

(19)



**Евразийское  
патентное  
ведомство**

(11) **045968**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента  
**2024.01.23**

(21) Номер заявки  
**202293484**

(22) Дата подачи заявки  
**2022.12.27**

(51) Int. Cl. **G06F 21/64** (2013.01)  
**G06F 40/10** (2020.01)  
**H04N 1/00** (2006.01)

---

(54) **СПОСОБ И СИСТЕМА ЗАЩИТЫ ИНФОРМАЦИИ ОТ УТЕЧКИ ПРИ ПЕЧАТИ ДОКУМЕНТОВ С ПОМОЩЬЮ СМЕЩЕНИЯ СИМВОЛОВ**

---

(31) **2022106205**

(32) **2022.03.10**

(33) **RU**

(43) **2023.09.29**

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

(56) US-A1-20040001606  
US-A1-20070047818  
US-A1-20130028466  
US-A1-20080205699  
RU-C2-2446464  
RU-C1-2758666

(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 уникального кода пользователя в документ 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 УИД (уникальный идентификатор) пользователя; осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой распознают буквы, содержащиеся в цифровом документе, кодируют УИД пользователя путем смещения букв по горизонтальной оси на заданное количество пикселей, передают цифровой документ на печать с закодированным УИД пользователя.

---

**B1**

**045968**

**045968**

**B1**

### Область техники

Изобретение относится к области защиты информации, в частности к решениям для предотвращения утечки информации при печати документов.

### Уровень техники

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.

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.

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

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

Технический результат заключается в повышении эффективности защиты данных от утечки, за счет внедрения уникального кода пользователя в документ 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 УИД (уникальный идентификатор) пользователя;

осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой распознают буквы, содержащиеся в цифровом документе,

кодируют УИД пользователя путем смещения букв по горизонтальной оси на заданное количество пикселей,

передают цифровой документ на печать с закодированным УИД пользователя.

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

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

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

В другом частном примере реализации способа на основании разряда двоичного кода определяются буквы для их смещения.

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 смещенные по горизонтальной оси буквы;

выполняют определение и извлечение закодированного УИД.

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

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

### Краткое описание чертежей

Фиг. 1 иллюстрирует блок-схему способа кодирования цифровой метки.

- Фиг. 2 иллюстрирует пример смещения букв для кодирования УИД.  
 Фиг. 3 иллюстрирует блок-схему декодирования цифровых меток.  
 Фиг. 4 иллюстрирует диаграмму частоты раскрытия позиций УИД.  
 Фиг. 5 иллюстрирует общий вид вычислительного устройства.

### Осуществление изобретения

На фиг. 1 представлен способ (100) защиты информации в цифровых документах от утечки с помощью кодирования УИД пользователя путем смещения букв в документе. 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.

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.

После этапа распознавания цифрового документа на этапе (103) осуществляется процесс кодирования УИД. УИД представляет собой, например, числовой табельный номер сотрудника - цифровой код ТАВ, 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 табл. 1.

Таблица 1

Схематичное изображение табельного номера

$TAB_8$								
Цифра	$n_1$							
Позиция	1	2	3	4	5	6	7	8

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 табл. 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

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  $b_i = \{c_1, c_2, c_3, c_4\}$ . Главная особенность алгоритма кодировки заключается в том, что на каждую букву навешивается только один разряд в табельном номере. Схематично схема кодирования представлена в табл. 3.

