

(19)



**Евразийское
патентное
ведомство**

(21) **202293484** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.09.29

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

(22) Дата подачи заявки
2022.12.27

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

(31) 2022106205

(72) Изобретатель:

(32) 2022.03.10

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

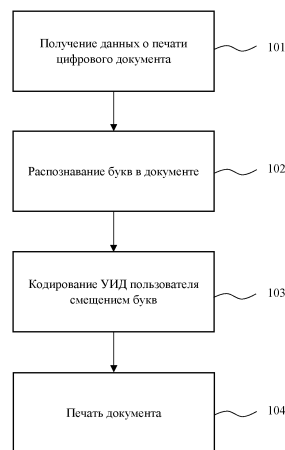
(33) RU

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

100



A1

202293484

202293484

A1

СПОСОБ И СИСТЕМА ЗАЩИТЫ ИНФОРМАЦИИ ОТ УТЕЧКИ ПРИ ПЕЧАТИ ДОКУМЕНТОВ С ПОМОЩЬЮ СМЕЩЕНИЯ СИМВОЛОВ

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 уникального кода пользователя в документ 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 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.

[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 определяются буквы для их смещения.

[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 смещенные по горизонтальной оси буквы;
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. 2 иллюстрирует пример смещения букв для кодирования УИД.

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

[0018] Фиг. 4 иллюстрирует диаграмму частоты 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 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) осуществляется 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_1	n_1	n_1	n_1	n_1	n_1	n_1
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 $b_i = \{c_1, c_2, c_3, c_4\}$.

[0025] Главная особенность алгоритма кодировки заключается в том, что на каждую букву навешивается только один разряд в табельном номере. Схематично схема кодирования представлена в таблице 3.

Таблица 3. Схема кодирования

Разряд	Буква	Позиция
c_1	а	n_1
c_2	з	
c_3	п	

c_4	ч	n_2
c_1	б	
c_2	и	
c_3	р	
c_4	ш	
c_1	в	n_3
c_2	й	
c_3	с	
c_4	щ	
c_1	г	n_4
c_2	к	
c_3	т	
c_4	ь	
c_1	д	n_5
c_2	л	
c_3	у	
c_4	ы	
c_1	е	n_6
c_2	м	
c_3	ф	
c_4	ь	
c_1	н	n_7
c_2	х	
c_3	э	
c_4	я	
c_1	ж	n_8
c_2	о	
c_3	ц	
c_4	ю	

[0026] 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.

[0027] Шифрование текста происходит следующим образом.

[0028] Пусть документ D содержит l страниц, тогда документ D – есть массив страниц, $D = \{p_1, p_2, p_3 \dots p_l\}, l \in N$.

[0029] 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 $Wrus_{p_i} \in S_{p_i}$.

[0030] 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 $Wrus_{p_i}$, которые соответствуют позиции из таблицы 3. Например, Pos_1 заполняется всеми символами из $Wrus_{p_1}$, которые имеют значения $\{a\}$, вне зависимости от регистра.

[0031] 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 $\sum_{i=0}^{32} \frac{Pos_i}{32} \cdot P$, $Pos_1, Pos_2 \dots Pos_{32} \rightarrow Pos_1^P, Pos_2^P \dots Pos_{32}^P$.

[0032] 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 .

[0033] Рассмотрим пример с кодированием УИД пользователя - 00013400.

[0034] В соответствии с Таблицей 2 УИД переводится в двоичную форму для последующего определения разрядов по Таблице 3. В соответствии с Таблицей 3 определяются буквы, соответствующие 1 в двоичном коде цифры УИД. Для УИД «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.

[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 практически неразличимый для человеческого глаза.

[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 определение букв, которые смещены по горизонтали.

[0038] 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 $L_i = |Ch_i - Ch_{i+1}|$, $i \in [1 \dots n - 1]$. 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. Смещённым будет считаться тот символ из W , чьи индексы i будут совпадать $L_i \rightarrow Ch_i$.

[0039] После этого выполняется декодирование УИД на этапе (304) в соответствии с примерами в Таблицах 2 и 3. Это позволяет установить табельный номер сотрудника (пользователя), с компьютерного устройства которого была осуществлена печать документа.

[0040] Математическое обоснование метода.

[0041] 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.

[0042] 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 (Таблица 6).

Таблица 5. 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

	раскрытия позиции		страниц е		страниц е		страниц е	
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

[0048] 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;
- 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;

[0049] В каждом случае рассматривалась возможность извлечения метки с:

- распечатки напрямую;
- с фотографии распечатки;
- переданной по мессенджеру распечатки фотографии.

[0050] Печать проводилась на офисном черно белом лазерном принтере Lexmark MX711de на офисной бумаге «Снегурочка» с белизной CIE 146 по ISO 11475.

[0051] 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.

