[0098] Затем способ переходит к шагу (155), на котором сканируют найденные биржи аккаунтов. Для этого используют список URL бирж аккаунтов, составленный на шаге (154). Сканирование выполняют любым общеизвестным способом, при помощи какой-либо программы, реализующей функции веб-парсера, то есть автоматического "сборщика" публикаций с различных веб-сайтов, такой, например, как CloudScrape или Scrapinghub.

[0099] В результате выполнения шага (155) получают и сохраняют в базе веб-страницы просканированных сайтов, которые содержат, помимо прочего, списки предлагаемых к продаже или аренде аккаунтов.

[0100] Следует заметить, что аккаунты, продаваемые или предлагаемые к сдаче в аренду на биржах аккаунтов, заведомо управляются ботами. Поэтому на следующем шаге (156) анализируют сохраненные на шаге (155) веб-страницы и извлекают из них названия предлагаемых к продаже или аренде аккаунтов, из которых формируют список аккаунтов, управляемых ботами. Собственно анализ веб-страниц может выполняться любым общеизвестным способом, например, при помощи скрипта, осуществляющего парсинг (разбор) веб-страницы, извлечение из нее названий аккаунтов и сохранение их в отдельный список.

[0101] Список аккаунтов, управляемых ботами, перед завершением шага (156) используется для того, чтобы пометить известные аккаунты, полученные на шаге (120), тэгом "Бот". Тэги проставляются в используемой базе данных любым общеизвестным способом, в соответствии с используемой архитектурой базы данных.

[0102] Те аккаунты, которые присутствуют в списке, полученном на шаге (156), но отсутствуют в базе данных (то есть аккаунты, которые управляются ботами, но пока не были найдены или не были использованы) также сохраняют в базе данных с тэгом "Бот" и в дальнейшем используют наряду со всеми остальными известными аккаунтами.

[0103] Возможен также альтернативный вариант реализации описываемого способа, в котором шаг (156) пропускают, переходя от шага (155) к шагу (157).

[0104] Способ переходит к шагу (157), на котором сканируют найденные биржи фрилансеров. Для этого используют список URL бирж фрилансеров, составленный на шаге (154). Сканирование выполняют любым общеизвестным способом, при помощи

какой-либо программы, реализующей функции веб-парсера, то есть автоматического "сборщика" публикаций с различных веб-сайтов, такой, например, как CloudScrape или Scrapinghub.

5 [0105] В результате выполнения шага (157) получают и сохраняют в базе веб-страницы просканированных бирж фрилансеров, которые содержат, помимо прочего, тексты заданий фрилансерам на размещение отзывов заранее заданной направленности на страницах тех или иных веб-ресурсов.

10 [0106] Следует заметить, что веб-ресурсы, на которых фрилансерам предлагают размещать отзывы заранее заданной направленности, как правило, относятся к категории рекламных площадок, то есть представляют собой интернет-СМИ, публикующие, помимо обычных новостей, оплаченные публикации заранее заданной направленности.

15 [0107] Поэтому на следующем шаге (158) анализируют сохраненные на шаге (157) веб-страницы и извлекают из них ссылки (URL) на рекламные площадки, из которых формируют список рекламных площадок. Собственно анализ веб-страниц может выполняться любым общеизвестным способом, например, при помощи скрипта, осуществляющего парсинг (разбор) веб-страницы, извлечение из нее URL и сохранение их в отдельный список.

20 [0108] Затем способ переходит к шагу (159), на котором сформированный таким образом список анализируют и очищают. Для этого из списка удаляют повторные вхождения, то есть исключают повторяющиеся ссылки (URL). Кроме того, на данном шаге полученные ссылки усекают до домена второго уровня, таким образом, что URL вида

`https://reklamnoe-smi.ru/index.html`

25 оказывается преобразована в строку вида

`reklamnoe-smi.ru.`

30 [0109] На этом шаг (150) завершается. Результатом выполнения шагов (130), (140) и (150) становится рубрикация данных, сохраненных после выполнения шага (110), то есть отнесение по меньшей мере части найденных источников публикаций к по меньшей мере одному типу источников. Как уже было отмечено, источник публикаций может одновременно относиться к разным типам источников, источнику может быть проставлено более одного тэга.

35 [0110] В одном возможном варианте реализации способа для дальнейшей обработки данных используют только те источники, тип которых был определен на шагах (130), (140) и (150). В другом возможном варианте реализации для дальнейшей обработки данных используют все источники.

40 [0111] Затем способ переходит к шагу (160), на котором дополнительно к списку, полученному на шаге (156), выявляют среди найденных на шаге (120) аккаунтов те аккаунты, которые управляются ботами. Выполнение шага (160) будет подробно описано ниже со ссылкой на Фиг. 1Г.

45 [0112] Выполнение шага (160), как это показано на Фиг. 1Г, начинается на шаге (161), на котором отбирают среди найденных публикаций те, которые сделаны в социальных сетях. Поскольку ранее в результате выполнения этапа (130) была выполнена рубрикация источников публикаций, являющихся социальными сетями, выполнение шага (161) представляет собой отбор из базы данных публикаций, для которых в базе данных проставлена пометка "Социальная сеть". Технически подобный отбор может осуществляться любым общеизвестным образом, выбранным в зависимости от архитектуры используемой базы данных, например, отправкой соответствующего

SQL-запроса и получение ответа на него.

[0113] Следует заметить, что "публикациями в социальных сетях" в контексте выполнения данного шага (161) считают результаты таких действий, выполнение которых требует от пользователя социальной сети наибольших затрат времени. К ним относятся следующие публикации, типичные для современных социальных сетей:

- запись (оригинальное сообщение),
- комментарий (ответ на чью-то запись или комментарий),
- репост, то есть размещение от своего имени какой-либо записи, сделанной произвольным пользователем, с указанием аккаунта отправителя и ссылкой на оригинальную запись.

[0114] Такие альтернативные средства волеизъявления пользователей социальных сетей как эмодзи (смайлики) и лайки/дизлайки (голоса "за" и "против") в ходе выполнения данного шага не учитываются.

[0115] После получения публикаций в социальных сетях, способ переходит к шагу (162), на котором из общего массива публикаций отбирают все публикации, сделанные одним аккаунтом. Поскольку ранее, в ходе шага (120) были определены аккаунты, с которых сделаны все публикации, то технически шаг (162) представляет собой фильтрацию полученного массива публикаций по автору (аккаунту). Он может выполняться любым общеизвестным образом, выбранным в зависимости от архитектуры используемой базы данных. Собственно название аккаунта поочередно берут из общего хранящегося в базе данных списка аккаунтов, сформированного, как было описано ранее, на этапах (120) и (156).

[0116] Перед осуществлением фильтрации полученного массива публикаций (на Фиг.1Г этот шаг для простоты восприятия не показан) могут дополнительно проверять, имеет ли уже данный аккаунт тэг "Бот". Такой тэг мог быть проставлен ранее, в ходе выполнения шага (156). Если такой тэг присутствует, то шаг (163) не выполняют и переходят к следующему аккаунту из списка аккаунтов.

[0117] Затем на шаге (163) подсчитывают количество  $M$  публикаций, сделанных данным аккаунтом с заданным интервалом, например, с интервалом в 1 секунду или менее. Для этого публикации любым общеизвестным способом упорядочивают по дате и времени публикации, после рассчитывают временные интервалы между каждыми двумя соседними по времени публикациями. Например, если аккаунт сделал публикации с условными обозначениями П1, П2, П3 и П4, то будут рассчитаны интервалы между публикациями П1 и П2, между П2 и П3, а также между П3 и П4. Затем подсчитывают количество  $M$  интервалов, длительность которых меньше или равна заранее выбранному значению, например, меньше или равна 1 секунде.

[0118] Затем способ переходит к шагу (164), на котором сравнивают подсчитанное значение  $M$  с заранее заданным порогом. Этот порог может быть выбран эмпирически на этапе настройки системы и равен, например 4. В случае, если значение  $M$  для анализируемого аккаунта превышает этот заранее заданный порог, то способ переходит к шагу (167), на котором относят данный аккаунт к тем аккаунтам, которые управляются ботами, и затем возвращается к шагу (162).

[0119] Если на шаге (164) значение  $M$  для анализируемого аккаунта оказывается меньше заранее заданный порог, то способ переходит к шагу (165), на котором подсчитывают период времени  $T$ , в течение которого аккаунт делал публикации с частотой не реже заданной частоты  $F$ . Величина  $F$  при этом может быть выбрана заранее, на этапе настройки системы. Например,  $F$  может быть выбрана равной одной публикации в час или одной публикации в два часа.