Таблица 3

Схема кодирования

Разряд	Буква	Позиция
$c_1$	а	$n_1$
$c_2$	з	
$c_3$	п	

C <sub>4</sub>	ч	
C <sub>1</sub>	б	n <sub>2</sub>
C <sub>2</sub>	и	
C <sub>3</sub>	р	
C <sub>4</sub>	ш	
C <sub>1</sub>	в	n <sub>3</sub>
C <sub>2</sub>	й	
C <sub>3</sub>	с	
C <sub>4</sub>	щ	
C <sub>1</sub>	г	n <sub>4</sub>
C <sub>2</sub>	к	
C <sub>3</sub>	т	
C <sub>4</sub>	ъ	
C <sub>1</sub>	д	n <sub>5</sub>
C <sub>2</sub>	л	
C <sub>3</sub>	у	
C <sub>4</sub>	ы	
C <sub>1</sub>	е	n <sub>6</sub>
C <sub>2</sub>	м	
C <sub>3</sub>	ф	
C <sub>4</sub>	ь	
C <sub>1</sub>	н	n <sub>7</sub>
C <sub>2</sub>	х	
C <sub>3</sub>	э	
C <sub>4</sub>	я	
C <sub>1</sub>	ж	n <sub>8</sub>
C <sub>2</sub>	о	
C <sub>3</sub>	ц	
C <sub>4</sub>	ю	

Таким образом, буквы (20), у которых разряд  $c_i=1$  в тексте двигаются влево вправо на некоторое количество пикселей, как это представлено на фиг. 2.

Шифрование текста происходит следующим образом.

Пусть документ D содержит l страниц, тогда документ D - есть массив страниц,  $D=\{p_1, p_2, p_3, \dots, P_l\}$ ,  $l \in \mathbb{N}$ .

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  $p_i$ .

Символы распознаются с помощью, к примеру, OCR, из них выявляются русские буквы  $W_{rus_{p_i}} \in S_{p_i}$ .

Далее создаются 32 массива  $Pos_1, Pos_2, \dots, Pos_{32}$ , каждый из которых соответствует каждой позиции табельного номера. Каждый массив  $Pos$  заполняется теми символами из  $W_{rus_{p_i}}$ , которые соответствуют позиции из табл. 3. Например,  $Pos_1$  заполняется всеми символами из  $W_{rus_{p_i}}$ , которые имеют значения {а}, вне зависимости от регистра.

Массивы  $Pos_1, Pos_2, \dots, Pos_{32}$  перемешиваются, к примеру, тасованием Кнута. Пусть  $l_{Pos_1}, l_{Pos_2}, l_{Pos_3}, \dots, l_{Pos_{32}}$  - размерности полученных массивов, P - процент символов на внедрение метки  $P \in [0,3 \dots 0,7]$ , тогда каждый массив из  $Pos_1, Pos_2, \dots, Pos_{32}$  обрезается с конца до размерности  $\sum_{i=0}^{32} \frac{Pos_i}{32} \cdot P, Pos_1, Pos_2, \dots, Pos_{32} \rightarrow Pos_1^P, Pos_2^P, \dots, Pos_{32}^P$ .

Полученные массивы  $Pos_1^P, Pos_2^P, \dots, Pos_{32}^P$  используются для смещения символов по вышеописанному способу. После внедрения всех меток на искомой странице  $p_i$  тоже самое проделывается для следующей страницы  $p_{i+1}$  и так далее до конца документа  $p_l$ . Рассмотрим пример с кодированием UID пользователя - 00013400. В соответствии с табл. 2 UID переводится в двоичную форму для последующего определения разрядов по табл. 3. В соответствии с табл. 3 определяются буквы, соответствующие 1 в двоичном коде цифры UID. Для UID "00013400" получаются буквы: Ч Ш С Т Л Ы Ю Я. Данные буквы смещаются на заданное количество пикселей по горизонтали на каждой странице документа, направленного на печать. Изменение документа выполняется с помощью технологии OCR путем распознавания букв для замены и

их замены в цифровом документе.

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 практически неразличимый для человеческого глаза.

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.

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 определение букв, которые смещены по горизонтали.

