



# МЕТОДИКА ТРАНСФОРМАЦИИ ЭЛЕКТОРАЛЬНЫХ ДАННЫХ, РАЗМЕЩАЕМЫХ НА САЙТЕ ЦЕНТРАЛЬНОЙ ИЗБИРАТЕЛЬНОЙ КОМИССИИ, В ФОРМУ ПРИГОДНУЮ ДЛЯ ИХ АВТОМАТИЗИРОВАННОГО АНАЛИЗА

*В статье на примере открытых официальных данных, представленных на сайте Центральной избирательной комиссии (ЦИК) Российской Федерации (РФ), по результатам выборов Президента РФ в 2018 г. описана автоматизированная методика трансформации электоральных в форму пригодную для их дальнейшего автоматизированного анализа, а также программные инструменты, обеспечивающие работоспособность предложенной методики. Обоснован выбор структур электоральных данных на уровнях избирательных комиссий (ИК) субъектов РФ, территориальных избирательных комиссий (ТИК) и участковых избирательных комиссий (УИК), извлеченных их единого файла, содержащего все электоральные данные, выгруженные с сайта ЦИК РФ.*

*Приведен пример использования преобразованных электоральных данных, подтверждающий отсутствие намерений их фальсификации на этапе формирования ИК, ТИК и УИК.*

**Ключевые слова:** выборы, электоральные данные, избирательная комиссия, анализ данных, информационный поиск.

# METHODOLOGY OF TRANSFORMATION OF ELECTORAL DATA POSTED ON THE WEBSITE OF THE CENTRAL ELECTION COMMISSION INTO A FORM SUITABLE FOR THEIR AUTOMATED ANALYSIS

*The article presents the automated method for the transformation of electoral data provided by the Central Election Commission (CEC) of the Russian Federation (RF) based on the results of the presidential elections in RF in 2018. There are described software tools that ensure the operability of the proposed methodology and the structure of electoral data at the levels of election commissions (EC) of RF regions, territorial election commissions (TEC) and district election commissions (DEC), this data was extracted from a single file with all electoral data downloaded from the CEC website of RF.*

*There is an example with the use of transformed electoral data confirming the absence of intentions to falsify data at the stage of formation of the EC, TEC and DEC.*

**Keywords:** *election, election data, election commission, data retrieval, data mining.*

## **Введение**

Проблемы, связанные с оценкой достоверности результатов тех или иных выборов в органы власти и местного самоуправления в РФ и выявления возможных фальсификаций электоральных данных, в современной политической ситуации являются предметом острых научных и политических дискуссий, в которых принимают участие политики, политологи, юристы, специалисты по прикладной статистике (электоральные статистики), а также политически активные избиратели) (см., например, [1,2]). При этом, в большинстве случаев, мнения участников дискуссий по обсуждаемой проблеме, в зависимости занимаемой ими политической позиции, оказываются диаметрально противоположными: от полного отрицания каких-либо фальсификаций электоральных данных (например, члены

ЦИК и избирательных комиссий нижних уровней), до полного отрицания возможности проведения в Российской Федерации (РФ) честных выборов (например, кандидаты, не получившие поддержки избирателей, а также члены либерально настроенного сообщества электоральных статистиков (см., например, [3]).

В [4] на основе анализа действующей нормативно-правовой базы в области защиты информации, доказано, что задачи обеспечения достоверности электоральных данных, как на этапе их получения, так и на этапе интерпретации, по своей сути, изоморфны задачам защиты информации, а потому для их решения необходимо применять соответствующие методы защиты информации. В этой связи, разработка методов защиты электоральных данных, к которым, безусловно, следует

отнести методы оценки достоверности выводов электоральных статистиков, является актуальной.

При этом основная проблема состоит в отсутствии электоральных данных, полученных по результатам «абсолютно честных» выборов, в которых отсутствие фальсификаций доказано с помощью методов, отличных от методов электоральной статистики (например, на основе данных, представленных независимыми наблюдателями, данных экзит-поллов и т.д.). Их анализ, как очевидно, мог бы позволить определить набор «эталонных» статистических характеристик, по значениям которых можно будет говорить об отсутствии или наличии фальсификаций электоральных данных.