[0052] При передаче фотографий использовался мессенджер Telegram со сжатием изображения при отправлении.

[0053] В ходе эксперимента подбирались параметры, такие как размер меток, их оптимальные места и способы нанесения. Результаты последней фазы эксперимента показаны в таблице 9.

Таблица 9. Результат эксперимента.

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

[0054] 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 мессенджером.

[0055] 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.

[0056] 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).

[0057] 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.

[0058] ОЗУ (502) представляет собой оперативную память и предназначено для хранения исполняемых процессором (501) машиночитаемых инструкций для выполнение 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.).

[0059] 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.

[0060] 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.

[0061] 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.

[0062] 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.

[0063] Дополнительно могут применяться также средства спутниковой навигации в составе устройства (500), например, GPS, ГЛОНАСС, BeiDou, Galileo.

[0064] 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 УИД 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.

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.

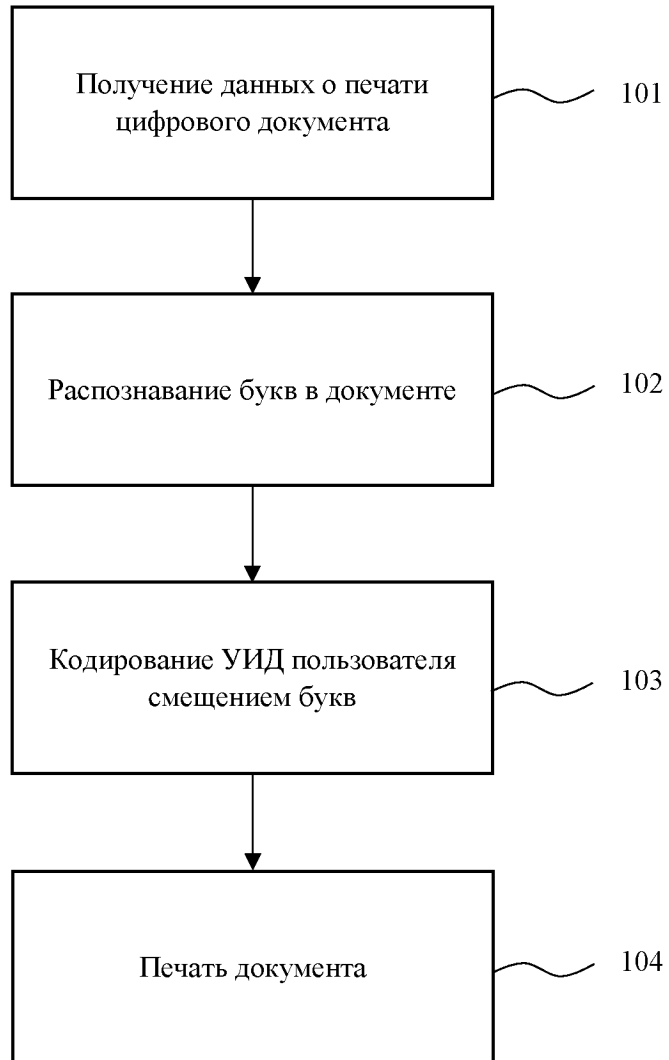

7. Способ по п. 6, характеризующийся тем, что распознавание полученного изображения выполняется с помощью OCR.

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.

