

(19)



**Евразийское
патентное
ведомство**

(21) **202293485** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2023.09.26

(22) Дата подачи заявки
2022.12.27

(51) Int. Cl. **G06F 21/16** (2013.01)
G06F 21/00 (2013.01)
G06F 21/88 (2013.01)
G06V 30/224 (2022.01)
G06V 30/244 (2022.01)
G06V 30/20 (2022.01)
G06F 18/00 (2023.01)
G06F 18/20 (2023.01)

(54) **СПОСОБ И СИСТЕМА ЗАЩИТЫ ИНФОРМАЦИИ ОТ УТЕЧКИ ПРИ ПЕЧАТИ ДОКУМЕНТОВ С ПОМОЩЬЮ ВНЕДРЕНИЯ ЦИФРОВЫХ МЕТОК**

(31) **2022106206**

(32) **2022.03.10**

(33) **RU**

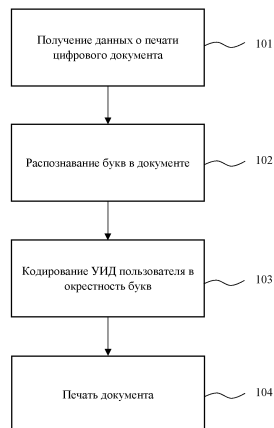
(71) Заявитель:
**ПУБЛИЧНОЕ АКЦИОНЕРНОЕ
ОБЩЕСТВО "СБЕРБАНК
РОССИИ" (ПАО СБЕРБАНК) (RU)**

(72) Изобретатель:
**Анистратенко Михаил Артурович,
Анистратенко Александр Артурович,
Оболенский Иван Александрович,
Борисов Дмитрий Алексеевич,
Сысоев Валентин Валерьевич (RU)**

(74) Представитель:
Герасин Б.В. (RU)

(57) Изобретение относится к области защиты информации, в частности к решениям для предотвращения утечки информации при печати документов. Технический результат заключается в повышении эффективности защиты данных от утечки за счет внедрения цифровых меток в документ, кодирующих уникальный идентификатор пользователя, для последующей его идентификации при анализе распечатанных документов. Заявленный результат достигается за счет способа кодирования информации для защиты от ее утечек при печати документов, выполняемого с помощью процессора компьютерного устройства, при этом способ содержит этапы, на которых: получают на компьютерном устройстве пользователя информацию о печати по меньшей мере одного цифрового документа, содержащего по меньшей мере текст, при этом компьютерное устройство связано с уникальным идентификатором (УИД) пользователя; осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой распознают буквы, содержащиеся в цифровом документе; кодируют УИД пользователя в набор цифровых меток, которые располагаются на контурах букв и/или вблизи контуров букв цифрового документа; передают цифровой документ на печать с закодированным УИД пользователя.

100



A1

202293485

202293485

A1

СПОСОБ И СИСТЕМА ЗАЩИТЫ ИНФОРМАЦИИ ОТ УТЕЧКИ ПРИ ПЕЧАТИ ДОКУМЕНТОВ С ПОМОЩЬЮ ВНЕДРЕНИЯ ЦИФРОВЫХ МЕТОК

ОБЛАСТЬ ТЕХНИКИ

[0001] Заявленное решение относится к области защиты информации, в частности к решениям для предотвращения утечки информации при печати документов.

УРОВЕНЬ ТЕХНИКИ

[0002] Технологии предотвращения утечек (англ. Data Leak Prevention, DLP) представляют собой технологии предотвращения утечек конфиденциальной информации из информационной системы вовне, а также технические устройства (программные или программно-аппаратные) для такого предотвращения утечек.

[0003] Из патентной заявки US 20080091954 A1 (Morris et al., 17.04.2008) известно решение для проверки целостности данных, представленных на печатных документах. Решение базируется на применении уникального идентификатора, с помощью которого осуществляется анализ содержимого документа. Каждому сегменту документа присваивается цифра или группа цифр, и каждой странице или сегменту документа может быть присвоена одна цифра в общем идентификаторе. Совокупность цифр, связанных с документом, объединяется в строку аутентификации. При получении запроса на последующую обработку документа выполняется аутентификация и проверка целостности документа путем считывания представленного документа для получения строки аутентификации, и последующего сравнения новой строки с ранее сохраненной строкой. После успешного сопоставления документ считается действительным, аутентифицированным и неизменным.

[0004] Недостатком данного решения является невозможность его использования для предотвращения утечек с целью идентификацию сотрудника, допустившего факт утечки при печати документов. Также, другим недостатком является недостаточная эффективность защиты документов, что обусловлено применением кода для сравнения аутентичности документа, что позволяет только установить факт неизменности и подлинности документа, но не предотвратить утечку информации.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[0005] Заявленное изобретение направлено на решение технической проблемы, заключающейся в создании эффективного средства для защиты цифровой информации от утечки при ее печати.

[0006] Технический результат заключается в повышении эффективности защиты данных от утечки, за счет внедрения цифровых меток в документ, кодирующих уникальный идентификатор пользователя, для последующей его идентификации при анализе распечатанных документов.

[0007] Заявленный результат достигается за счет способа кодирования информации для защиты от ее утечек при печати документов, выполняемого с помощью процессора компьютерного устройства, при этом способ содержит этапы, на которых:

получают на компьютерном устройстве пользователя информацию о печати по меньшей мере одного цифрового документа, содержащего по меньшей мере текст, при этом компьютерное устройство связано с уникальным идентификатором (УИД) пользователя;

осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой

распознают буквы, содержащиеся в цифровом документе;

кодируют УИД пользователя в набор цифровых меток, которые располагаются на контурах букв и/или вблизи контуров букв цифрового документа;

передают цифровой документ на печать с закодированным УИД пользователя.