[0120] Например, если аккаунт сделал публикации с условными обозначениями П1, П2 ... П400, то на шаге (164) они будут упорядочены по дате и времени публикации, после чего будет рассчитаны временные интервалы между каждыми двумя соседними по времени публикациями: между П1 и П2, между П2 и П3, и так далее, до интервала между П499 и П400. Затем находят периоды времени Т1, Т2, Т3 и т.д., такие, что внутри каждого такого периода частота публикаций превышает заранее заданную величину F. Иными словами, находят все периоды времени, на протяжении которых данный аккаунт размещал публикации чаще, чем с заданной частотой F. Это может быть сделано любым общеизвестным образом.

[0121] Затем определяют продолжительность периода времени Т как максимальную продолжительность периода среди всех найденных периодов времени Т1, Т2, Т3 и т.д.

[0122] После чего способ переходит к шагу (166), на котором определяют, превышает ли продолжительность периода времени Т заранее заданный порог. Например, этот порог может быть выбран равным 36 часам или 48 часам. Иначе говоря, на данном этапе проверяется, как долго с данного аккаунта размещали какие-либо публикации непрерывно, без перерыва на то время, которое необходимо человеку для сна.

[0123] В том случае, если Т превышает заранее заданный порог, то способ переходит к шагу (167), на котором относят данный аккаунт к тем аккаунтам, которые управляются ботами, то есть проставляют для него в базе данных тэг "Бот", и затем возвращается к шагу (162). В противном случае, если Т не превышает заранее заданный порог, способ возвращается к шагу (162).

[0124] В целях упрощения блок-схемы на Фиг. 1Г условно не показана проверка условия "достигнут конец списка аккаунтов?", которая может выполняться каждый раз перед шагом (162), на котором выбирают очередной аккаунт для анализа. При выполнении этого условия выполнение шага (160) завершается и способ возвращается к шагу (170), как это было описано выше применительно к Фиг. 1А.

[0125] Следует заметить, что шаги (110), (120), (130), (140), (150) и (160) для простоты описания показаны применительно к предварительному этапу как простая последовательность. Однако, возможен вариант реализации системы, в котором эти шаги выполняются не один раз, а циклически, в том числе и параллельно выполнению шагов, которые будут описаны далее применительно к рабочему этапу. Данная деятельность может осуществляться непрерывно, что позволит постоянно пополнять базы данных и в любой момент времени иметь в базе данных "свежие", актуальные сведения.

[0126] На заключительном шаге (170) подготовительного этапа (100) сохраняют все сведения, полученные на предыдущих шагах, в базе данных. Это может выполняться любым общеизвестным образом. На этом подготовительный этап (100) завершается.

[0127] Для реализации описываемого способа выявления атак на репутацию после завершения предварительного этапа (100) выполняют рабочий этап (200), как это описано ниже со ссылкой на Фиг. 2А.

[0128] Предварительный этап (200) начинается с шага (210), на котором получают по меньшей мере одно слово или словосочетание, характеризующие цель атаки на репутацию. Это может быть выполнено любым общеизвестным способом. Например, заранее подготовленная в соответствии с принятым системой форматом строка, содержащая слова и словосочетания, характеризующая цель атаки на репутацию, может поступать в систему, реализующую способ, из базы данных, где хранятся задания для системы, реализующей способ, по окончании выполнения системой предыдущего задания.

[0129] Возможны также, без ограничений, любые альтернативные варианты реализации данного шага, в том числе импорт слов и словосочетаний из текста электронного письма, направленного на заранее организованный адрес электронной почты, ассоциированный с системой, реализующей описываемый способ, и т.д.

5 [0130] Затем способ переходит к шагу (220). Перед началом шага (220) любым общеизвестным образом получают и сохраняют в базе данных показания системных часов, т.е. текущее время на момент начала выполнения шага. Затем сканируют интернет и находят веб-страницы, содержащие полученные слова и словосочетания. Технически это может выполняться любым общеизвестным способом, например, аналогично тому,  
10 как это было описано выше для шага (110). Затем способ переходит к шагу (230).

[0131] На шаге (230) найденные веб-страницы анализируют и выделяют по меньшей мере: заголовок, аккаунт автора, дату и время, источник публикации, текст публикации; это выполняют аналогично описанному выше для шага (120). Извлеченную информацию сохраняют в базе данных.

15 [0132] На этом шаг (230) завершается и способ переходит к шагу (240), на котором из найденных текстов публикаций извлекают ссылки (URL) и формируют списки ссылок. Это выполняют любым общеизвестным способом, позволяющим извлечь из сохраненных на предыдущем шаге публикаций все ссылки (URL), например, посредством заблаговременно подготовленного скрипта, находящего в теле каждой публикации  
20 такие сочетания символов, как http, https и www, и извлекающие всю строку, начинающуюся с этих символов и заканчивающуюся пробелом или символами "возврат каретки" или "перенос строки".

[0133] Извлеченные таким образом ссылки сохраняют, например, в базе данных, и способ переходит к шагу (250). На шаге (250) анализируют публикации, найденные на  
25 шаге (220) и ссылки, извлеченные на шаге (240) и подсчитывают значения количественных характеристик и динамику их изменения.

[0134] К количественным характеристикам при этом относят по меньшей мере: (2)

- общее количество публикаций N,
- количество публикаций, сделанных ботами, Nb,
- 30 • количество публикаций, сделанных на агрегаторах компромата, Nk,
- количество публикаций, сделанных группами связанных источников публикаций, Ng
- количество публикаций, сделанных группами связанных источников, которые также являются агрегаторами компромата, Ngk
- 35 • количество публикаций, сделанных на рекламных площадках, Nr,
- количество публикаций, сделанных на рекламных площадках, входящих в группу связанных источников, Ngr,
- количество публикаций, сделанных на агрегаторах отзывов пользователей, No,
- количество публикаций, сделанных на агрегаторах утечек, Nu,
- 40 • количество публикаций, сделанных на площадках для найма сотрудников на удаленную работу, Nh,
- общее количество публикаций, являющихся дублями друг друга, Nd,
- общее количество публикаций на агрегаторах компромата, являющихся дублями друг друга, Ndk,
- 45 • общее количество публикаций на агрегаторах компромата, являющихся дублями друг друга и сделанных ботами, Ndbk,
- общее количество ссылок, являющихся дублями друг друга, Nld
- количество аккаунтов, с которых были размещены найденные публикации, Na

- количество аккаунтов, управляемых ботами, с которых были размещены найденные публикации, Nab

- количество аккаунтов, с которых были размещены публикации, найденные на агрегаторах компромата, Nak

5     • количество аккаунтов, управляемых ботами, с которых были размещены публикации на агрегаторах компромата, Nabk

- количество аккаунтов, с которых были размещены публикации, найденные на рекламных площадках, Nar.

10    [0135] Под динамикой изменения названных величин в данном случае понимают значения этих величин, вычисленные на протяжении заранее заданного интервала времени  $t$  с заранее заданным шагом (временным интервалом между итерациями)  $ts$ . В качестве неограничивающего примера, интервал  $t$  может быть задан равным 10 минутам, а шаг  $ts$  -- равным 1 минуте.

15    [0136] Конкретные способы вычисления названных величин будут подробнее описаны ниже, со ссылками на Фиг. 2Б, Фиг. 2 В.

[0137] Несложно видеть, что вышеназванные количественные характеристики можно условно объединить в три основные группы: характеристики, имеющие смысл количества тех или иных публикаций, характеристики, имеющие смысл количества дублей (повторов) и характеристики, имеющие смысл количества аккаунтов.

20    [0138] Характеристики, имеющие смысл количества публикаций вычисляют, как это показано на Фиг. 2Б. Поскольку в ходе предварительного этапа, а именно, на шагах (130), (140), (150), (160) различные источники публикаций и аккаунты были размечены, то есть по меньшей мере для некоторых из них в базе данных были проставлены пометки, такие как "Агрегатор компромата", "Группа связанных источников", "Бот" и так далее, 25 извлечение из базы данных количества публикаций, относящихся к тем или иным источникам, технически реализуется как поиск в базе данных записей с соответствующей пометкой (тэгом).

30    [0139] На первом этапе (251) выбирают по меньшей мере один критерий (пометку, тэг) для фильтрации. Ее выбирают из заранее подготовленного списка пометок, например, поочередно выбирая одну пометку за другой.

[0140] Затем способ переходит к этапу (252), на котором строят запрос к базе данных, содержащий выбранную пометку и получают из базы данных список публикаций, соответствующий этому запросу, и на этапе (253) получают оценку длины этого списка, то есть количества публикаций. Затем на этапе (254) собственно полученную оценку 35 сохраняют; опционально при этом могут также сохранять и сам полученный список.

[0141] В качестве примера можно подробнее описать подсчет количества публикаций Nb, сделанных ботами. В ходе предварительного этапа, а именно шагов (156) и (160), были определены аккаунты, управляемые ботами, и для каждого из таких аккаунтов в базе данных была проставлена пометка "Бот".