Отметим, что проведение «абсолютно честных» выборов, процедура организации и проведения которых гарантировала недопущение каких-либо процедурных нарушений (как преднамеренных, так и не преднамеренных) и признание отсутствия таковых всеми сторонами избирательного процесса в реальных условиях, как очевидно, невозможно. В этой связи для оценки адекватности известных методов выявления электоральных аномалий представляется целесообразным использовать данные, представленные избирательными комиссиями (ИК) соответствующих уровней по результатам таких выборов, в которых преимущество одного из кандидатов являлось столь значительным, что его победа на выборах ни у кого не вызывала ни малейших сомнений, а потому использование каких-либо фальсификаций не требовалось. Такими были выборы Президента РФ в 2018 г., победа в которых за явным преимуществом была одержана действующим Президентом РФ В.В. Путиным.

Для автоматизированной выгрузки с сайта ЦИК РФ [5] соответствующих электоральных данных, представленных ИК субъектов РФ, территориальными избирательными комиссиями (ТИК) и участковыми избирательными комиссиями (УИК) о результатах выборов Президента РФ в 2018 г. (число которых превышало 97 тыс.) в файл CVS-формата была использована специальная технология, подробно описанная в [6]. (Отметим, что за время прошедшее с момента разработки данной технологии появилась новая версия сайта ЦИК РФ и обсуждаемые электоральные данные размещены на странице [7].) Наличие электоральных данных по каждой ИК субъекта РФ, ТИК и УИК позволяет провести количе-

ственный анализ особенностей избирательного процесса в каждом субъектах РФ. Например, в [8] на основе анализа электоральных данных, представленных ИК Свердловской области (которые были извлечены из соответствующего файла CVS-формата) вычислены оценки плотностей вероятностей второй после запятой цифры в электоральном показателе «отношение числа проголосовавших избирателей на данном избирательном участке к общему числу зарегистрированных избирателей», а также появления одинаковых цифр в первом и во втором разрядах после запятой (метод Бебера и Скакко), которые позволили сделать обоснованный вывод об отсутствии фальсификаций электоральных данных в СО на обсуждаемых выборах.

В данной статье обсуждаются автоматизированная методика извлечения из электоральных данных, полученных в ходе выборов Президента РФ в 2018 г., размещенных на сайте ЦИК РФ, а также их декомпозиция в соответствие со структурой ИК, ТИК и УИК, а также приводится пример анализа этих электоральных данных.

## **2. Структура электоральных данных о результатах выборов Президента РФ в 2018 г., выгруженных в файл CVS-формата, и ее модификация**

Напомним, что система избирательных комиссий по выборам Президента РФ в 2018 г. имела следующую иерархическую структуру: ЦИК, избирательные комиссии (ИК) субъектов РФ, территориальные ИК, участковые ИК. Всего в 2018 г. было сформировано: 85 ИК субъектов РФ. Кроме того, было сформировано еще 2 ИК без образования в них ТИК: ИК «Территория за пределами РФ», которой были отнесены 394 УИК, и ИК «Город Байконур (Республика Казахстан)», на территории которого были сформированы 7 участковых УИК. Таким образом, было сформировано: 87 ИК, 2776 ТИК, 97314 УИК.

При разработке программного инструмента, обеспечивающего автоматическую выгрузку обсуждаемых электоральных данных в файл CVS-формата, предполагалось, что эти данные имеют следующую иерархическую структуру: вершина дерева – данные ЦИК (сводные электоральные данные по всей РФ), ветви первого уровня – данные ИК субъектов РФ (сводные электоральные данные по соответствующему субъекту РФ), ветви второго уровня – данные территориальных ИК (сводные данные по территориальным ИК соответ-

ствующих субъектов РФ), ветви третьего уровня – данные участковых ИК, а информационные модели электоральных данных, размещаемых на вершине, а также первом и втором уровнях описанной выше иерархической структуры, одинаковы.

Для автоматической загрузки электоральных данных, представленных ИК после выборов Президента РФ в 2018 г., с сайта ЦИК [4] был использован обычный офисный ком-

пьютер (AMD Ryzen 5 4 ГГц, 32 Гб ОЗУ, Интернет со скоростью 50Мб/с. На выгрузку данных потребовалось 8 часов, в течение которых было выполнено 201156 http-запросов и скачано около 10 Гб сырых данных. Далее из первичных («сырых») данных извлечено 11Мб полезных данных, которые были сохранены в CSV-файл (в формате RFC 4180) со следующей структурой (таблица 1).