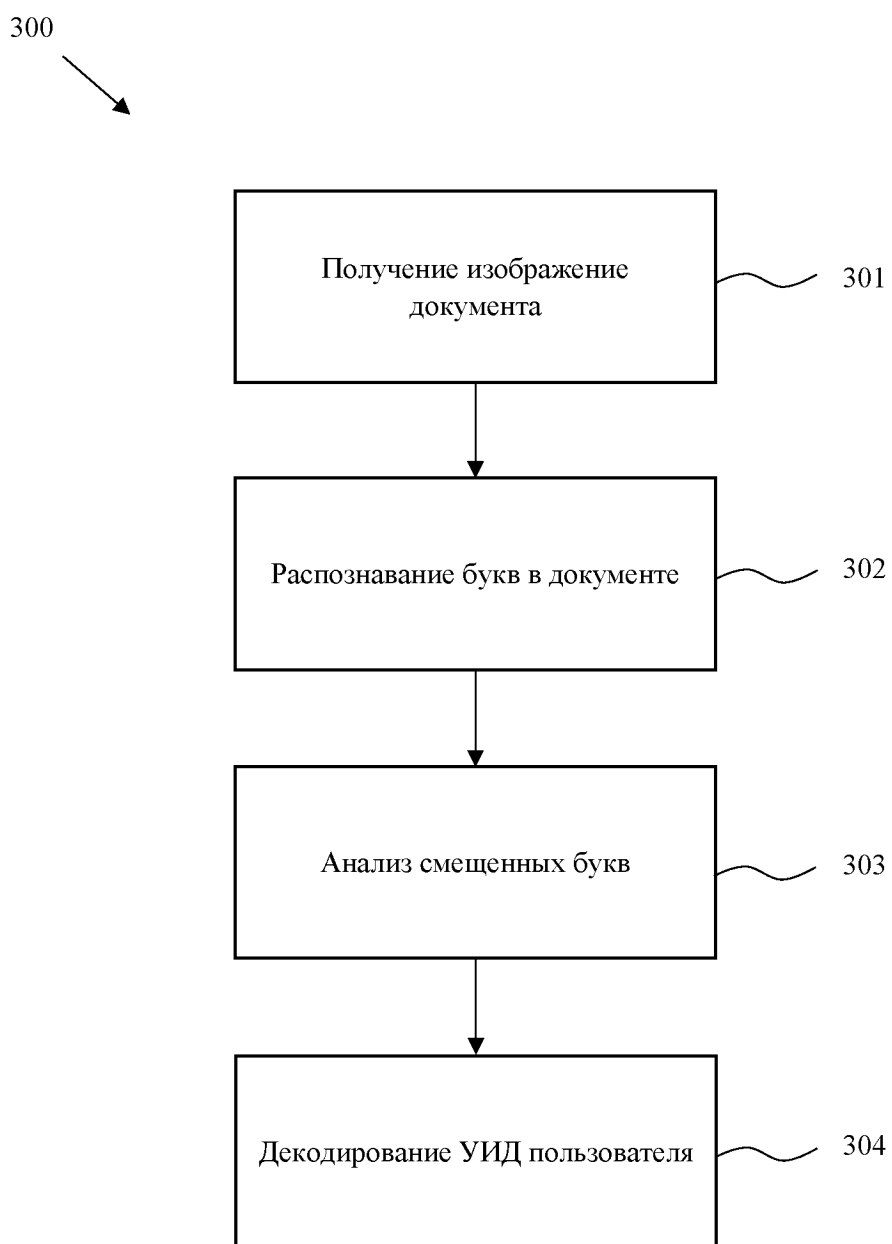
100



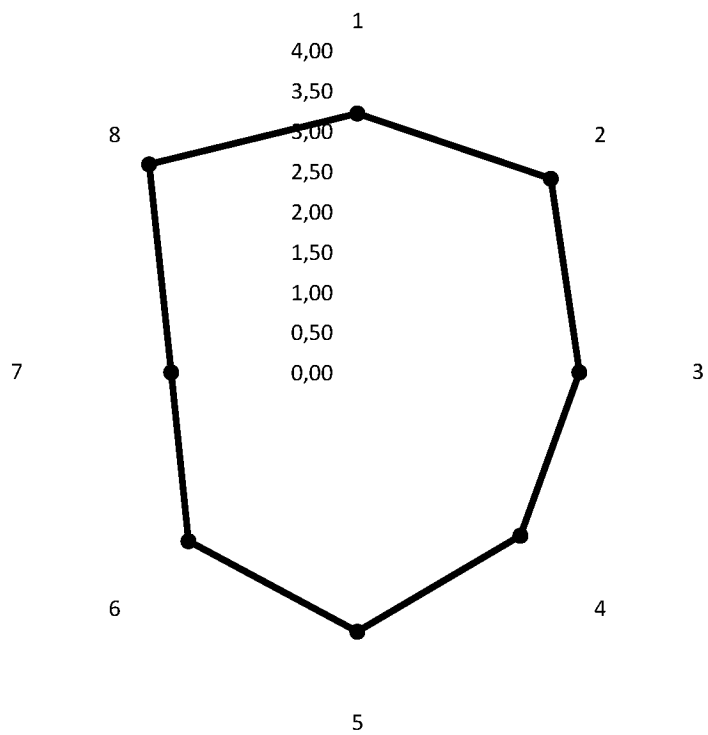
Фиг. 1



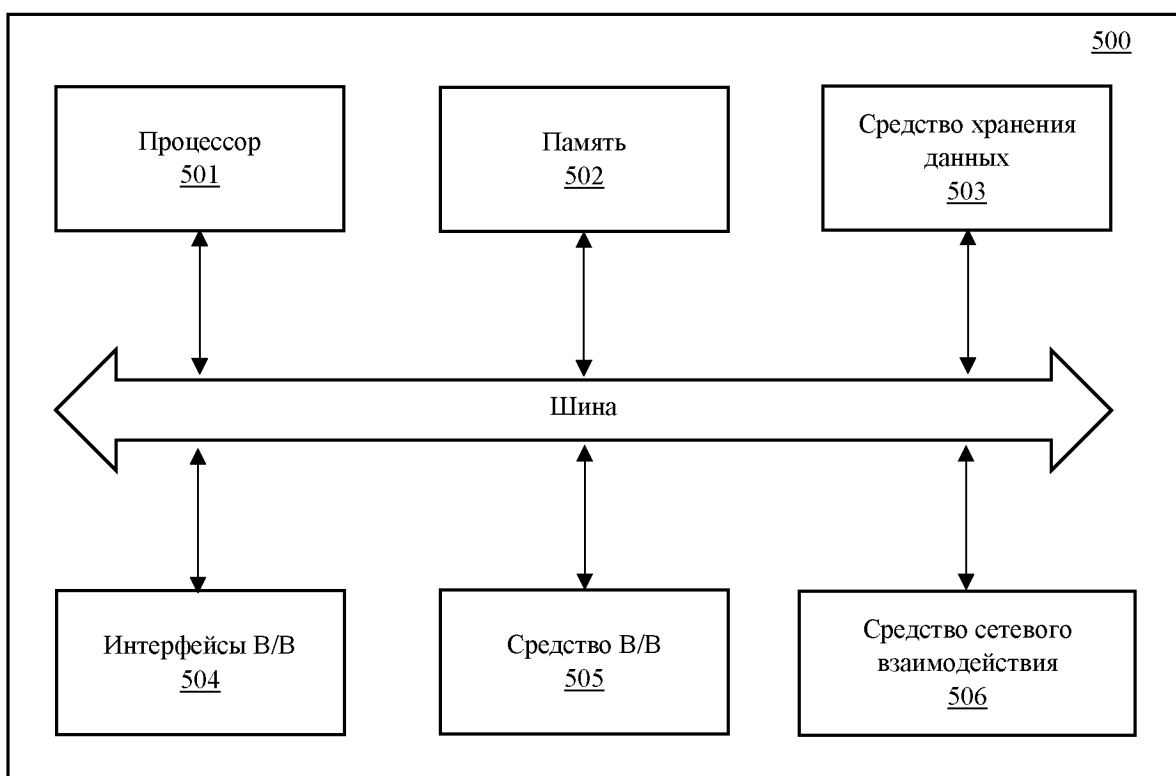
Фиг. 2



Фиг. 3



Фиг. 4



Фиг. 5

ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ

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

Номер евразийской заявки:

202293484**А. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ:***G06F 21/64 (2013.01)**G06F 40/10 (2020.01)**H04N 1/00 (2006.01)*

Согласно Международной патентной классификации (МПК)

Б. ОБЛАСТЬ ПОИСКА:

Просмотренная документация (система классификации и индексы МПК)

G06F 21/00, 21/10, 21/16, 21/60, 40/00, 40/10, 3/00, 3/12, G06V 30/00, 30/10, G06K 19/00, 19/06, H04N 1/00-1/44

Электронная база данных, использовавшаяся при поиске (название базы и, если возможно, используемые поисковые термины)
Espacenet, ЕАПАТИС, ЕРОQUE Net, Reaxys, Google**В. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ**

Категория*	Ссылки на документы с указанием, где это возможно, релевантных частей	Относится к пункту №
X	US 2004/0001606 A1 (KENNETH L. LEVY) 01.01.2004,	1-2, 4-9