40    [0142] На этапе (251) из списка пометок получают пометку "Бот"; затем на этапе (252) к базе данных строят запрос и из базы данных получают список публикаций, найденных в ходе шага (220), которые притом были сделаны с аккаунтов, имеющих пометку "Бот". Это может быть сделано любым общеизвестным образом, в зависимости от архитектуры используемой базы данных, например, отправкой соответствующего 45 SQL-запроса.

[0143] Затем определяют длину списка, то есть количество полученных таким образом публикаций. Оно и будет количеством публикаций Nb, сделанных ботами. Его сохраняют в базе данных; дополнительно может быть сохранен и собственно список публикаций.

[0144] В другом примере, для вычисления общего количества найденных публикаций N на этапе (252) запрос к базе данных могут не строить, притом принимать N равным общему количеству веб-страниц, сохраненных на текущей итерации шага (220).

Например, на первой итерации способа на шаге (220) может быть найдено 100 веб-страниц, и в базе данных будет сохранено значение  $N=100$ . На второй итерации способа количество найденных страниц может стать равным 110, и в базе данных будет сохранено значение  $N=110$ . На третьей итерации способа количество найденных страниц может стать равным, например, 130, и в базе данных будет сохранено значение  $N=130$ .

[0145] Следует отметить, что значения всех количественных характеристик, вычисляемых на этапе (250) сохраняют в базе данных в виде вектора, то есть последовательности чисел. Например, в результате описанных выше итераций для характеристики N будет сохранена следующая последовательность значений:  
 $N=(100, 110, 130)$ .

[0146] Еще в одном примере, для определения количества публикаций, сделанных группами связанных источников, которые также являются агрегаторами компромата (Ngk), используют два тэга: "Агрегатор компромата" и "Группа связанных источников". При построении запроса к базе данных эти тэги объединяют логическим И, таким образом получая список публикаций, где для каждого из источников ранее, на шагах (130) и (140) были проставлены обе эти пометки.

[0147] Затем определяют длину списка, то есть количество полученных таким образом публикаций. Оно и будет количеством публикаций Ngk, сделанных группами связанных источников, которые также являются агрегаторами компромата. Его сохраняют в базе данных; дополнительно может быть сохранен и собственно список публикаций.

[0148] Количественные характеристики, имеющие смысл количества дублей (повторов), рассчитывают в два этапа. На первом этапе из базы данных получают то множество записей, внутри которого необходимо найти дубли. Например, чтобы вычислить общее количество публикаций, являющихся дублями друг друга и размещенных на агрегаторах компромата (Ndk), получают список публикаций, размещенных на агрегаторах компромата.

[0149] В данном примере могут использовать список, полученный при вычислении оценки количества публикаций, сделанных на агрегаторах компромата (Nk) и сохраненный на этапе (254). В другом примере, для вычисления общего количества ссылок, являющихся дублями друг друга (Nld) могут использовать данные, полученные в ходе шага (240), на котором были извлечены и сохранены имеющиеся в найденных публикациях ссылки (URL). В этом случае любым общеизвестным способом строят запрос к базе данных и получают список ссылок, найденных на шаге (240).

[0150] На втором этапе определяют количество дубликатов внутри полученного списка. Для вычисления общего количества публикаций, являющихся дублями друг друга и размещенных на агрегаторах компромата (Ndk), это может выполняться полностью аналогично описанному ранее шагу (142). Для вычисления количества дубликатов в списке ссылок может использоваться аналогичный алгоритм, с той лишь разницей, что поиск в базе данных ведут не по полю "Публикация", а по полю "Гиперссылка".

[0151] Как показано на Фиг. 2 В, вычисление характеристик, имеющих смысл количества аккаунтов, начинают с этапа (251), на котором выбирают по меньшей мере один критерий (пометку, тэг) для фильтрации, например, пометку "Бот". Ее выбирают из заранее подготовленного списка пометок.

[0152] Затем способ переходит к этапу (255), на котором строят запрос к базе данных,

содержащий выбранную пометку и получают из базы данных список аккаунтов, соответствующий этому запросу.

[0153] Затем на этапе (256) полученный список фильтруют, любым общеизвестным образом исключая из него повторы. После чего на этапе (257) получают оценку длины этого списка, то есть количества аккаунтов, соответствующих заданному критерию. Затем на этапе (258) собственно полученную оценку сохраняют; опционально при этом могут также сохранять и список аккаунтов.

[0154] Скажем, для определения количества аккаунтов, управляемых ботами, с которых размещены публикации на агрегаторах компромата (Nak) из списка публикаций, сделанных на агрегаторах компромата. Указанный список публикаций при этом мог быть получен ранее, как это было описано применительно к этапам (251)...(254), либо построен заново, путем запроса из базы данных всех публикаций, сделанных источниками, имеющими пометку "Агрегатор компромата". Затем из этого списка извлекают список аккаунтов, с которых они были сделаны, притом имеющих пометку "Бот". Список аккаунтов любым общеизвестным образом фильтруют, удаляя из него повторы и оставляя в списке по одному вхождению каждого аккаунта. Длину полученного после такой фильтрации списка принимают за искомое количество аккаунтов Nak и сохраняют его.

[0155] Возможен вариант реализации, при котором этап (251) пропускают. Так, чтобы определить общее количество аккаунтов, с которых были размещены найденные публикации (Na), из базы данных извлекают полный перечень аккаунтов, с которых были сделаны публикации. Затем из этого списка любым общеизвестным образом удаляют повторы, то есть оставляют в нем по одному вхождению каждого аккаунта. Количество строк полученного списка считают количеством аккаунтов, с которых были размещены найденные публикации Na, и сохраняют его.

[0156] Таким образом, возвращаясь к Фиг. 2А, на шаге (250) вычисляют значения перечисленных характеристик и сохраняют их в базе данных, после чего любым общеизвестным способом, например, сравнивая показания системных часов в момент начала шага (220) и в текущий момент, вычисляют время  $T_r$ , фактически прошедшее с начала данной итерации, затем вычисляют оценку времени, прошедшего с начала этапа (210):

$$T_i = T_i + T_r,$$

[0157] Затем способ переходит к шагу (260), на котором проверяют, достигнута ли заранее заданная величина интервала времени  $t$ , сравнивая  $t$  и  $T_i$ . В том случае, если

$$T_i < t,$$

то есть заданный временной интервал еще не достигнут, выдерживают паузу  $dT$ , численно равную разности заранее заданной величиной шага (интервала между итерациями)  $t_s$  и времени  $T_r$ , фактически прошедшего с начала данной итерации:

$$dT = t_s - T_r,$$

после чего способ возвращается к этапу (220), на котором сканируют интернет и находят вебстраницы, содержащие полученное на этапе (210) по меньшей мере одно слово или словосочетание, характеризующее цель атаки на репутацию.

[0158] В том случае, если

$$T_i > t,$$

то есть заданный временной интервал достигнут, способ переходит к шагу (270).

[0159] На шаге (270) вычисляют на основании подсчитанных величин параметры,

характеризующие вероятность наличия атаки на репутацию.

[0160] Как было упомянуто ранее, значения всех количественных характеристик, вычисляемых на шаге (250) сохраняют в базе данных в виде векторов, то есть последовательностей чисел. Например, в результате выполнения шагов (220)...(260) на протяжении заданного временного интервала  $t$  были вычислены пять значений для каждой из численных характеристик, названных в списке (2):

- $N = (N1, N2, N3, N4, N5),$  (3)
- $Nb = (Nb1, Nb2, Nb3, Nb4, Nb5),$
- 10 •  $Nk = (Nk1, Nk2, Nk3, Nk4, Nk5),$
- $Ng = (Ng1, Ng2, Ng3, Ng4, Ng5),$
- $Ngk = (Ngk1, Ngk2, Ngk3, Ngk4, Ngk5),$
- 15 •  $Nr = (Nr1, Nr2, Nr3, Nr4, Nr5),$
- $Ngr = (Ngr1, Ngr2, Ngr3, Ngr4, Ngr5),$
- $No = (No1, No2, No3, No4, No5),$
- $Nu = (Nu1, Nu2, Nu3, Nu4, Nu5),$
- 20 •  $Nh = (Nh1, Nh2, Nh3, Nh4, Nh5),$
- $Nd = (Nd1, Nd2, Nd3, Nd4, Nd5),$
- $Ndk = (Ndk1, Ndk2, Ndk3, Ndk4, Ndk5),$
- 25 •  $Ndbk = (Ndbk1, Ndbk2, Ndbk3, Ndbk4, Ndbk5),$
- $Nld = (Nld1, Nld2, Nld3, Nld4, Nld5),$
- $Na = (Na1, Na2, Na3, Na4, Na5),$
- $Nab = (Nab1, Nab2, Nab3, Nab4, Nab5),$
- 30 •  $Nak = (Nak1, Nak2, Nak3, Nak4, Nak5),$
- $Nabk = (Nabk1, Nabk2, Nabk3, Nabk4, Nabk5),$
- $Nar = (Nar1, Nar2, Nar3, Nar4, Nar5),$