[0008] В одном из частных примеров реализации способа распознавание цифрового документа выполняется с помощью оптического распознавания символов (OCR).

[0009] В другом частном примере реализации способа распознаются все символы на каждой странице цифрового документа.

[0010] В другом частном примере реализации способа каждый символ УИД пользователя кодируется в двоичный код.

[0011] В другом частном примере реализации способа на основании разряда двоичного кода определяется область размещения цифровых меток.

[0012] Заявленный технический результат также достигается за счет осуществления способа защиты информации от утечек на печатных документах, выполняемого с помощью процессора компьютерного устройства, при этом способ содержит этапы, на которых:

получают по меньшей мере часть изображения печатного документа с закодированным УИД пользователя вышеуказанным способом;
выполняют распознавание полученного изображения;
определяют буквы, содержащие цифровые метки в своей окрестности;
выполняют определение и извлечение закодированного УИД.

[0013] В одном из частных примеров выполнения способа распознавание цифрового документа выполняется с помощью OCR.

[0014] Заявленное решение также осуществляется с помощью соответствующих систем, содержащих процессор и память, которые хранят машиночитаемые инструкции для реализации каждого из вышеописанных способов.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[0015] Фиг. 1 иллюстрирует блок-схему способа кодирования цифровой метки.

[0016] Фиг. 2А – 2В иллюстрируют примеры размещения цифровых меток в цифровом документе.

[0017] Фиг. 3 иллюстрирует блок-схему декодирования цифровых меток.

[0018] Фиг. 4 иллюстрирует диаграмму час раскрытия позиций УИД.

[0019] Фиг. 5 иллюстрирует общий вид вычислительного устройства.

ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ

[0020] На Фиг. 1 представлен способ (100) защиты информации в цифровых документах от утечки с помощью кодирования УИД пользователя в виде цифровых меток в документ. На первом этапе (101) получается информация о печати цифрового документа. Выполнение способа (100) осуществляется на компьютерном устройстве пользователя, например, сотрудника, при этом к устройству привязан УИД пользователя, позволяющий его идентифицировать. Исполнение этапа (101) одушевляется с помощью программной логики, исполняемой компьютерным устройством и может быть реализовано, например, в виде программного агента или модуля, обеспечивающего получение сигналов от процессора, свидетельствующих об отправке цифрового документа на печать. Цифровой документ представляет собой, как правило, файл и может содержать текст, графику или их сочетания.

[0021] После получения на устройстве команды на перехват и анализ документа до его отправки на принтер на этапе (102) выполняется распознавание упомянутого цифрового документа. Обработка документа выполняется с помощью технологии OCR для обеспечения распознавания букв и символов в цифровом документе.

[0022] После этапа распознавания цифрового документа на этапе (103) осуществляется процесс кодирования УИД. УИД представляет собой, например, числовой табельный номер сотрудника - цифровой код TAB , состоящий, например, из 8-ми цифр. Данный код можно представить как массив цифр $TAB_8 = \{n_1, n_2, \dots, n_m\}$, $TAB_8 \in [0 \dots 9]$, $m = 8$. Схематичный вид кода представлен в таблице 1.

Таблица 1. Схематичное изображение табельного номера:

TAB_8								
Цифра	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8
Позиция	1	2	3	4	5	6	7	8

[0023] Каждый элемент табельного номера представляет собой число от 0 до 9, соответственно, каждый элемент табельного номера можно отобразить в двоичном виде размерностью в 4 бит, т.е. он будет представлять собой двоичное число от 1 до 1100, являющееся гомоморфизмом со сдвигом, представленным в таблице 2.

Таблица 2. Схема гомоморфизма табельного номера из десятичной в двоичную систему счисления.

TAB_8^{DEC}	TAB_8^{BIN}
0	0001
1	0010
2	0011
3	0100
4	0101
5	0110
6	0111
7	1000
8	1001
9	1010

[0024] Отображение 0 в 0001 необходимо для того, чтобы фиксировать наличие 0 в табельном номере. Для кодирования элемента табельного номера в двоичном коде $TAB_8^{BIN} = \{b_1, b_2, b_3, \dots, b_i\}$, $i = 8$, необходимо 4 разряда $b_i = \{c_1, c_2, c_3, c_4\}$, пример которых представлен в Таблице 3.

Таблица 3. Схематическое деление бинарного числа на разряды.

b_i				
Сигнал	c_1	c_2	c_3	c_4
Разряд	I	II	III	IV

[0025] Таким образом, возможно кодировать любое число в букву посредством двоичного кодирования. Пример такого разделения для последующего кодирования представлено на Фиг. 2А – Фиг. 2В. Каждая распознанная буква (20) делится на 4 четверти в плоскости по часовой стрелке, начиная с левого нижнего угла.

[0026] При наличии 1-цы в I разряде двоичного представления цифры табельного номера s_1 метка размещается в I четверти. Аналогичные операции проводятся со всеми разрядами двоичного представления цифры.

[0027] Метод нанесения метки в пространство возле буквы заключается в том, что как по казано на Фиг. 2Б-2В наносится цифровая метка в виде линии (21) на поверхности буквы или точки (22) в окрестности буквы в заданной четверти.

[0028] Пример кодирования меток в буквы представлен в Таблице 4.