Y	реферат, параграфы [0004], [0011]-[0012], [0031], [0042], [0048]-[0049], [0052], [0058], [0065]-[0066], [0069]-[0070], [0072], [0090], пункт 17 формулы	3
Y	US 2007/0047818 A1 (JONATHAN J. HULL et al.) 01.03.2007, параграфы [0094], [0130]	3
A	US 2013/0028466 A1 (SURSEN CORP.) 31.01.2013	1-9
A	US 2008/0205699 A1 (FUJITSU LIMITED) 28.08.2008	1-9
A	RU 2446464 C2 (КОРПОРАЦИЯ "САМСУНГ ЭЛЕКТРОНИКС КО., ЛТД") 27.03.2012	1-9
A	RU 2758666 C1 (ПУБЛИЧНОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "СБЕРБАНК РОССИИ" (ПАО СБЕРБАНК)) 01.11.2021	1-9

 последующие документы указаны в продолжении

* Особые категории ссылочных документов:

«А» - документ, определяющий общий уровень техники

«D» - документ, приведенный в евразийской заявке

«Е» - 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

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

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

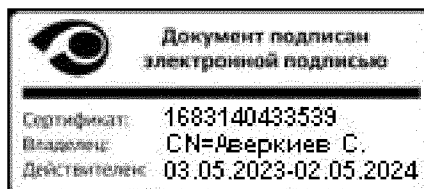
«&» - документ, являющийся патентом-аналогом

«L» - документ, приведенный в других целях

Дата проведения патентного поиска: 03 июля 2023 (03.07.2023)

Уполномоченное лицо:

Начальник Управления экспертизы



С.Е. Аверкиев