35 [0161] На шаге (270) в каждой последовательности, показанной в списке (3), вычисляют абсолютную  $D$  (в единицах) и относительную  $Dr$  (в процентах) разность между соседними значениями. Например, для последовательности, вычисленной для общего количества публикаций  $N$ :

$$N = (N1, N2, N3, N4, N5),$$

40 в данном примере будут вычислены:

45

$$D1 = N2 - N1;$$

$$Dr1 = 100 * (N2 - N1) / N1,$$

$$D2 = N3 - N2;$$

$$Dr2 = 100 * (N3 - N2) / N2,$$

$$D3 = N4 - N3;$$

$$Dr3 = 100 * (N4 - N3) / N3,$$

$$D4 = N5 - N4;$$

$$Dr4 = 100 * (N5 - N4) / N4.$$

[0162] После вычисления всех значений абсолютной  $D$  и относительной  $Dr$  разности для каждой последовательности чисел (3), полученных для количественных характеристик (2), шаг (270) завершается и способ переходит к шагу (280).

[0163] На шаге (280) определяют, превышает ли по меньшей мере одно из значений  $D$  и  $Dr$  заранее заданное для него пороговое значение.

[0164] Например, для численной характеристики  $N_{dk}$ , имеющей смысл общего количества публикаций на агрегаторах компромата, являющихся дублями друг друга, может быть задано пороговое значение 7 для абсолютной разности  $D$ , и пороговое значение 5% для относительной разности  $Dr$ .

[0165] В то же время, для численной характеристики  $N_e$ , имеющей смысл количества публикаций, сделанных на рекламных площадках, может быть задано пороговое значение 3 для абсолютной разности  $D$ , и пороговое значение 6% для относительной разности  $Dr$ .

[0166] Притом для численной характеристики  $N_d$ , имеющего смысл общего количества публикаций, являющихся дублями друг друга, может быть задано пороговое значение 95 для абсолютной разности  $D$ , и пороговое значение 20% для относительной разности  $Dr$ .

[0167] Иными словами, для каждой из названных (2) количественных характеристик могут быть заданы соответствующие ей пороговые значения для относительной и абсолютной разности.

[0168] Сами эти значения могут быть подобраны эмпирически на этапе настройки системы.

[0169] Если ни одно из значений  $D$  и  $Dr$  соответствующее ему пороговое значение не превышает, способ возвращается к этапу (220), на котором сканируют интернет и находят вебстраницы, содержащие полученное на этапе (210) по меньшей мере одно слово или словосочетание, характеризующее цель атаки на репутацию.

[0170] В другой возможной реализации описываемого способа в этом случае способ (200) завершается.

[0171] Еще в одной возможной реализации описываемого способа (не показано на Фиг. 2А) система, реализующая способ (200), формирует сообщение о том, что атака на репутацию по заданной цели не обнаружена и переходит к ожиданию дальнейших команд пользователя, например, ввода новых слов и/или словосочетаний, характеризующих цель атаки на репутацию.

[0172] В том случае, если на шаге (280) определяют, что по меньшей мере одно из значений  $D$  и  $Dr$  превышает заранее заданное для него пороговое значение, то способ переходит к шагу (290).

[0173] На шаге (290) вычисляют, на основании подсчитанных величин, а именно

значений абсолютной  $D$  и относительной  $D_r$  разности, для разных количественных параметров, оценок способа атаки и характера атаки. Кроме того, на данном этапе формируют и отправляют оповещение об атаке на репутацию, а также способе и характере ее осуществления.

5 [0174] Неограничивающий пример способа вычисления (300) оценок способа атаки и характера атаки будет описан ниже со ссылкой на Фиг. 3.

[0175] Следует заметить, что алгоритм, показанный на Фиг. 3, в показанном виде использован только для простоты иллюстрации общего принципа; показанные на Фиг. 3 две

10 характеристики  $N_d$  (общее количество публикаций, являющихся дублями друг друга) и  $N_{ld}$  (общее количество ссылок, являющихся дублями друг друга) также приведены для простоты иллюстрации и не ограничивают способ (300).

[0176] В реализации способа могут быть использованы все количественные характеристики, приведенные в списке (2). Кроме того, описываемый способ может также включать любые другие, кроме показанных на Фиг. 3, логические зависимости между перечисленными в списке (2) характеристиками и быть реализован с учетом любых наперед заданных соотношений между численными значениями приведенных в списке (2) количественных характеристик.

[0177] Аналогично, показанные на Фиг. 3 способы атаки, условно названные "Посев" и "Разгон", не составляют исчерпывающего перечня возможных способов атаки на репутацию, и приведены исключительно для примера. Описанный способ позволяет идентифицировать и выявить, без ограничений, любые известные специалистам в данной предметной области способы атаки на репутацию.

[0178] Способ (300) начинается на этапе (310), на котором определяют, к каким 25 количественным характеристикам из перечисленных в списке (2) относятся значения абсолютной  $D$  и/или относительной  $D_r$  разности, превысившие заранее заданный порог.

[0179] Например, если порог превысила величина  $N_{ld}$ , соответствующая общему количеству ссылок, являющихся дублями друг друга (320), то на этапе (340) атаке присваивают тип "Разгон". (Так может быть назван тип атаки, заключающейся в 30 распространении по большому количеству веб-площадок одной и той же гиперссылки, ведущей на один материал, служащий для воздействия на целевую аудиторию).

[0180] Затем способ переходит к этапу (360), на котором определяют, в зависимости от того, какая из величин  $D$  и  $D_r$  превысила порог, уровень атаки. В данном случае, если заданный порог превысило значение абсолютной разности  $D$ , то способ переходит к этапу (397), на котором атаке присваивают уровень "Предупреждение". В противном случае, если заданный порог превысило значение относительной разности  $D_r$ , то способ переходит к этапу (398), на котором атаке присваивают уровень "Угроза". После этого способ завершается.

[0181] Если же на этапе (310) определяют, что порог превысила величина  $N_d$ , соответствующая общему количеству публикаций, являющихся дублями друг друга 40 (330), то на следующем этапе (350) атаке присваивают тип "Посев". (Это тип атаки, смысл которой в распространении по большому количеству веб-площадок одного и того же текста, содержимое которого призвано воздействовать на целевую аудиторию).

[0182] Затем способ переходит к этапу (370), на котором определяют, в зависимости от того, какая из величин  $D$  и  $D_r$  превысила порог, уровень атаки. В данном случае, если заданный порог превысило значение абсолютной разности  $D$ , то способ переходит к этапу (398), на котором атаке присваивают уровень "Угроза". В противном случае, если заданный порог превысило значение относительной разности  $D_r$ , то способ

переходит к этапу (399), на котором атаке присваивают наивысший уровень "Атака". После этого способ завершается.

[0183] Важно, что выбор на этапе (310) не является бинарным, как для простоты восприятия показано на Фиг. 3. На этом этапе может быть выбрано любое количество характеристик, из числа перечисленных в списке (2), соответствующие которым значения D и/или D<sub>r</sub> превысили порог. Если оказались выбраны две, три или более характеристик, то последовательности действий, соответствующие этапам (320) и (330), выполняются одновременно.

[0184] Соответственно, такой атаке может быть присвоено несколько типов; применительно к Фиг. 3, например, атака может относиться одновременно к типам "Посев" и "Разгон".

[0185] Аналогично, возможна ситуация, когда атаке присваивают несколько разных уровней; применительно к Фиг. 3, например, могут быть присвоены уровни "Предупреждение" и "Атака". В такой ситуации система, реализующая описываемый способ, выбирает наивысший из присвоенных уровней, и использует его при формировании оповещения об атаке.

[0186] Как следует из Фиг. 3, возможна реализация описываемого способа, при которой оповещение об атаке на репутацию, которое является результатом работы описываемой системы, может иметь один из трех уровней важности: "Предупреждение", "Угроза", "Атака". Указанные уровни важности указывают на уровень интенсивности атаки.

[0187] В другой возможной реализации (не показана на Фиг. 3) оповещение об атаке на репутацию может иметь численное выражение, характеризующее уровень интенсивности атаки, например, "Зафиксирована атака на [название цели атаки] с интенсивностью I=71%". При этом данное число I может быть получено, например, путем нормирования значений абсолютной D или относительной D<sub>r</sub> разности какой-либо из характеристик, перечисленных в списке (2) к максимальному значению, найденному за заданный временной интервал t:

$$P = 100% * (D/D_{max});$$

[0188] или любым другим способом, опирающимся на численные значения перечисленных в списке (2) количественных характеристик, например, на вычисленные для каждого из них значения среднего арифметического за заданный временной интервал t, и т.д.