Таблица 4. Схема позиционного кодирования

Позиция	Значение буквы				
	а	з	п	ч	
1	а	з	п	ч	
2	б	и	р	ш	
3	в	й	с	щ	
4	г	к	т	ь	
5	д	л	у	ы	
6	е	м	ф	ь	
7	ё	н	х	э	я
8	ж	о	ц	ю	

[0029] Выше представленная таблица 4 означает, что каждую позицию числа в табельном номере возможно кодировать на любую из 4-х букв. Выбор букв для нанесения метки осуществляется постранично. Пусть документ D содержит l страниц, тогда документ D – есть массив страниц, $D = \{p_1, p_2, p_3 \dots p_l\}, l \in N$.

[0030] На каждой странице $p_i, i \in [1, l]$ считывается посимвольно текст и записывается в массив символов $S_{p_i} = \{s_1, s_2, s_3, \dots s_{l_{p_i}}\}$, где l_{p_i} - количество символов на странице p_i , из них выявляются русские буквы $Wrus_{p_i} \in S_{p_i}$.

[0031] Далее создаются 8 массивов $Pos_1, Pos_2 \dots Pos_8$, каждый из которых соответствует каждой позиции табельного номера. Каждый массив Pos заполняется теми символами из $Wrus_{p_i}$, которые соответствуют позиции из таблицы 4. Например, Pos_1 заполняется всеми символами из $Wrus_{p_i}$, которые имеют значения {а, з, п, ч}, вне зависимости от регистра.

[0032] Массивы $Pos_1, Pos_2 \dots Pos_8$ перемешиваются, к примеру, тасованием Кнута. Пусть $l_{Pos_1}, l_{Pos_2}, l_{Pos_3} \dots l_{Pos_8}$ – размерности полученных массивов, P – процент символов на внедрение метки $P \in [0,3 \dots 0,7]$, тогда каждый массив из $Pos_1, Pos_2 \dots Pos_8$ обрезается с конца до размерности $\sum_{i=0}^8 \frac{Pos_i}{8} \cdot P, Pos_1, Pos_2 \dots Pos_8 \rightarrow Pos_1^P, Pos_2^P \dots Pos_8^P$.

[0033] Полученные массивы $Pos_1^P, Pos_2^P \dots Pos_8^P$ используются для нанесения цифровых меток вышеописанным способом. Внесение цифровых меток осуществляется с помощью вырезания букв с помощью OCR, внесения меток в пиксельные координаты и внесение букв с цифровыми метками обратно в документ, направляемый на печать. После внедрения всех меток (21, 22) на искомой странице p_i тоже самое выполняется для следующей страницы p_{i+1} и так далее до конца документа p_l .

[0034] В Таблице 5 приведен пример кодирования меток для УИД пользователя - 00013400.

Таблица 5. Пример кодирования цифровых меток в окрестности букв.

Позиция	Значение буквы				
1	а	з	ц	ч	
2	б	и	р	ш	
3	в	й	с	щ	
4	г	к	т	ь	
5	д	л	у	ы	
6	е	м	ф	ь	
7	ё	н	х	э	я
8	ж	о	ц	ю	

[0035] После внесения в документ, направленный на печать цифровых меток, кодирующих УИД, на этапе (104) выполняется его направление на печать. Распечатанный документ будет содержать закодированный УИД неразличимый для человеческого глаза. Размер цифровых меток может выбираться произвольно (например, метки радиусом от 1-2 пикселей).

[0036] На Фиг. 3 приведена последовательность этапов, выполняемых при выполнении способа (300) распознавании УИД на распечатанных документах. На этапе (301) вычислительное устройство, используемое для определения УИД в распечатанном документе, получает изображение такого документа. Изображение может содержать полностью или частично текст, с закодированным УИД, полученный, например, с помощью

фотографирования внешним устройством (смартфон, камера и т.п.) или при помощи сканирования с помощью OCR распечатанного документа.

[0037] Далее на этапе (302) также при помощи технологии OCR выполняется распознавание букв в документе, при этом если страниц в документе несколько, то распознается каждая страница документа. На этапе (303) выполняется считывание цифровых меток в окрестностях распознанных букв. Пример анализа цифровых меток может осуществляться по примеру, приведенному в Таблице 5, которая может применяться как таблица для сопоставления меток соответствующей цифре УИД пользователя.

[0038] После этого выполняется декодирование УИД на этапе (304) и установление по нему табельного номера сотрудника и соответствующего пользователя, с компьютерного устройства которого была осуществлена печать документа.

[0039] Математическое обоснование метода

[0040] Для этого убедимся, что частоты раскрытия позиций $TAB_8 = \{n_1, n_2, \dots, n_m\}$, $m = 8$ равномерно распределены для всех m , что позволяет показать вероятность извлечения табельного номера (УИД) из текста страницы.

[0041] Для математического обоснования было проведено исследование по частоте встречающихся букв в тексте с разным содержанием, к примеру, рассмотрим, такое распределение характерное для литературных произведений. Список литературных произведений, участвующих в эксперименте: Сильмариллион. Дж.Р.Р.Толкин, Двадцать тысяч лье под водой. Жюль Г.Верн, Двадцать лет спустя. Александр Дюма, Три мушкетера. Александр Дюма, Унесенные ветром. Маргарет Митчелл, Айвенго. Вальтер Скотт, Герой нашего времени. Н.В. Гоголь, Война и мир. Л.Н.Толстой, Обитаемый остров. Борис и Аркадий Стругацкие, Преступление и наказание. Ф.М.Достоевский, Живые и мертвые. К.М.Симонов, всего 8 366 594 символов, 3919 страниц. Математическая лингвистика показала следующие вероятности частоты встречи букв русского алфавита в текстах (Таблица 6).