Принцип выявления смещенных букв в слове заключается в следующем: пусть  $W$  - слово в тексте,  $W = \{Ch_1, Ch_2, \dots, Ch_n\}$  - слово представляет собой массив символов,  $n$  - количество символов в слове. Тогда пусть  $L_w = \{L_1, L_2, \dots, L_{n-1}\}$  массив расстояний между символами в слове, где  $L_i = |Ch_i - Ch_{i+1}|$ ,  $i \in [1..n-1]$ . Далее анализируется массив  $L_w$  и методами выявления аномалий находятся аномалии в  $L_i$ , таким, как, к примеру, локальный уровень выброса, метод  $k$ -ближайших соседей или их аналоги, в том числе построенные на нейронных сетях. Смещенным будет считаться тот символ из  $W$ , чьи индексы  $i$  будут совпадать  $L_i \rightarrow Ch_i$ .

После этого выполняется декодирование УИД на этапе (304) в соответствии с примерами в табл. 2 и 3. Это позволяет установить табельный номер сотрудника (пользователя), с компьютерного устройства которого была осуществлена печать документа.

Математическое обоснование метода.

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.

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.

Математическая лингвистика показала следующие вероятности частоты встречи букв русского алфавита в текстах (табл. 4).

Таблица 4

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		

Для получения значения частоты раскрытия позиций  $TAB_8 = \{n_1, n_2, \dots, n_m\}$  выполняются следующие действия. Из табл. 3 и 4 известны буквы, в которые кодируются разряды. Вычисляется среднее арифметическое частот множества букв, по которым вскрыт разряд, так как для вскрытия разряда нужно иметь все буквы разряда, где есть значение 1 в  $TAB_{BIN}^8$ . В результате вышеописанных действий получается табл. 5.

Таблица 5

Таблица частоты раскрытия позиций табельного номера

Частота встречи букв

Частота раскрытия разряда

$n_i, i$	Частота встречи букв								Частота раскрытия разряда		
1	а	8,31	з	1,71	п	2,71	ч	0,44		3,29	
2	б	1,65	и	6,84	р	4,32	ш	0,97		3,45	
3	в	4,59	й	1,11	с	5,24	щ	0,32		2,82	
4	г	1,72	к	3,42	т	6,06	ъ	0,03		2,81	
5	д	3,06	л	4,99	у	2,95	ы	1,81		3,20	
6	е	8,42	м	3,16	ф	0,13	ь	1,93		3,41	
7	ё	0,02	н	6,46	х	0,84	э	0,27	я	1,95	1,91
8	ж	1,01	о	11,42	ц	1,56	ю	0,57			3,64

На основании табл. 5 формируется диаграмма, представленная на фиг. 4.

Диаграмма показывает, что частота раскрытия всех позиций распределена относительно равномерно.

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.

Вычислим количество каждой буквы русского алфавита экспериментальной выборки.

Таблица 6

## Буквенно-позиционная количественная характеристика экспериментальной выборки

$n_i, i$		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			

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;

$1/2$  страницы;

$1/4$  страницы.

Таблица 7

## Пояснения и вероятностей распознавания текста, закодированного методом нанесения точки в пространство возле буквы

$n_i, i$	MIN необходимый набор символ на страницы для раскрытия позиции	передача по мессенджеру (метки не теряются)	целая страница		$1/2$		$1/4$	
			букв текущей позиции на странице	позиция распознала сь или нет	букв текущей позиции на странице	позиция распознала сь или нет	букв текущей позиции на странице	позиция распознала сь или нет
1	31	31	31	1	15,5	1	7,7	1
2	19	19	19	1	9,6	1	4,8	1
3	7	7	7	1	3,4	1	1,7	1
4	1	1	1	0	0,3	0	0,2	0
5	42	42	42	1	20,8	1	10,4	1
6	3	3	3	1	1,5	1	0,8	0
7	2	2	2	1	0,8	0	0,4	0
8	8	8	8	1	4,0	1	2,0	1
Вероятности распознавания:				87,5		75		62,5

Пример экспериментального применения.

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;

с разным оформлением текста: обычный, курсив, полужирный, подчеркнутый;

разной размерности: 12px, 14px;

разным межстрочным интервалом: 0,5, 1,15, 1,5;

разным межзнаковым интервалом: обычный, разреженный, уплотненный;

В каждом случае рассматривалась возможность извлечения метки с распечатки напрямую;

фотографии распечатки;

переданной по мессенджеру распечатки фотографии.