[0189] Формирование и отправка оповещения может выполняться по меньшей мере одним из перечисленных способов: по электронной почте, посредством отправки SMS, посредством отправки MMS, посредством отправки push-уведомления, сообщением в программе обмена мгновенными сообщениями, посредством создания события API.

[0190] Следует отметить, что использование такого средства оповещения, как события API, позволяет реализовать дополнительную интеграцию описываемой системы с различными сторонними инструментами, такими как платформы мониторинга общественного мнения, платформы управления безопасностью, SIEM-решения и так далее. Собственно формирование всех перечисленных оповещений, таких как электронные письма, SMS, MMS, push-уведомления и т.д. может быть выполнено любым общеизвестным образом.

[0191] На этом описываемый способ завершается.

[0192] Еще в одной возможной реализации описываемого способа (не показано на Фиг. 2А) система, реализующая способ (200), после формирования и отправки оповещения переходит к ожиданию дальнейших команд пользователя, например, ввода

новых слов и/или словосочетаний, характеризующих цель атаки на репутацию.

[0193] Еще в одной возможной реализации описываемого способа (не показано на Фиг. 2А) система, реализующая способ (200), после формирования и отправки оповещения возвращается к шагу (220) и продолжает работу по описанному выше

5 алгоритму.

[0194] На Фиг. 4 представлена пример общей схемы вычислительного устройства (400), обеспечивающего обработку данных, необходимую для реализации заявленного решения.

[0195] В общем случае устройство (400) содержит такие компоненты как один или более процессоров (401), по меньшей мере одно оперативное запоминающее устройство или память (402), средство хранения данных (403), интерфейсы ввода/вывода (404), средство В/В (405), средства сетевого взаимодействия или, что то же самое, передачи данных (406).

[0196] Процессор (401) устройства выполняет основные вычислительные операции, необходимые для функционирования устройства (400) или функциональности одного или более его компонентов. Процессор (401) исполняет необходимые машиночитаемые команды, содержащиеся в оперативной памяти (402).

[0197] Память (402), как правило, выполнена в виде ОЗУ и содержит необходимую программную логику, обеспечивающую требуемую функциональность.

[0198] Средство хранения данных (403) может выполняться в виде HDD, SSD дисков, рейд массива, сетевого хранилища, флэш-памяти, оптических накопителей информации (CD, DVD, MD, Blue-Ray дисков) и т.п.

[0199] Интерфейсы (404) представляют собой стандартные средства для подключения и работы с серверной частью, например, USB, RS232, RJ45, LPT, COM, HDMI, PS/2, Lightning, Fire Wire и т.п. Выбор интерфейсов (404) зависит от конкретного исполнения устройства (400), которое может представлять собой персональный компьютер, мейнфрейм, серверный кластер, тонкий клиент, смартфон, ноутбук и т.п.

[0200] В качестве средств В/В данных (405) могут использоваться клавиатура, джойстик, дисплей (сенсорный дисплей), проектор, тачпад, манипулятор мышь, трекбол, световое перо, динамики, микрофон и т.п.

[0201] Средства сетевого взаимодействия (406) выбираются из устройств, обеспечивающих прием и передачу данных по сети, например, Ethernet-карта, WLAN/Wi-Fi модуль, Bluetooth модуль, BLE модуль, NFC модуль, IrDa, RFID модуль, GSM модем и т.п. С помощью средств (406) обеспечивается организация обмена данными по проводному или беспроводному каналу передачи данных, например, WAN, PAN, ЛВС (LAN), Интранет, Интернет, WLAN, WMAN или GSM.

[0202] Компоненты устройства (400) сопряжены посредством общей шины передачи данных (410).

[0203] В заключение следует отметить, что приведенные в описании сведения являются только примерами, которые не ограничивают объем настоящего изобретения, описанного формулой. Специалисту в данной области становится понятным, что могут существовать и другие варианты осуществления настоящего изобретения, согласующегося с сущностью и объемом настоящего изобретения.

[0204] Примерные системы и способы, проиллюстрированные в данном документе, могут описываться с точки зрения компонентов функциональных блоков. Следует принимать во внимание, что такие функциональные блоки могут быть реализованы посредством любого числа аппаратных и/или программных компонентов, сконфигурированных с возможностью выполнять указанные функции. Например,

система может использовать различные компоненты интегральной схемы, например, запоминающие элементы, элементы обработки, логические элементы, таблицы поиска и т.п., которые могут выполнять множество функций под управлением одного или более микропроцессоров либо других устройств управления. Аналогично, программные элементы системы могут реализовываться с помощью любого языка программирования или подготовки сценариев, такого как C, C++, C#, Java, JavaScript, VBScript, Macromedia Cold Fusion, COBOL, Microsoft Active Server Pages, язык ассемблера, Perl, PHP, AWK, Python, Visual Basic, хранимых процедур SQL, PL/SQL, любых сценариев оболочки UNIX и расширяемого языка разметки (XML), при этом различные алгоритмы реализуются с любой комбинацией структур данных, объектов, процессов, процедур или других программных элементов.

[0205] Кроме того, система выявления атаки на репутацию может работать на одном вычислительном устройстве, либо на нескольких, связанных между собой по сети. Дополнительно следует отметить, что система может использовать любое число традиционных технологий для передачи данных, передачи служебных сигналов, обработки данных, управления сетью и т.п.

[0206] В данном контексте под устройствами понимаются любые вычислительные устройства, построенные на базе программно-аппаратных средств, например, такие как: персональные компьютеры, серверы, смартфоны, ноутбуки, планшеты и т.д.

[0207] В качестве устройства обработки данных может выступать процессор, микропроцессор, ЭВМ (электронно-вычислительная машина), ПЛК (программируемый логический контроллер) или интегральная схема, сконфигурированные для исполнения определенных команд (инструкций, программ) по обработке данных. Процессор может быть многоядерным, для параллельной обработки данных.

[0208] В роли устройства памяти могут выступать, но не ограничиваясь, жесткие диски (HDD), флеш-память, ПЗУ (постоянное запоминающее устройство), твердотельные накопители (SSD) и т.д.

[0209] Следует отметить, что в указанное устройство могут входить и любые другие известные в данном уровне техники устройства, например, такие как датчики, устройства ввода/вывода данных, устройства отображения (дисплеи) и т.п. Устройство ввода/вывода данных может представлять собой, но не ограничиваясь, например, манипулятор мышь, клавиатуру, тачпад, стилус, джойстик, трекпад и т.п.

[0210] В настоящих материалах заявки было представлено предпочтительное раскрытие осуществление заявленного технического решения, которое не должно использоваться как ограничивающее иные, частные воплощения его реализации, которые не выходят за рамки испрашиваемого объема правовой охраны и являются очевидными для специалистов в соответствующей области техники.

#### (57) Формула изобретения

1. Способ выявления информационной атаки, выполняемый вычислительным устройством и содержащий шаги, на которых:

на предварительном этапе:

- сканируют сеть Интернет и находят источники публикаций,

- выявляют в составе найденных источников публикаций источники, используемые

для информационных атак,

- находят аккаунты, с которых размещались записи в выявленных источниках публикаций, используемых для информационных атак,

- выявляют среди найденных аккаунтов те, которые управляются ботами,

- сохраняют полученные сведения об источниках, используемых для информационных атак, и управляемых ботами аккаунтах в базе данных;

на рабочем этапе:

- 5 - получают слова и словосочетания, характеризующие цель информационной атаки,
- сканируют интернет и находят публикации, содержащие слова и словосочетания, характеризующие цель информационной атаки,
- извлекают из найденных публикаций гиперссылки,
- подсчитывают, используя сведения об источниках, используемых для информационной атаки, и управляемых ботами аккаунтах, количественные
- 10 характеристики публикаций и динамику их изменения,
- вычисляют на основании подсчитанных количественных характеристик параметры, характеризующие вероятность наличия информационной атаки, и в ответ на превышение по меньшей мере одним вычисленным параметром заранее заданного порогового значения
- 15 - определяют, на основании вычисленных параметров, тип атаки и уровень атаки,
- формируют и отправляют оповещение об информационной атаке, а также о типе атаки и уровне атаки.

2. Способ по п. 1, отличающийся тем, что к источникам публикаций, используемым для информационной атаки, относятся по меньшей мере следующие:

- 20 - агрегаторы компромата,
- социальные сети, агрегаторы утечек данных,
- рекламные площадки,
- группы связанных источников,
- агрегаторы отзывов пользователей,
- 25 - площадки для найма сотрудников на удаленную работу.