Таблица 6. Таблица частоты встречи букв русского алфавита в художественной литературе

Буква	Частота встречи, %	Буква	Частота встречи, %
а	8,31	р	4,32
б	1,65	с	5,24
в	4,59	т	6,06
г	1,72	у	2,95
д	3,06	ф	0,13

е	8,42	х	0,84
ё	0,02	ч	1,56
ж	1,01	ц	0,44
з	1,71	ш	0,97
и	6,84	щ	0,32
й	1,11	ъ	0,03
к	3,42	ы	1,81
л	4,99	ь	1,93
м	3,16	э	0,27
н	6,46	ю	0,57
о	11,42	я	1,95
п	2,71		

[0042] Для получения значения частоты раскрытия позиций $TAB_8 = \{n_1, n_2, \dots, n_m\}$ выполняются следующие действия. Из таблиц 4 и 5 известны буквы, в которые кодируются разряды. Для получения частоты раскрытия разрядов для алгоритма нанесения метки в пространстве возле буквы, частоты букв, в которые кодируются метки, складываются, т.к. позиция выкрывается при обнаружении метки хотя бы в одной из них. В результате вышеописанных действий получается таблица 7.

Таблица 7. Таблица частоты раскрытия позиций табельного номера.

n_i, i	Частота встречи букв								Частота раскрытия разряда		
1	а	8,31	з	1,71	п	2,71	ч	0,44			13,17
2	б	1,65	и	6,84	р	4,32	ш	0,97			13,78
3	в	4,59	й	1,11	с	5,24	щ	0,32			11,26
4	г	1,72	к	3,42	т	6,06	ъ	0,03			11,23
5	д	3,06	л	4,99	у	2,95	ы	1,81			12,81
6	е	8,42	м	3,16	ф	0,13	э	0,27			11,84
7	ё	0,02	н	6,46	х	0,84	ь	1,93	я	1,95	11,57
8	ж	1,01	о	11,42	ц	1,56	ю	0,57			14,56

[0043] На основании таблицы 7 формируется диаграмма, представленная на Фиг. 4. Диаграмма показывает, что частота раскрытия всех позиций распределена относительно равномерно.

[0044] Вычислим количество каждой буквы русского алфавита экспериментальной выборки:

Таблица 8. Буквенно-позиционная количественная характеристика экспериментальной выборки.

n_i, i		Кол-во символов	Симв / стр		Кол-во символов	Симв / стр		Кол-во символов	Симв / стр		Кол-во символов	Симв / стр		Кол-во символов	Симв / стр
1	а	689 971	176	з	142 242	36	п	213 560	54	ч	121 417	31			
2	б	140 050	36	и	572 432	146	р	372 611	95	ш	74 968	19			
3	в	372 447	95	й	92 969	24	с	448 533	114	щ	26 501	7			
4	г	152 827	39	к	283 925	72	т	516 921	132	ъ	2 588	1			
5	д	255 254	65	л	420 003	107	у	234 845	60	ы	162 890	42			
6	е	709 671	181	м	257 188	66	ф	12 025	3	ь	165 110	42			
7	ё	5 953	2	н	536 626	137	х	75 243	19	э	24 582	6	я	171 316	44
8	ж	88 798	23	о	940 740	240	ц	31 199	8	ю	51 195	13			

[0045] Для метода нанесения точки в пространство возле буквы принимается следующее допущение: процент P символов на внедрение метки $P = 0,3$, при передаче через мессенджеры теряется определенный процент $M = 0,7$ меток. На основании вышеописанного можно вычислить вероятность распознавания текста, если для дешифрования доступно:

целая страница;

$\frac{1}{2}$ страницы;

$\frac{1}{4}$ страницы.

Таблица 9. Пояснения и вероятностей распознавания текста, закодированного методом нанесения точки в пространство возле буквы

n_i, i	выбирается 30 % общего количества в разряде	передача по мессенджеру (70% меток теряется)	целая страница		$\frac{1}{2}$		$\frac{1}{4}$	
			букв текущей позиции и на странице	позиция распознана или нет	букв текущей позиции и на странице	позиция распознана или нет	позиция распознана или нет	позиция распознана или нет
1	89	27	27	1	13,4	1	6,7	1
2	89	27	27	1	13,3	1	6,7	1

3	72	22	22	1	10,8	1	5,4	1
4	73	22	22	1	11,0	1	5,5	1
5	82	25	25	1	12,3	1	6,2	1
6	88	26	26	1	13,1	1	6,6	1
7	49	15	15	1	7,4	1	3,7	1
8	85	26	26	1	12,8	1	6,4	1
				100		100		100

[0046] Пример Экспериментального применения.

[0047] В ходе тестирования было распечатано и анализировано около 500 страниц разного содержания:

текст, разреженный текст, текст с таблицами, текст с графиками, текст с формулами;
с разными типами шрифтов: Arial, Calibri, Times New Roman;

разное оформление текста: обычный, курсив, полужирный, подчеркнутый;

разной размерности: 12px, 14px;

разным межстрочным интервалом: 0.5, 1.15, 1,5;

разным межзнаковым интервалом: обычный, разреженный, уплотнённый;

[0048] В каждом случае рассматривалась возможность извлечения метки с:

распечатки напрямую;

с фотографии распечатки;

переданной по мессенджеру распечатки фотографии.

[0049] Печать проводилась на офисном черно белом лазерном принтере Lexmark MX711de на офисной бумаге «Снегурочка» с белизной CIE 146 по ISO 11475.

[0050] Фотографирование производилось на телефон Samsung A51 при офисном освещении, бумага лежит горизонтально на столе, фотографирование случайное под разными, незначительными углами, порядком 2-4% в 3-х измерениях.

[0051] При передаче фотографий использовался мессенджер Telegram со сжатием изображения при отправлении.