Фрагмент скриншот программы Excel с

Таблица 1

### Описание данных

Наименование столбца	Тип данных	Описание
I1	Строка	наименование ИК субъекта РФ
I2	Строка	Наименование территориальной ИК субъекта РФ
I3	Строка	Наименование участковой ИК 3 уровня, сформированной территориальной ИК субъекта РФ
c1	Целое число	Число избирателей, включенных в список избирателей
c2	Целое число	Число избирательных бюллетеней, полученных участковой избирательной комиссией
c3	Целое число	Число избирательных бюллетеней, выданных избирателям, проголосовавшим досрочно
c4	Целое число	Число избирательных бюллетеней, выданных в помещении для голосования в день голосования
c5	Целое число	Число избирательных бюллетеней, выданных вне помещения для голосования в день голосования
c6	Целое число	Число погашенных избирательных бюллетеней
c7	Целое число	Число избирательных бюллетеней в переносных ящиках для голосования
c8	Целое число	Число бюллетеней в стационарных ящиках для голосования
c9	Целое число	Число недействительных избирательных бюллетеней
c10	Целое число	Число действительных избирательных бюллетеней
c11	Целое число	Число утраченных избирательных бюллетеней
c12	Целое число	Число избирательных бюллетеней, не учтенных при получении
c13	Целое число	Число избирателей, проголосовавших за Бабурина Сергея Николаевича
c14	Целое число	Число избирателей, проголосовавших за Грудина Павла Николаевича
c15	Целое число	Число избирателей, проголосовавших за Жириновского Владимира Вольфовича
c16	Целое число	Число избирателей, проголосовавших за Путина Владимира Владимировича
c17	Целое число	Число избирателей, проголосовавших за Собчак Ксению Анатольевну
c18	Целое число	Число избирателей, проголосовавших за Сурайкина Максима Александровича
c19	Целое число	Число избирателей, проголосовавших за Титова Бориса Юрьевич
c20	Целое число	Число избирателей, проголосовавших за Явлинского Григория Алексеевича

открытым файлом, содержащим результаты выгрузки электоральных данных, представленных ИК после выборов Президента РФ в 2018 г., подтверждающий работоспособность

разработанного программного инструмента приведен на рисунке 1.

Данный файл содержал 100579 строк. В первой строке данного файла размещены на-

	A	B	C	D	E	F	G	H	I	J
1	I1	I2	I3	c1	c2	c3	c4	c5	c6	c7
7	Алтайский край	Алейская	УИК №559	215	220	0	143	13	64	13
8	Алтайский край	Алейская	УИК №560	469	470	0	350	33	87	33
9	Алтайский край	Алейская	УИК №561	428	420	0	326	42	52	42
10	Алтайский край	Алейская	УИК №562	76	79	0	66	3	10	3
11	Алтайский край	Алейская	УИК №563	886	820	0	585	139	96	139
12	Алтайский край	Алейская	УИК №565	220	210	0	136	29	45	29
13	Алтайский край	Алейская	УИК №566	880	860	0	561	117	182	117
14	Алтайский край	Алейская	УИК №567	357	340	0	287	18	35	18
15	Алтайский край	Алейская	УИК №568	160	160	0	131	13	16	13
16	Алтайский край	Алейская	УИК №569	178	170	0	143	19	8	19
17	Алтайский край	Алейская	УИК №570	450	460	0	361	33	66	33
18	Алтайский край	Алейская	УИК №571	779	720	0	546	43	131	43
19	Алтайский край	Алейская	УИК №572	52	67	0	45	3	19	3
20	Алтайский край	Алейская	УИК №573	455	460	0	332	48	80	48
21	Алтайский край	Алейская	УИК №574	583	560	0	411	42	107	42
22	Алтайский край	Алейская	УИК №575	464	480	0	333	77	70	77
23	Алтайский край	Алейская	УИК №578	141	130	0	107	8	15	8
24	Алтайский край	Алейская	УИК №579	73	80	0	55	6	19	6
25	Алтайский край	Алейская	УИК №580	439	430	0	327	25	78	25

Рис. 1. Фрагмент скриншота Excel с открытым файлом CSV-формата

звания столбцов, представленные в таблице 1. В остальных ячейках столбцов I1, I2, I3 размещены: названия субъекта РФ; названия созданной данным субъектом ТИК; упорядоченные по алфавиту, названия УИК, упорядоченные по алфавиту, созданных данным ТИК, в столбцах c1-c20 – соответствующие количественные показатели, описание которых представлено в таблице 2. Результаты анализ данного файла и далее их сравнение с аналогичными электоральными данными, размещенными на сайте ЦИК РФ [6], позволил сделать вывод из всех зарегистрированных УИК 21 не представили электоральные данные. В этой связи строки, соответствующие данным избирательным комиссиям были удалены из обсуждаемого файла.

Таким образом, общее число УИК, представивших электоральные данные по результатам выборов Президента РФ в 2018 г. составило 97293.