3. Способ по п. 2, отличающийся тем, что к группам связанных источников относят группы источников, не менее заданного количества раз разместивших идентичные публикации с разницей во времени публикации, не превышающей заранее заданное пороговое значение.

- 30 4. Способ по п. 1, отличающийся тем, что к аккаунтам, которые управляются ботами, относят аккаунты, сделавшие за заранее заданный промежуток времени не менее заранее заданного количества публикаций.

- 5. Способ по п. 4, отличающийся тем, что к аккаунтам, которые управляются ботами, также относят аккаунты, делавшие публикации на протяжении заданного промежутка
- 35 времени с частотой, превышающей заданное пороговое значение.

6. Способ по п. 1, отличающийся тем, что к количественным характеристикам публикаций относят по меньшей мере следующие величины:

- общее количество публикаций,
- количество публикаций, сделанных ботами,
- 40 - количество публикаций, сделанных на агрегаторах компромата,
- количество публикаций, сделанных группами связанных источников публикаций,
- количество публикаций, сделанных группами связанных источников, которые также являются агрегаторами компромата,
- количество публикаций, сделанных на рекламных площадках,
- 45 - количество публикаций, сделанных на рекламных площадках, входящих в группу связанных источников,
- количество публикаций, сделанных на агрегаторах отзывов пользователей,
- количество публикаций, сделанных на агрегаторах утечек,

- количество публикаций, сделанных на площадках для найма сотрудников на удаленную работу,
- общее количество публикаций, являющихся дублями друг друга,
- общее количество публикаций на агрегаторах компромата, являющихся дублями друг друга,
- общее количество публикаций на агрегаторах компромата, являющихся дублями друг друга и сделанных ботами,
- общее количество ссылок, являющихся дублями друг друга,
- количество аккаунтов, с которых были размещены найденные публикации,
- количество аккаунтов, управляемых ботами, с которых были размещены найденные публикации,
- количество аккаунтов, с которых были размещены публикации, найденные на агрегаторах компромата,
- количество аккаунтов, управляемых ботами, с которых были размещены публикации на агрегаторах компромата,
- количество аккаунтов, с которых были размещены публикации, найденные на рекламных площадках.

7. Способ по п. 1, отличающийся тем, что динамику изменения количественных характеристик вычисляют на основании значения этих характеристик, вычисленных на протяжении заранее заданного интервала времени с заранее заданным шагом.

8. Способ по п. 1, отличающийся тем, что параметры, характеризующие вероятность наличия информационной атаки, для каждой количественной характеристики вычисляют как абсолютную, выраженную в единицах, и относительную, выраженную в процентах, разность между соседними значениями данной характеристики.

9. Способ по п. 1, отличающийся тем, что оповещение об информационной атаке передают посредством по меньшей мере одного из следующих способов коммуникации:

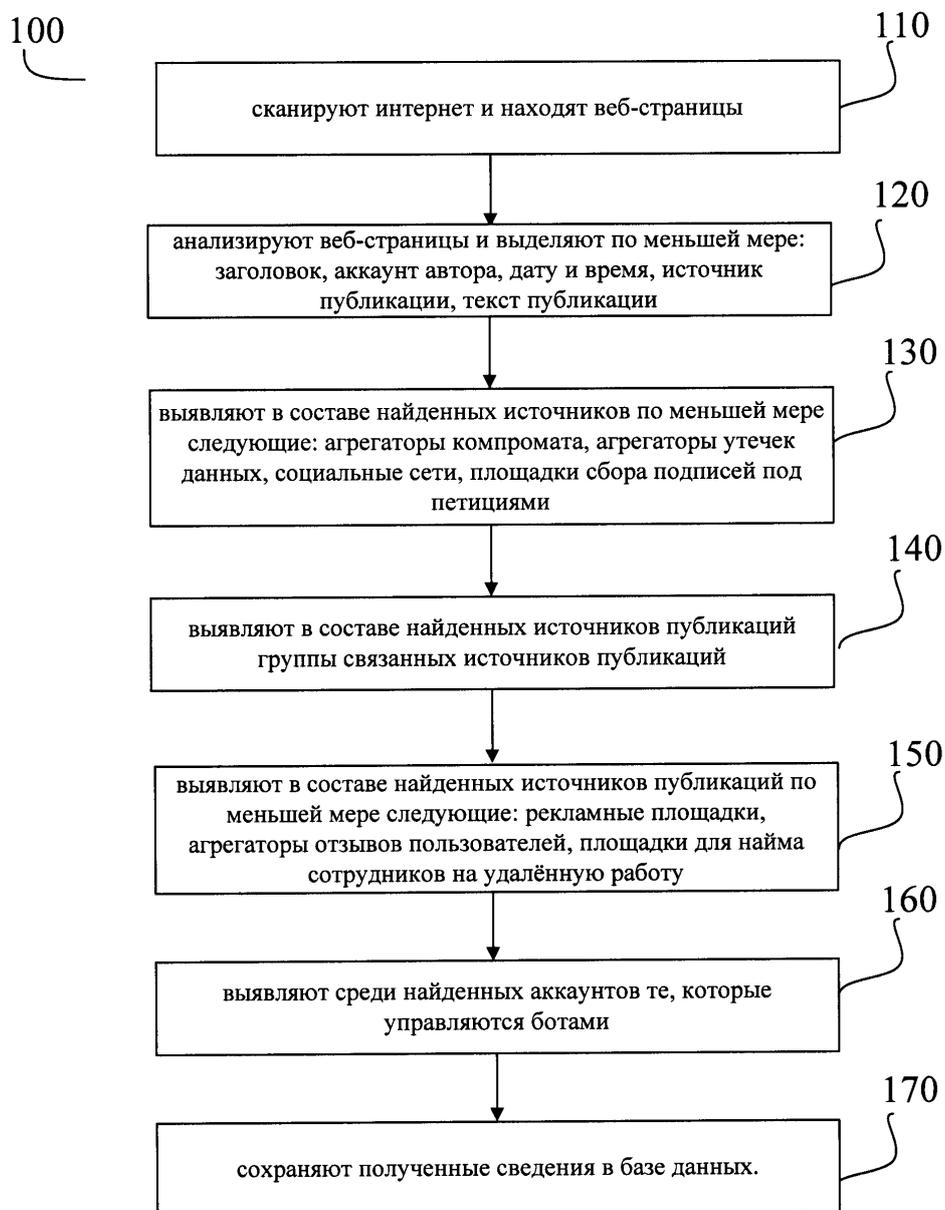
- электронной почты (e-mail),
- SMS,
- MMS,
- push-уведомления,
- сообщения в программе обмена мгновенными сообщениями,
- события API

10. Способ по п. 1, отличающийся тем, что оповещение об информационной атаке может иметь численное выражение, характеризующее уровень интенсивности атаки.

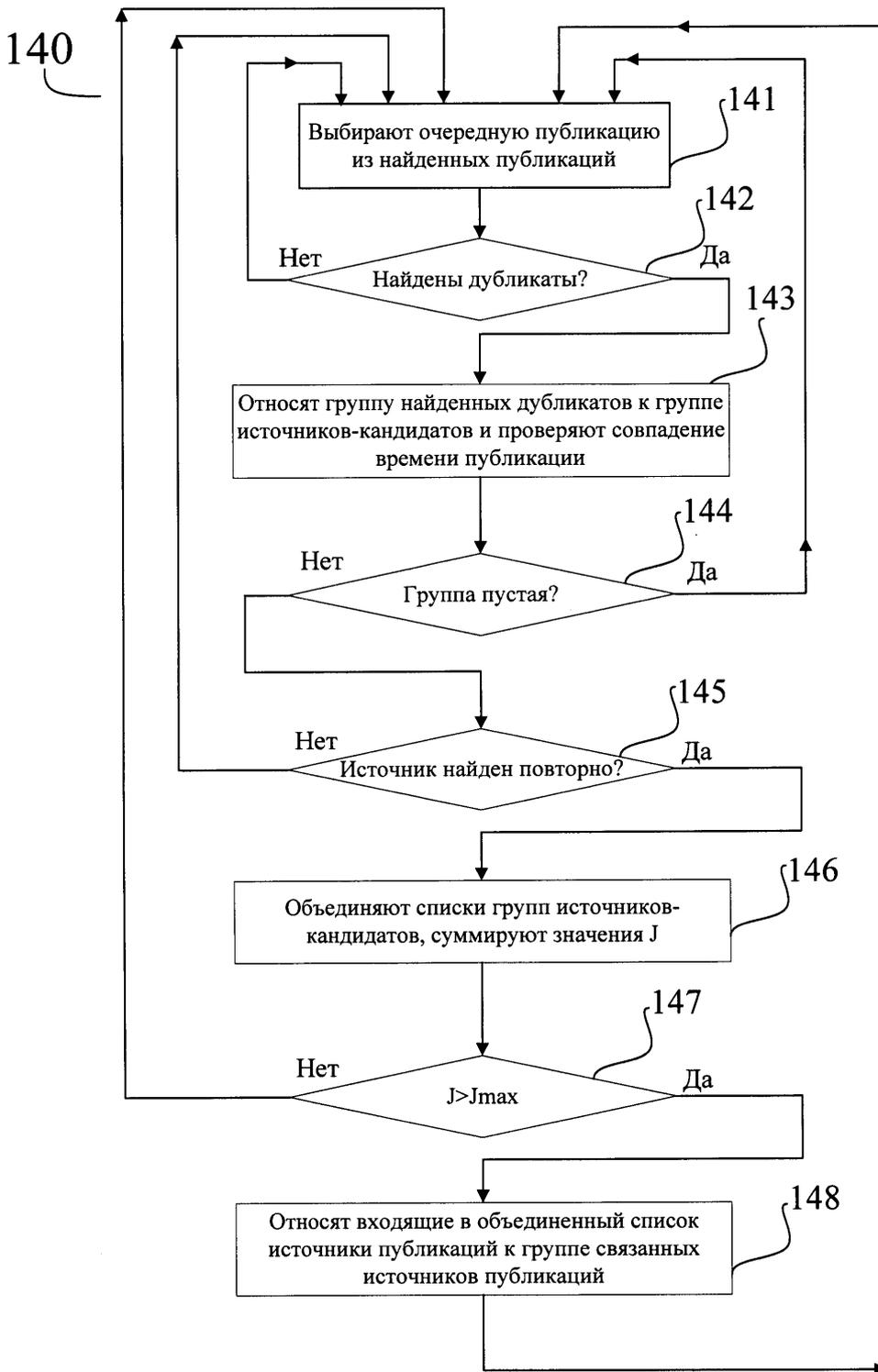
11. Способ по п. 1, отличающийся тем, что оповещение об информационной атаке может иметь один из трех уровней: "Предупреждение", "Угроза", "Атака".