Печать проводилась на офисном черно белом лазерном принтере Lexmark MX711de на офисной бумаге "Снегурочка" с белизной CIE 146 по ISO 11475.

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.

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

В ходе эксперимента подбирались параметры, такие как размер меток, их оптимальные места и способы нанесения. Результаты последней фазы эксперимента показаны в табл. 8.

Таблица 8

		Результат эксперимента																												
ТЕКСТ	обнаруженная буква			з				р	ш	й				к	ь	л	ы			м	ф	ь	н				ж	ю	100%	
	разряды	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	0	0	0	1	0
РАЗРЕЖЕННЫЙ ТЕКСТ	обнаруженная буква			з				р	ш	й				к	ь	л	ы			м	ф	ь	н				ж	ю	100%	
	разряды	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	0	0	0	1	0
ТАБЛИЦА	обнаруженная буква			з				р	ш	й				к	ь	л	ы			м	ф	ь	н				ж	ю	100%	
	разряды	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	0	0	0	1	0
ГРАФИК	обнаруженная буква			з				р	ш	й				к	ь	л	ы			м	ф	ь	н				ж	ю	100%	
	разряды	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	1	0	0	1	0
ФОРМУЛА	обнаруженная буква			з				р	ш	й				к	ь	л	ы			м	ф	ь	н				ж	ю	100%	
	разряды	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	1	0	0	1	0

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

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.

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).

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.

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.).

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.

Для организации работы компонент 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. 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.

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.

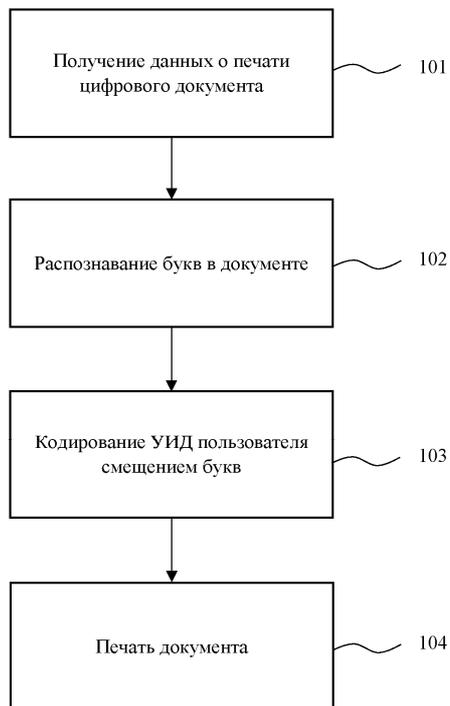
Дополнительно могут применяться также средства спутниковой навигации в составе устройства (500), например GPS, ГЛОНАСС, BeiDou, Galileo. Представленные материалы заявки раскрывают предпочтительные примеры реализации технического решения и не должны трактоваться как ограничивающие

иные, частные примеры его воплощения, не выходящие за пределы испрашиваемой правовой охраны, которые являются очевидными для специалистов соответствующей области техники.

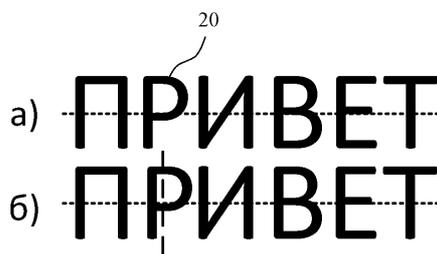
#### ФОРМУЛА ИЗОБРЕТЕНИЯ

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 УИД (уникальным идентификатором) пользователя; осуществляют до момента передачи цифрового документа на печать его обработку, в ходе которой распознают буквы, содержащиеся в цифровом документе, переводят УИД пользователя в двоичную форму, отображают каждый разряд двоичного представления на соответствующие символы, кодируют УИД пользователя путем смещения букв по горизонтальной оси влево или вправо на заданное количество пикселей на каждой странице 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 определяются буквы для их смещения.
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 смещенные по горизонтальной оси буквы; выполняют определение и извлечение закодированного УИД.
7. Способ по п.6, характеризующийся тем, что распознавание полученного изображения 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

