

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ – ПРОЦЕССОВ
УПРАВЛЕНИЯ

ОТЧЕТ
по лабораторной работе №4
по дисциплине «Алгоритмы и структуры данных»
на тему «Деобезличивание данных»
Вариант 13

Выполнила
студентка 2 курса
группы 21-Б15.ПУ
Павлова Ксения Андреевна

Преподаватель
Дик Александр Геннадьевич

Санкт-Петербург
2022

СОДЕРЖАНИЕ

Цель.....	3
Задачи.....	3
Блок-схема общего алгоритма.....	4
Примечания к блок-схеме общего алгоритма.....	5
Рекомендации программиста.....	6
Рекомендации пользователя	6
Контрольный пример	7
Вывод	8
Список литературы	9

Цель: написать программу деобезличивания набора данных.

Задачи:

1. Деобезличить первый столбец в таблице – номер телефона;
2. Деобезличить второй и третий столбцы данных в таблице, для каждой строчки которых использован сдвиг буквы по шифру Цезаря, идентичный для одной строки двух столбцов и не обязательно одинаковый для разных строк таблицы;
3. Для каждой строки таблицы вывести шаг, на который были сдвинуты 2 и 3 столбцы.

Блок-схема общего алгоритма.

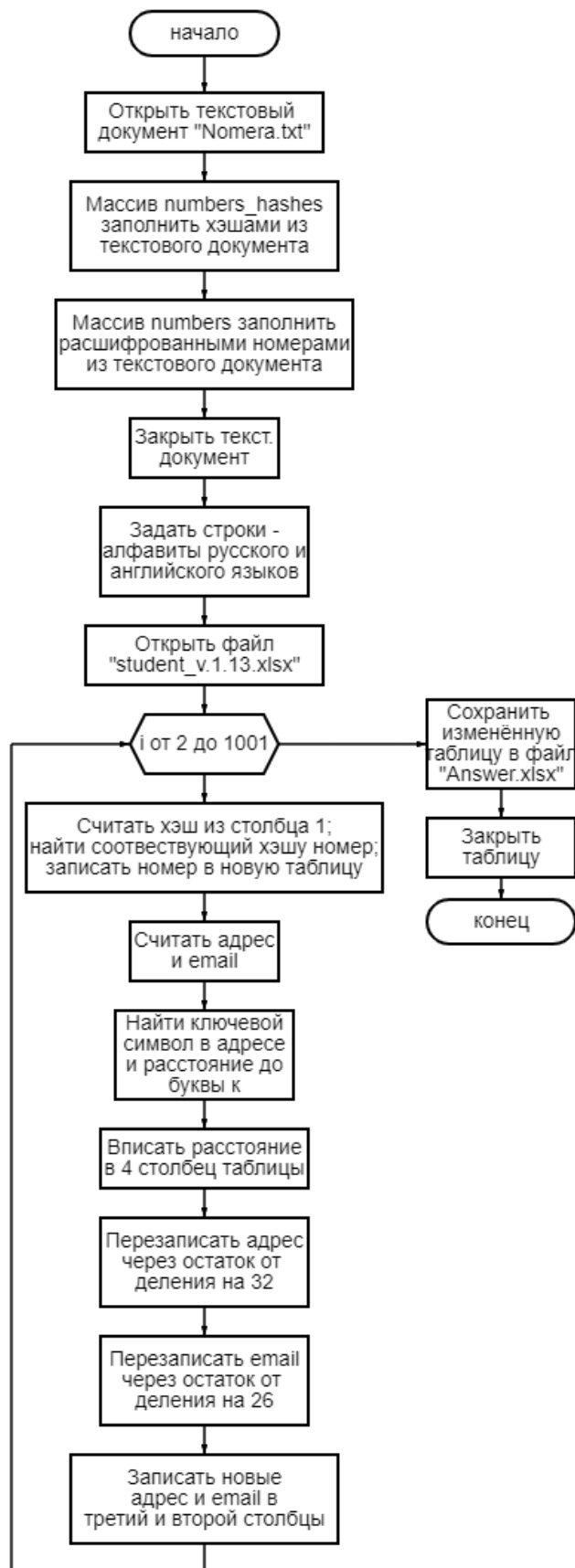


Рис. 1. Блок-схема общего алгоритма

Примечания к блок-схеме общего алгоритма (рис. 1).

Номера телефонов в 1 столбце были зашифрованы с помощью алгоритма хэширования «SHA-1».

Номера телефонов были восстановлены с помощью утилиты HashCat и записаны в текстовый файл «Nomera.txt» в формате «<хэш-значение>:<расшифрованный номер телефона>».

Работа с файлами формата xlsx происходит с помощью библиотеки xlsxwriter.

В алфавите русского языка отсутствует буква ё.

Так как в адресе всегда присутствует номер квартиры и обозначение кв., то ключевой символ – первый символ последней части адреса, разделенного по пробелам. Расстояние (сдвиг) в шифре Цезаря – разность номера ключевого символа и 10 (номер буквы к в алфавите).