[0052] В ходе эксперимента подбирались параметры, такие как размер меток, их оптимальные места и способы нанесения. Результаты последней фазы эксперимента показаны в таблице 10.

Таблица 10. Результат эксперимента.

ТЕКСТ	позиция	1	2	3	4	5	6	7	8	100%
	обнаруженная буква	а	и	с	т	д	е	н	о	
РАЗРЕЖЕННЫЙ ТЕКСТ	позиция	1	2	3	4	5	6	7	8	100%
	обнаруженная буква	а	и	с	т	л	е	н	о	
ТАБЛИЦА	позиция	1	2	3	4	5	6	7	8	100%

	обнаруженная буква	п	и	с	т	д	м	н	ж	
ГРАФИК	позиция	1	2	3	4	5	6	7	8	75%
	обнаруженная буква	а	и	в	т	л			ж	
ФОРМУЛА	позиция	1	2	3	4	5	6	7	8	62%
	обнаруженная буква	ч	и			к	ф	н		

[0053] Вышеописанная таблица показывает хорошие результаты анализа переданных по мессенджеру фотографий распечаток на офисном черно-белом принтере. В результате эксперимента были подобраны оптимальные параметры для внедрения метки, которые с одной стороны, были бы заметны на распечатках как дефекты принтера, с другой стороны, хорошо извлекались из переданных фотографий по мессенджером.

[0054] На Фиг. 5 представлен общий вид вычислительного устройства (500), пригодного для выполнения вышеуказанных способов. Устройство (500) может представлять собой, например, компьютер, сервер или иной тип пригодного вычислительного устройства.

[0055] В общем случае вычислительное устройство (500) содержит объединенные общей шиной информационного обмена один или несколько процессоров (501), средства памяти, такие как ОЗУ (502) и ПЗУ (503), интерфейсы ввода/вывода (504), устройства ввода/вывода (505), и устройство для сетевого взаимодействия (506).

[0056] Процессор (501) (или несколько процессоров, многоядерный процессор) могут выбираться из ассортимента устройств, широко применяемых в текущее время, например, компаний Intel™, AMD™, Apple™, Samsung Exynos™, MediaTek™, Qualcomm Snapdragon™ и т.п. В качестве процессора (501) может также применяться графический процессор, например, Nvidia, AMD, Graphcore и пр.

[0057] ОЗУ (502) представляет собой оперативную память и предназначено для хранения исполняемых процессором (501) машиночитаемых инструкций для выполнения необходимых операций по логической обработке данных. ОЗУ (502), как правило, содержит исполняемые инструкции операционной системы и соответствующих программных компонент (приложения, программные модули и т.п.).

[0058] ПЗУ (503) представляет собой одно или более устройств постоянного хранения данных, например, жесткий диск (HDD), твердотельный накопитель данных (SSD), флэш-память (EEPROM, NAND и т.п.), оптические носители информации (CD-R/RW, DVD-R/RW, BlueRay Disc, MD) и др.

[0059] Для организации работы компонентов устройства (500) и организации работы внешних подключаемых устройств применяются различные виды интерфейсов В/В (504).

Выбор соответствующих интерфейсов зависит от конкретного исполнения вычислительного устройства, которые могут представлять собой, не ограничиваясь: PCI, AGP, PS/2, IrDa, FireWire, LPT, COM, SATA, IDE, Lightning, USB (2.0, 3.0, 3.1, micro, mini, type C), TRS/Audio jack (2.5, 3.5, 6.35), HDMI, DVI, VGA, Display Port, RJ45, RS232 и т.п.

[0060] Для обеспечения взаимодействия пользователя с вычислительным устройством (500) применяются различные средства (505) В/В информации, например, клавиатура, дисплей (монитор), сенсорный дисплей, тач-пад, джойстик, манипулятор мышь, световое перо, стилус, сенсорная панель, трекбол, динамики, микрофон, средства дополненной реальности, оптические сенсоры, планшет, световые индикаторы, проектор, камера, средства биометрической идентификации (сканер сетчатки глаза, сканер отпечатков пальцев, модуль распознавания голоса) и т.п.

[0061] Средство сетевого взаимодействия (506) обеспечивает передачу данных устройством (500) посредством внутренней или внешней вычислительной сети, например, Интранет, Интернет, ЛВС и т.п. В качестве одного или более средств (506) может использоваться, но не ограничиваясь: Ethernet карта, GSM модем, GPRS модем, LTE модем, 5G модем, модуль спутниковой связи, NFC модуль, Bluetooth и/или BLE модуль, Wi-Fi модуль и др.

[0062] Дополнительно могут применяться также средства спутниковой навигации в составе устройства (500), например, GPS, ГЛОНАСС, BeiDou, Galileo.

[0063] Представленные материалы заявки раскрывают предпочтительные примеры реализации технического решения и не должны трактоваться как ограничивающие иные, частные примеры его воплощения, не выходящие за пределы испрашиваемой правовой охраны, которые являются очевидными для специалистов соответствующей области техники.

ФОРМУЛА

1. Способ кодирования информации для защиты от ее утечек при печати документов, выполняемый с помощью процессора компьютерного устройства, при этом способ содержит этапы, на которых:

получают на компьютерном устройстве пользователя информацию о печати по меньшей мере одного цифрового документа, содержащего по меньшей мере текст, при этом компьютерное устройство связано с уникальным идентификатором (УИД) пользователя;

осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой

распознают буквы, содержащиеся в цифровом документе;

кодируют УИД пользователя в набор цифровых меток, которые располагаются на контурах букв и/или вблизи контуров букв цифрового документа;

передают цифровой документ на печать с закодированным УИД пользователя.