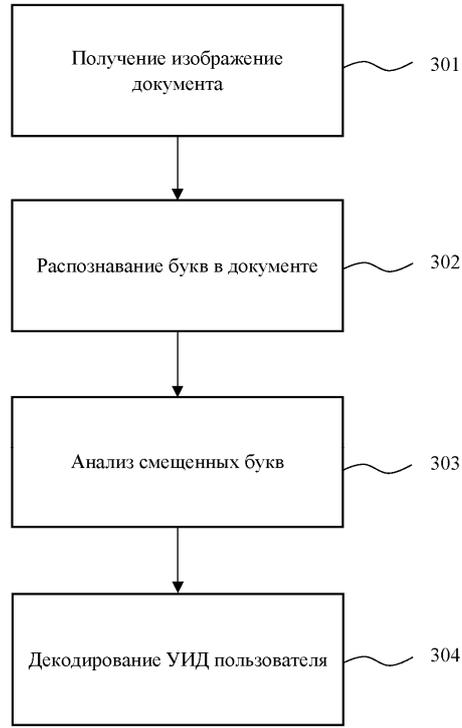


Фиг. 1

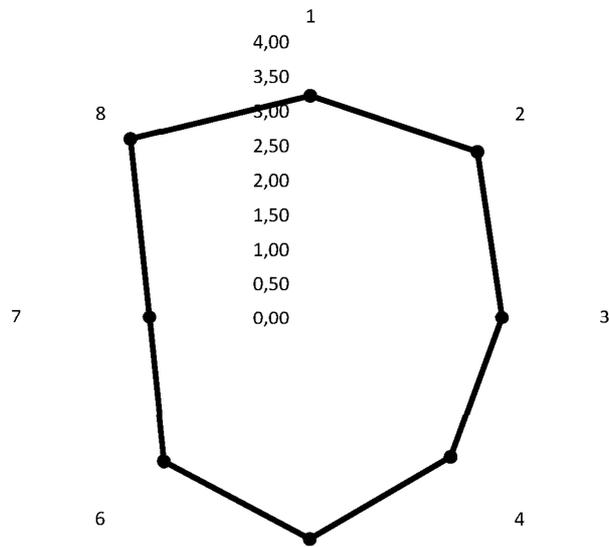


Фиг. 2

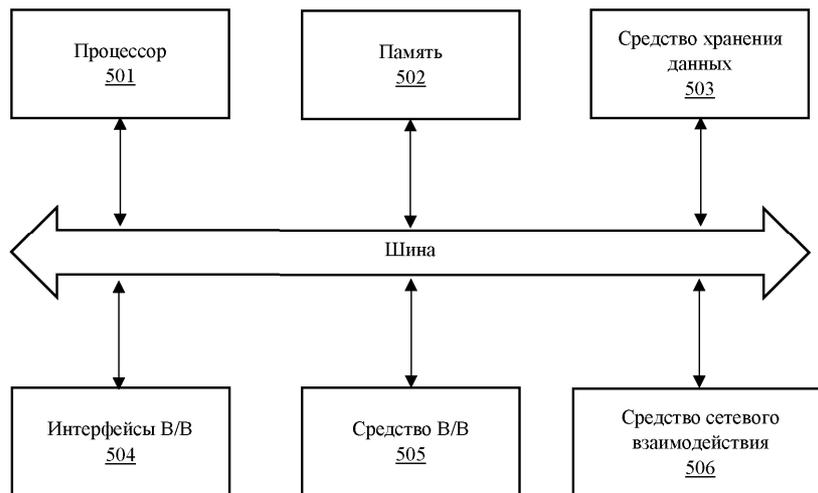
300



Фиг. 3



Фиг. 4



Фиг. 5