Сводные результаты по данной ТИК записывались в строку, следующую за строкой последней УИК, созданной данной ТИК. При этом в соответствующую ячейку столбца I3 заносилось значение null-string (null-string="") (рис. 1.). Сводные данные результаты по данной ИК, записывались в строку, следующую за строкой, содержащей сводные данные по последней ТИК, созданной данной ИК (рис. 2). В

результате число строк файла, содержащего обсуждаемые электоральные данные ЦИК, увеличилось до 1005879.

Дальнейший анализ данных показал, что ИК «Город Байконур (Республика Казахстан)» и ИК «Территория за пределами РФ» не создавали ТИК, но напрямую УИК – 9 и 369 в г. Байконур и за пределами РФ, соответственно. В связи с отсутствием обсуждаемых ТИК соответствующая информация по ТИК в файле отсутствовала (рис. 3).

Для автоматизации дальнейшей обработки электоральных данных о выборах Президента РФ в 2018 г. проведена следующая модификация файла CVS-формата: пустые ячейки с столбце I2, находящиеся в строках, содержащих электоральные данные УИК г. Байконур и УИК, созданных за пределами РФ, были заполнены текстом следующего содержания: «Без образования ТИК». Кроме того, в файл после строки с информацией, представленной последними из УИК г. Байконур и УИК территорий за пределами РФ были добавлены соответствующие строки, в ячейки которых находящиеся в столбцах c1c20 были оставлены пустыми (рис. 4).

Анализ результатов импорта обсуждаемого файла в рабочее пространство пакета MATLAB для его последующей обработки позволил обнаружить, что инструмент импорта

	A	B	C
1	I1	I2	I3
1901	Алтайский край	Яровская городская	УИК №548
1902	Алтайский край	Яровская городская	УИК №549
1903	Алтайский край	Яровская городская	УИК №550
1904	Алтайский край	Яровская городская	УИК №551
1905	Алтайский край	Яровская городская	УИК №552
1906	Алтайский край	Яровская городская	
1907	Алтайский край		
1908	Амурская область	Архаринского района	УИК №101

Рис. 2. Пример размещения сводных данных по Яровской городской ТИК Алтайского края и ИК Алтайского края

1	I1	I2	I3	c1	c2
12104	Город Байконур (Республика Казахстан)		УИК №8143	1769	1500
12105	Город Байконур (Республика Казахстан)		УИК №8144	1880	1500
12106	Город Байконур (Республика Казахстан)		УИК №8145	2047	1900
12107	Город Байконур (Республика Казахстан)		УИК №8146	2291	1900
12108	Город Байконур (Республика Казахстан)				
12109	город Москва	Академический район	УИК №2118	863	800

Рис. 3. Фрагмент файла CVS-формата, иллюстрирующий размещение электоральных данных каждой из УИК г. Байконур и сводных данных по этому городу

1	I1	I2	I3	c1	c2
12101	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8140	2132	2000
12102	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8141	2207	2000
12103	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8142	2249	2000
12104	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8143	1769	1500
12105	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8144	1880	1500
12106	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8145	2047	1900
12107	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8146	2291	1900
12108	Город Байконур (Республика Казахстан)	Без образования ТИК			
12109	Город Байконур (Республика Казахстан)	Пустое поле		14575	12800

Рис. 4. Фрагмент файла CVS-формата после добавления дополнительной строки в файл и заполнения пустых полей в столбце I2

данных выбранного пакета на смог определить тип данных, находящихся в столбце I2, в то время для пустых ячеек, находящихся в других столбцах таблицы такой проблемы не возникло. В этой связи в каждую пустую ячейку, находящуюся в столбце I2, был записан следующий текст: «Пустое поле».

В результате проведенных преобразований контента файла, содержащего электоральные данные, скаченные с сайта ЦИК, число его строк оказалось равным 100560.

Для удобства независимых исследователей электоральных данных о результатах выборов Президента РФ в 2018 сформулируем описанные выше действия с первичными данными в виде следующего алгоритма:

1. Открыть файл CVS-формата в инструменте для работы с электронными таблицами Excel.

2. Найти и удалить строки содержащие названия УИК, не представивших сведения о результатах выборов Президента РФ в 2018 г. (для их нахождения достаточно включить фильтр «Пустые» в любом из столбцов c1-c20).

3. Добавить новую строку после строки, содержащей поля «г. Байконур (Республика Казахстан)» (столбец I1), и заполнить поле данной строки, находящееся в столбце I2 текстом «Пустое поле».

4. Добавить новую строку после строки, содержащей поля «Территория за пределами РФ» (столбец I1) и заполнить поле данной

строки, находящееся в столбце I2, текстом «Пустое поле».