2. Способ по п.1, характеризующийся тем, что распознавание цифрового документа выполняется с помощью оптического распознавания символов (OCR).

3. Способ по п.2, характеризующийся тем, что распознаются все символы на каждой странице цифрового документа.

4. Способ по п.1, характеризующийся тем, что каждый символ УИД пользователя кодируется в двоичный код.

5. Способ по п.4, характеризующийся тем, что на основании разряда двоичного кода определяется область размещения цифровых меток.

6. Способ защиты информации от утечек на печатных документах, выполняемый с помощью процессора компьютерного устройства, при этом способ содержит этапы, на которых:

получают по меньшей мере часть изображения печатного документа с закодированным УИД пользователя способом по любому из пп. 1-5;

выполняют распознавание полученного изображения;

определяют буквы, содержащие цифровые метки в своей окрестности;

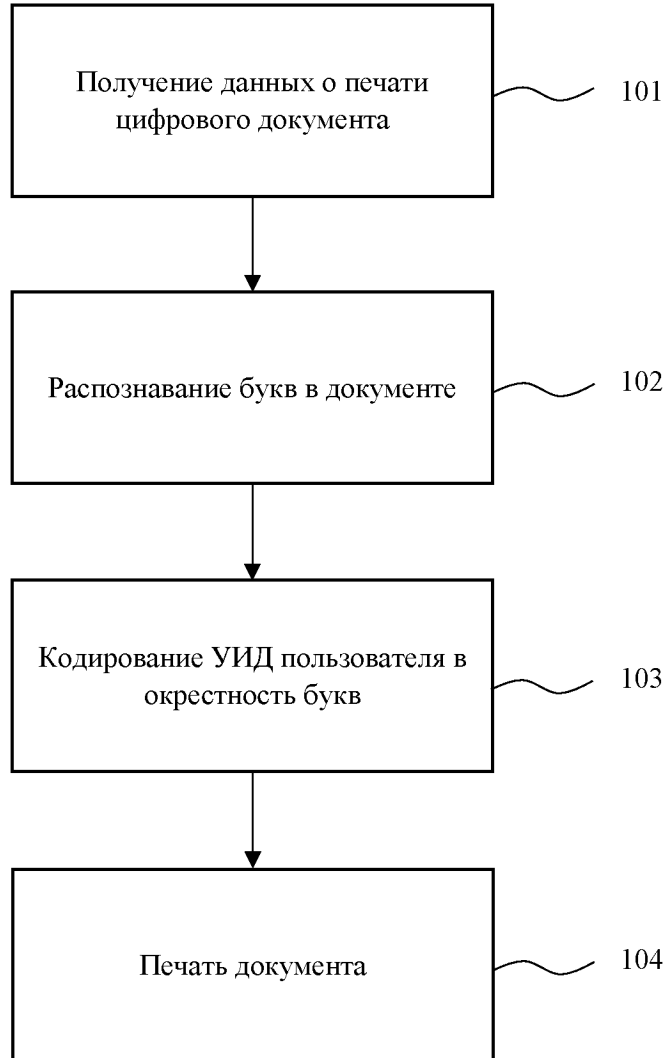

выполняют определение и извлечение закодированного УИД.

7. Способ по п.6, характеризующийся тем, что распознавание цифрового документа выполняется с помощью OCR.

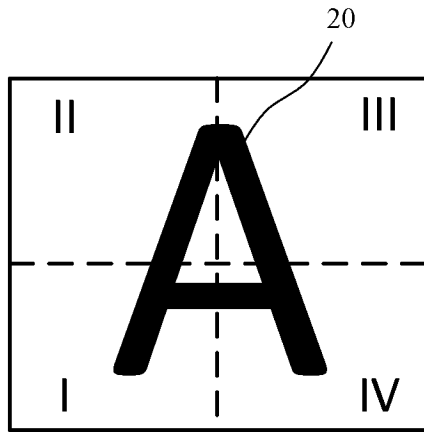
8. Система кодирования информации для защиты от ее утечек при печати документов, содержащая по меньшей мере один процессор, по меньшей мере одну память, связанную с процессором и содержащую машиночитаемые инструкции, которые при их исполнении процессором осуществляют способ по любому из пп. 1-5.

9. Система защиты информации от утечек на печатных документах, содержащая по меньшей мере один процессор, по меньшей мере одну память, связанную с процессором и содержащую машиночитаемые инструкции, которые при их исполнении процессором осуществляют способ по любому из пп. 6-7.