Номер символа, который заменяет зашифрованный в адресе, высчитывается как остаток разности номера зашифрованного символа и расстояния по модулю 32.

В случае, когда сдвиг неотрицательный, номер символа, который заменяет зашифрованный в email, высчитывается как остаток разности номера зашифрованного символа и расстояния по модулю 26. Иначе, сдвиг искусственно увеличивается на 32 и номер символа, который заменяет зашифрованный в email, высчитывается по той же формуле, с уже новым расстоянием. Тем самым происходит зацикливание алфавита и нейтрализуется разница в количестве символов в русском и английском алфавите.

В конечный файл вписывается начальный сдвиг.

Рекомендации программиста.

Для работы с кодом необходимо приложение PyCharm, HashCat и установленная библиотека xlswriter.

Рекомендации пользователя.

Перед запуском программы убедитесь, что в папке с файлом программы находится таблица «student_v.1.13.xlsx» с исходными зашифрованными данными и текстовый документ «Nomera.txt» с хэшами и соответствующими расшифрованными номерами.

Контрольный пример.

В данном разделе представлен пример работы данной программы.

На рис. 2 представлен результат работы программы.

Телефон	email	Адрес	Сдвиг
69196927521	fgertach@hotmail.com	ул. Плещеевская 89 кв. 23	17
69168988985	cabbott@tremblay.info	Садовая аллея 70 кв. 91	1
69257169569	mclaughlin.daren@yahoo.com	Солнечная ул. д. 17 кв. 406	2
69198277588	albin36@cummerata.info	1-й Богородский пер. д. 53 кв. 415	8
69169005940	aida18@bednar.biz	ул. М. Могученко 40 кв. 77	18
69656716483	mason.hartmann@hotmail.com	Рождественский пер. д. 86 кв. 497	19
69996998251	treichert@gmail.com	Южная ул. д. 39 кв. 204	17
69299527969	bogisch.lee@gmail.com	2-я Железнодорожная ул. д. 49 кв. 249	5
69367787674	clang@schneider.net	Тушинская ул. д. 38 кв. 139	17
69586301869	clemmie.jones@simonis.com	Весенняя ул. д. 7 кв. 84	7
69178103971	rosalinda.brown@hudson.biz	2-я Павлоградская ул. д. 83 кв. 434	1
69198213392	ina.heidenreich@gmail.com	Железнодорожная ул. д. 99 кв. 405	3
69018506587	idurgan@welch.com	ул. Милерина 86 кв. 413	5
69106694542	wisozk.kimberly@collins.com	Шайдаровская ул. д. 62 кв. 167	20
69777855330	alexis44@hotmail.com	Лесная ул. д. 79 кв. 328	19
69838657332	powlowski.autumn@cruickshank	Старовольная ул. д. 83 кв. 190	1
69867785430	guillermo.beer@yahoo.com	Боровский пр. д. 15 кв. 427	9
69368498333	conroy.dan@teeney.net	Зональная ул. д. 72 кв. 277	-10
6986727260	jarrett73@welch.com	Удмуртская ул. д. 9 кв. 178	15
69256054734	reinger.huthe@gmail.com	Египетский пер. д. 86 кв. 489	9
69296533927	jerde.daniela@gmail.com	Молодёжная ул. д. 46 кв. 283	2
69259133271	marie28@connelly.com	ул. Кирова 75 кв. 435	-9
69777161977	mellie28@wiegand.com	Южная ул. д. 88 кв. 94	-10
69017331399	rocio24@yahoo.com	Печатников пер. д. 37 кв. 334	11
69586114507	vandenvort.carmela@hotmail.co	ул. Зои и Александра	19
69776117504	slevie.nikolaus@runolfsson.com	Тестовская ул. д. 32 кв. 381	21
69296705401	felix.brandyn@krack.com	Лунальская ул. д. 42 кв. 218	15
69016489124	zachariah.aufderhar@gmail.co	Киевская ул. д. 2 кв. 366	14
69589619329	alfreda.wuckert@cummerata.co	Столбовая ул. д. 81 кв. 142	19
69176315182	anastasia55@gmail.com	ул. Эльдара Рязанова 74 кв. 395	5
69017554522	alexandre19@bradtke.com	Андроньевская пл. д. 80 кв. 446	4
69107547472	senger.leslie@kunde.com	7-я ул. Текстильщиков 56 кв. 170	7

Рис. 2. Вывод результата работы программы

Вывод: в ходе выполнения данной работы была написана программа деобезличивания набора данных.

Список литературы.

1. Introducing Python. Modern computing in simple packages / Bill Lobanovic.
2. <https://habr.com/ru/company/alexhost/blog/536490/>
3. Ссылка на код, необходимые и полученный файлы
<https://github.com/ksenkap/decryption>