12. Система выявления информационной атаки, выполненная с возможностью сканировать сеть Интернет и содержащая по меньшей мере:

- процессор,
- запоминающее устройство, содержащее:
  - по меньшей мере одну базу данных,
  - машиночитаемые команды, которые при исполнении их процессором обеспечивают выполнение способа по пп. 1-11.



Фиг.1А

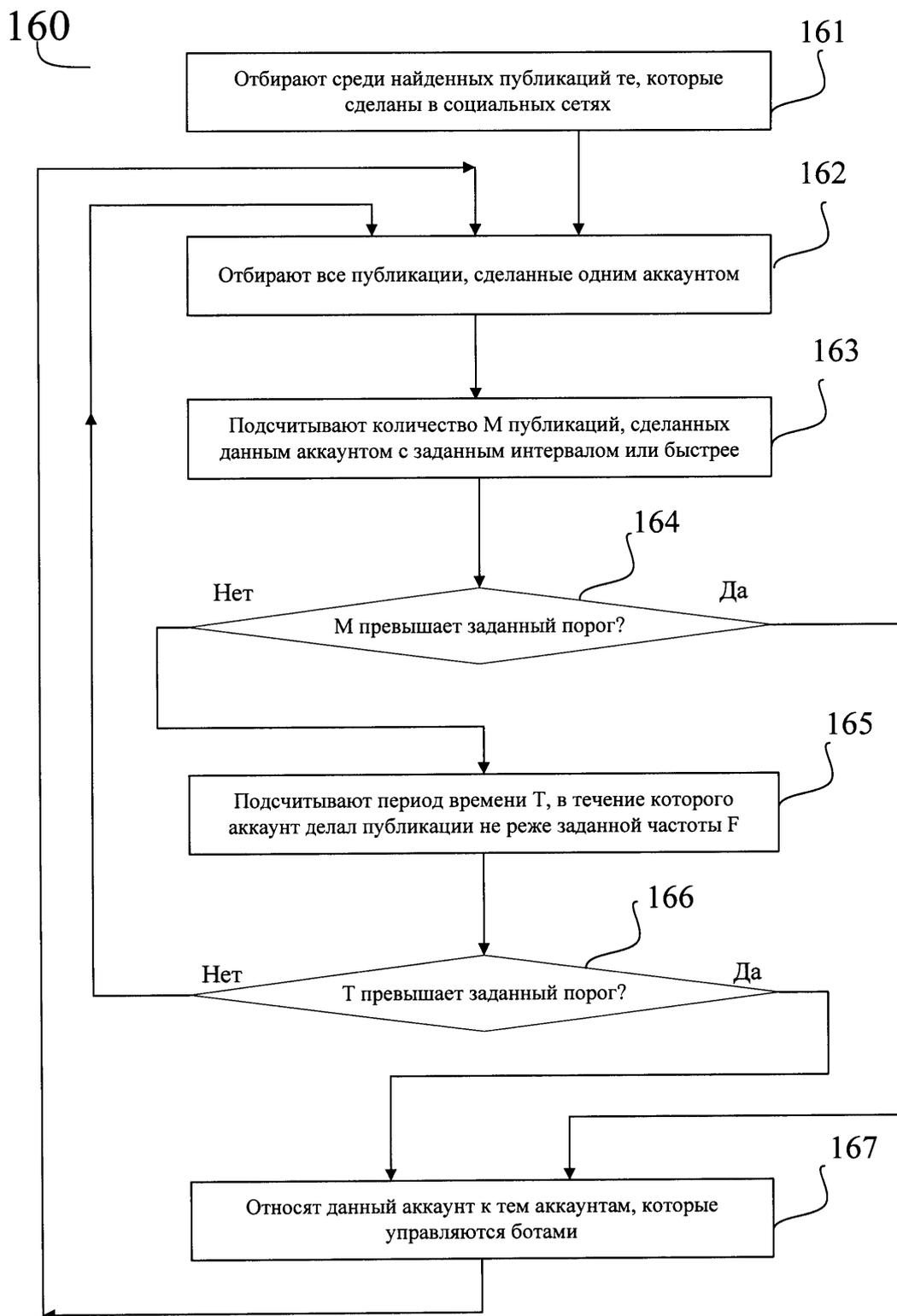


Фиг.1Б

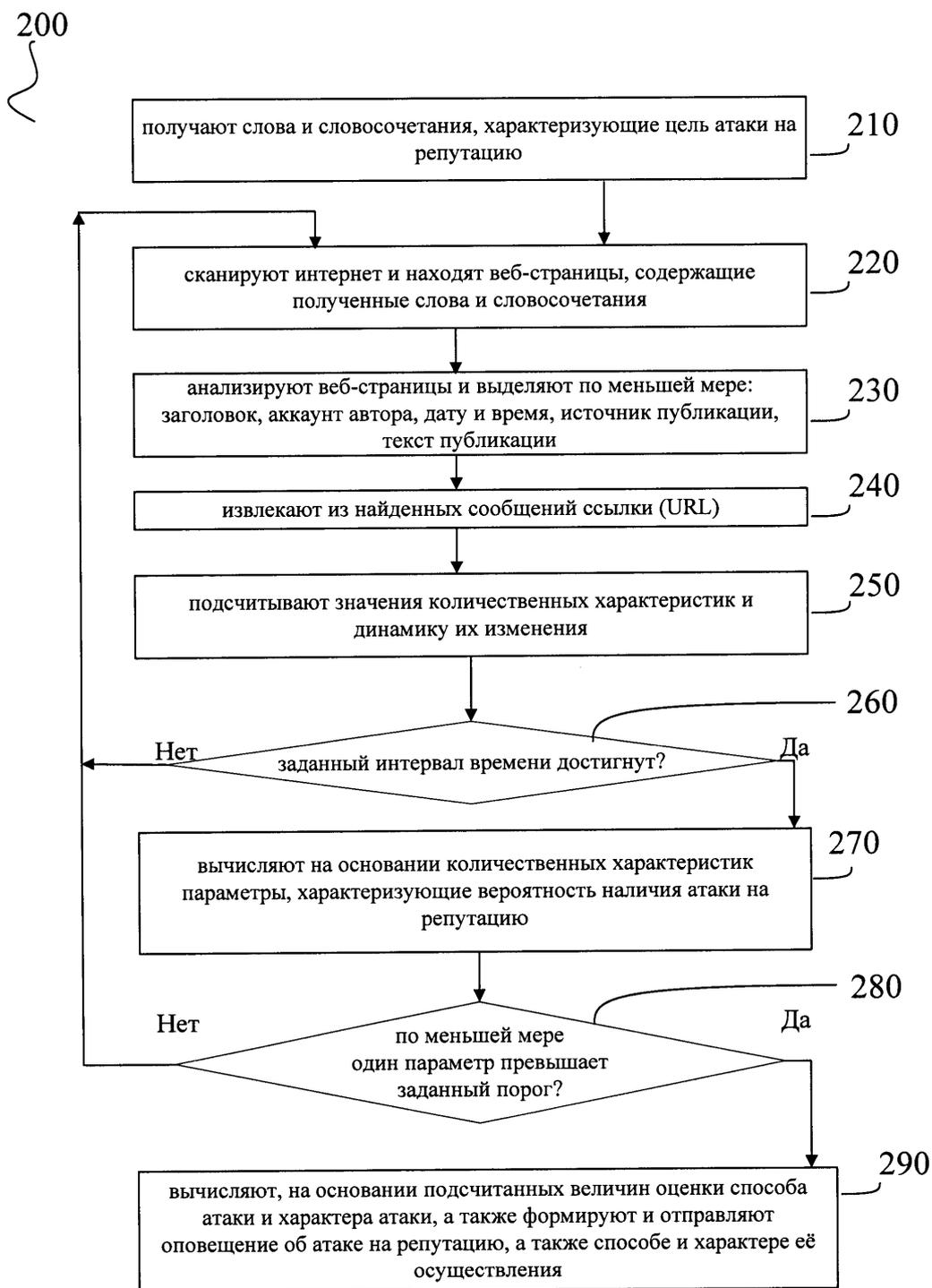
150



Фиг. 1В



Фиг.1Г



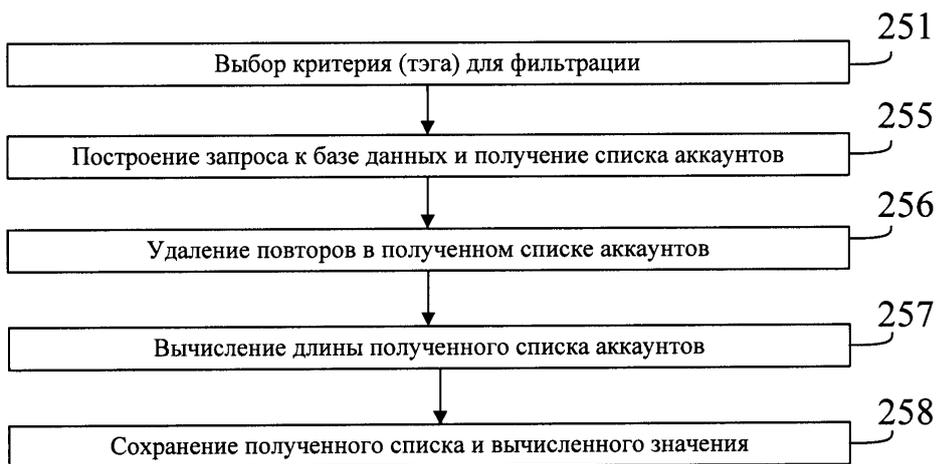
Фиг.2А

250

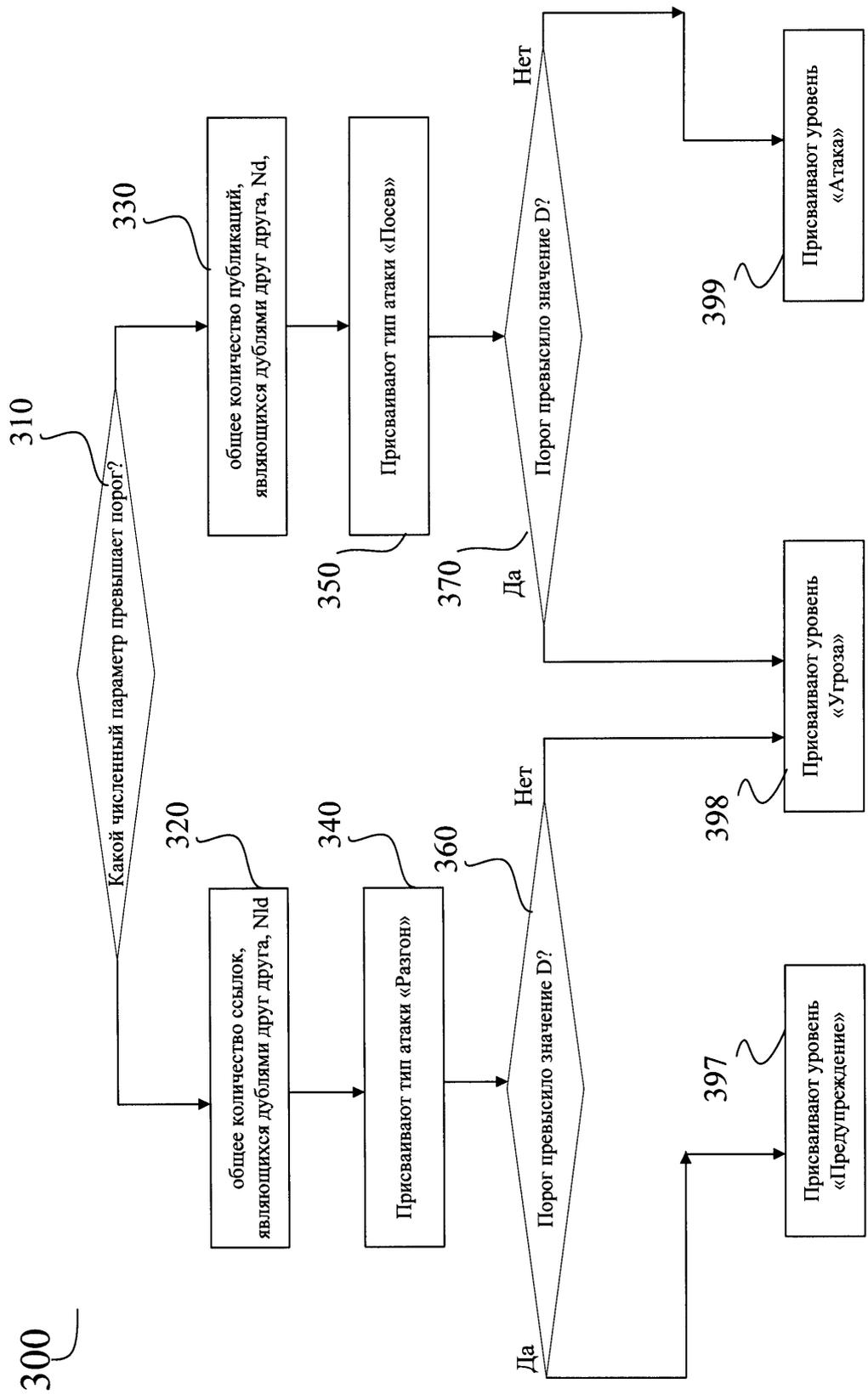


Фиг.2Б

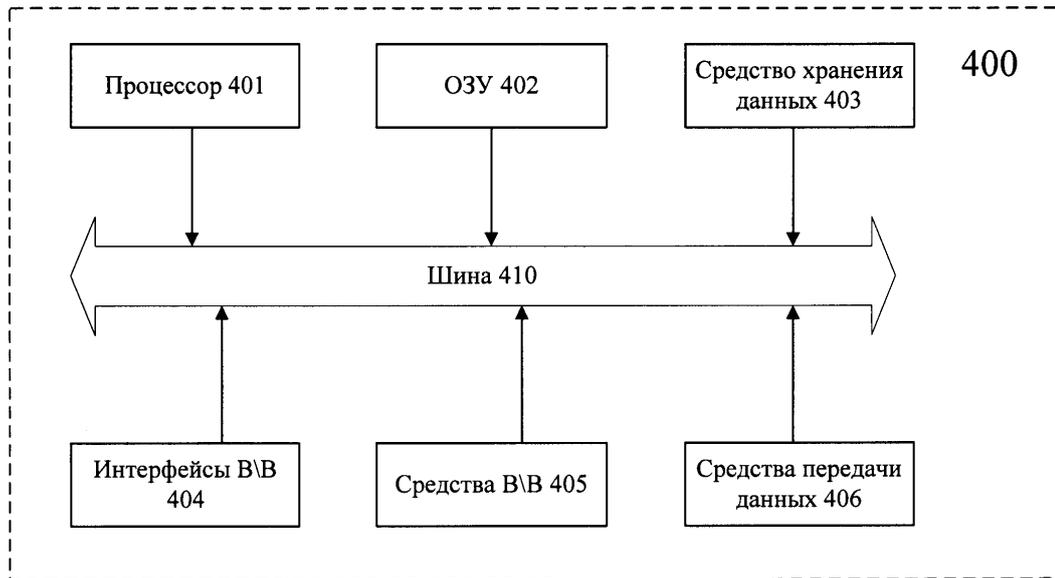
250



Фиг.2В



Фиг.3



Фиг.4