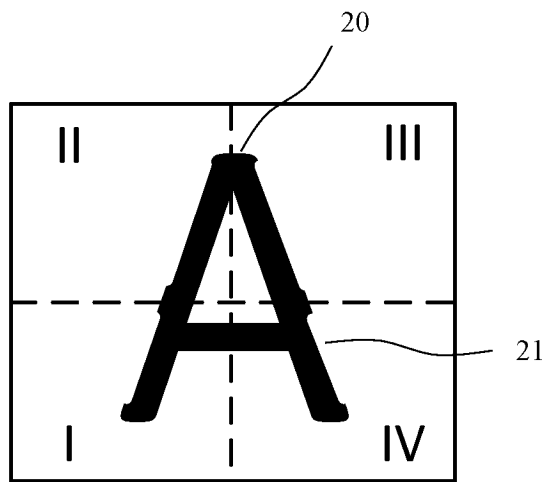
100



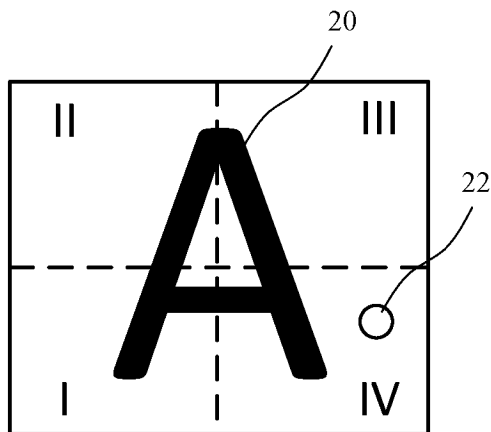
Фиг. 1



Фиг. 2А

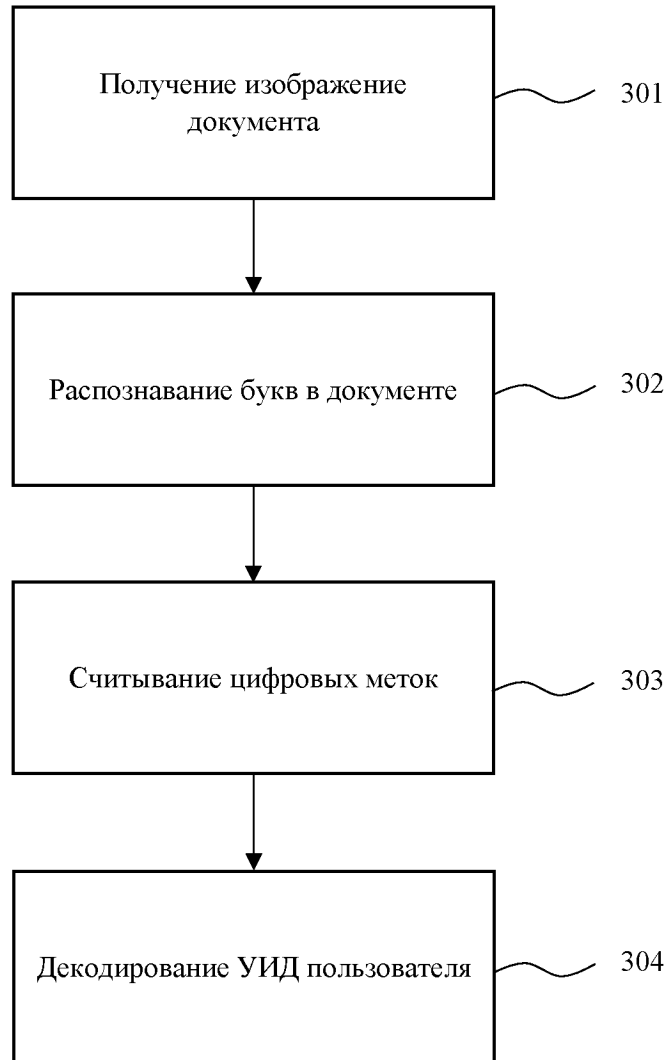



Фиг. 2Б

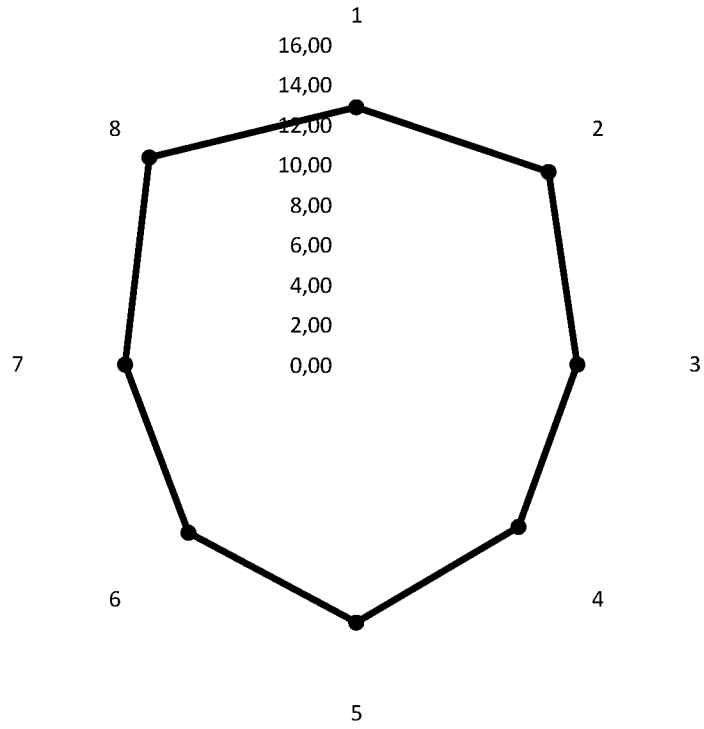


Фиг. 2В

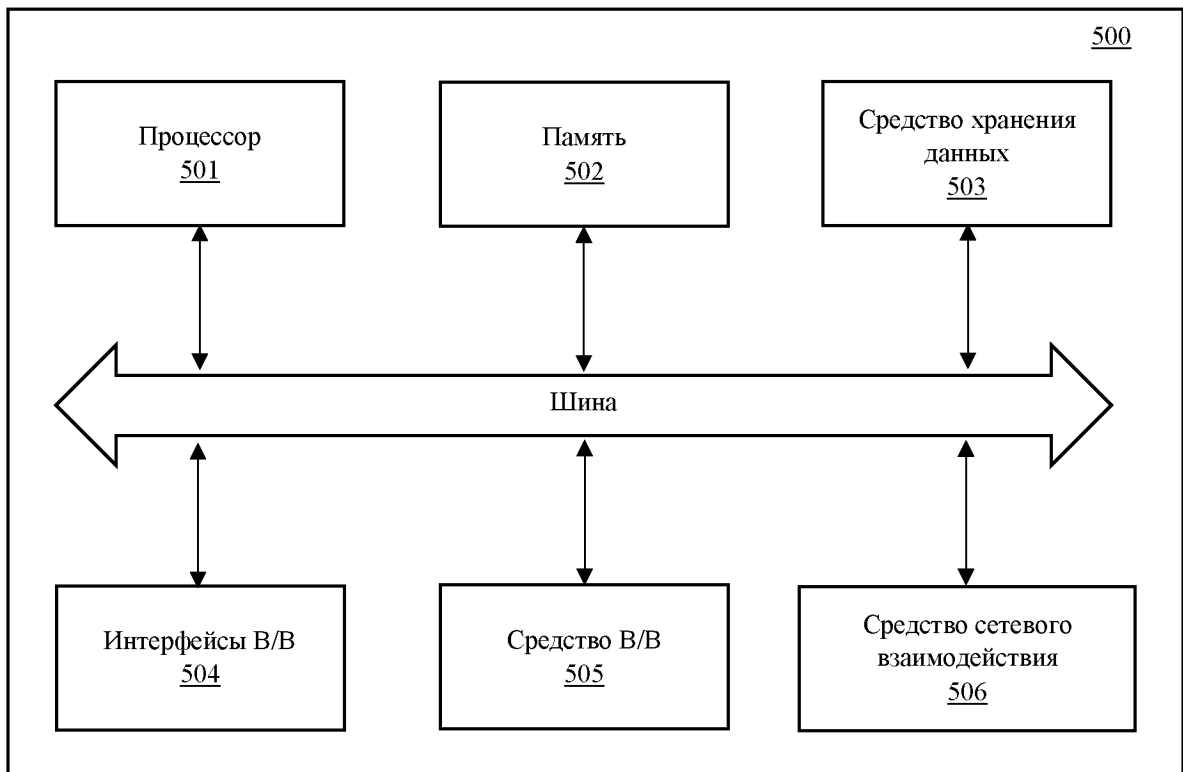
300



Фиг. 3



Фиг. 4



Фиг. 5

ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ

(статья 15(3) ЕАПК и правило 42 Патентной инструкции к ЕАПК)

Номер евразийской заявки:

202293485**А. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ:**

G06F 21/16 (2013.01)
 G06F 21/00 (2013/01)
 G06F 21/88 (2013.01)
 G06V 30/224 (2022.01)
 G06V 30/244 (2022.01)
 G06V 30/20 (2022.01)
 G06F 18/00 (2023.01)
 G06F 18/20 (2023.01)

Согласно Международной патентной классификации (МПК)

Б. ОБЛАСТЬ ПОИСКА:

Просмотренная документация (система классификации и индексы МПК)

G06F 21/00 - 21/16, 18/00 - 18/20, 21/88; G06V 30/00 - 30/20, 30/224 - 30/244

Электронная база данных, использовавшаяся при поиске (название базы и, если, возможно, используемые поисковые термины)
 EAPATIS, YANDEX, GOOGLE, USPTO

В. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ

Категория*	Ссылки на документы с указанием, где это возможно, релевантных частей	Относится к пункту №
A	US20120206751 A1 (CSR TECHNOLOGY INC; CSR IMAGING US LP) 16-08-2012 формула изобретения; описание, абзацы: [0019]; [0082]; [0086]; [0087]; [0093];	1-9
A	US2013080869 A1 (ARTAG SARL) 28-03-2013 формула изобретения; описание, абзацы: [0017-0019]; [0050-0052];	1-9
A	US20040189682 A1 (PORTAUTHORITY TECHNOLOGIES LLC) 30-09-2004 реферат; формула изобретения; описание, абзацы: [0011-0014]; [0042]; [0080]; [0109 - 0113]; [0129]; [0130]; [0140-0146]; [0148];	1-9
A	US20160035060 A1 (SEND ONLY OKED DOCUMENTS) 04-02-2016 реферат; формула изобретения; описание, абзацы: [0014]; [0015]; [0019]; [0023]; [0025]; [0026]; [0089]; [0090]; [0093]; [0107-0112]; [0171-0174]; [0173]; [0177]; [0224]; [0249-0252];	1-9
A	US20080310672 A1 (SURSEN CORP) 18-12-2008 реферат; формула изобретения; описание, абзацы [0024-0029], [0025]; [0043]; [0050]; [0055]; [0075]; [0085]; [0086]; [0105];	1-9
A	US20040001606 A1 (DIGIMARC CORP) 01-01-2004 реферат; формула изобретения; описание, абзацы [0025]; [0028]; [0037]; [0043]; [0045]; [0046]; [0048]; [0052]; [0053]; [0056]; [0066];	1-9

 последующие документы указаны в продолжении

* Особые категории ссылочных документов:

«А» - документ, определяющий общий уровень техники

«D» - документ, приведенный в евразийской заявке

«E» - более ранний документ, но опубликованный на дату подачи евразийской заявки или после нее

«O» - документ, относящийся к устному раскрытию, экспонированию и т.д.

"P" - документ, опубликованный до даты подачи евразийской заявки, но после даты испрашиваемого приоритета"

«Т» - более поздний документ, опубликованный после даты приоритета и приведенный для понимания изобретения

«X» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий новизну или изобретательский уровень, взятый в отдельности

«Y» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий изобретательский уровень в сочетании с другими документами той же категории

«&» - документ, являющийся патентом-аналогом

«L» - документ, приведенный в других целях

Дата проведения патентного поиска: **14/06/2023**

Уполномоченное лицо:

Начальник отдела механики,
 физики и электротехники



Д.Ф. Крылов