### 3. Автоматизированная методика трансформации электоральных данных, представленных Центральной избирательной комиссией по результатам выборов Президента Российской Федерации в 2018 г., в форму пригодную для их анализа

Из описания структуры файла, содержащего электоральные данные ЦИК РФ о результатах выборов Президента РФ в 2018 г., понятно, что целесообразно привести структуры информации, содержащуюся в данном файле, в соответствие структуре системы ИК, ТИК и УИК РФ. При этом, очевидно, что в связи с большим объемом первичных электораль-

ных данных реализация обсуждаемой декомпозиции электоральных данных вручную будет весьма трудоемкой. В этой связи требуется автоматизировать процесс решения данной задачи.

Для решения данной задачи авторы выбрали пакет MATLAB, являющийся сегодня одним из лидеров на рынке прикладного программного обеспечения, предназначенного, в том числе, для анализа данных.

На первом шаге осуществлялся импорт электоральных данных ЦИК из файла «Данные ЦИК.xlsx» (рис. 4) в рабочее пространство MATLAB. (Отметим, что здесь следует выбрать в меню «Output Type» тип экспортируемых данных «String Array».)

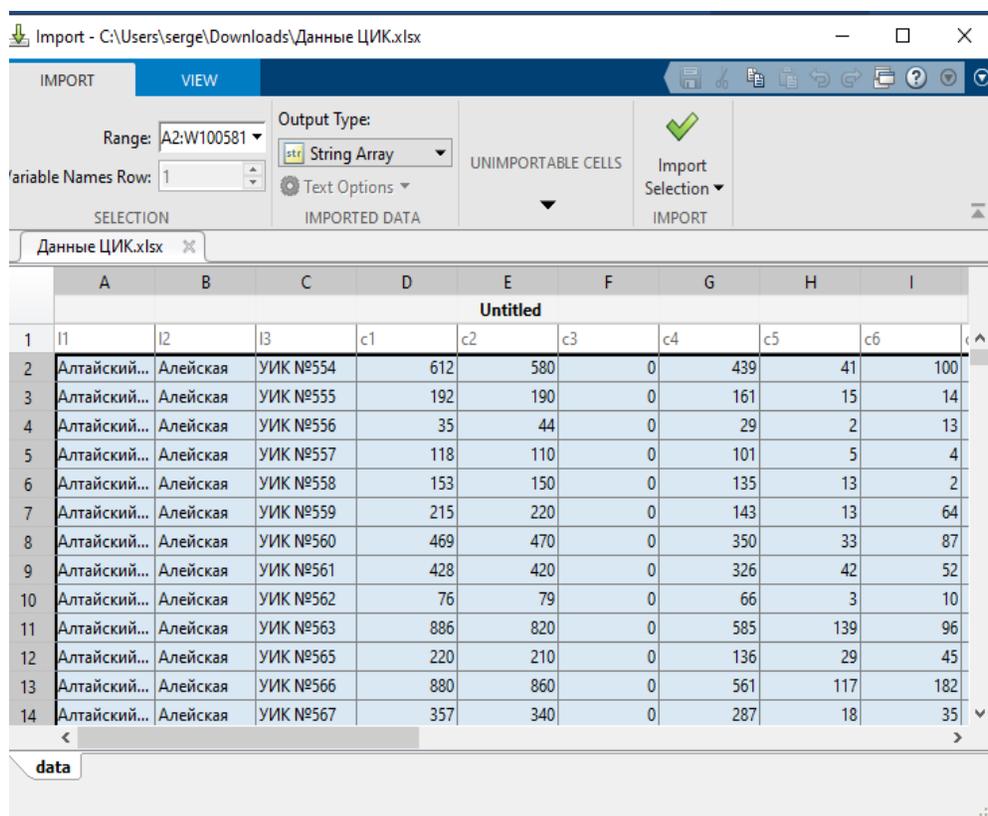


Рис. 5. Фрагмент интерфейса инструмента импорта данных в процессе импорта электоральных данных ЦИК из файла «Данные ЦИК.xlsx»

В результате импорта в рабочем пространстве MATLAB создавалась переменная «Untitled», содержащая электоральные данные ЦИК, которая далее переименовывалась в переменную «Data». Далее были написаны на m-пакета MATLAB 3 скрипта, которые были сохранены в файлах «DataElection.m», «DataElection2.m», «DataElection3.m», соответственно. Листинги данных файлов приведены в Приложении, соответственно. Первый

скрипт обеспечивает автоматическую выгрузку из переменной Data электоральных данных, представленных УИК, сформированными соответствующей ТИК, и их сохранение в папке «d:\Data» в mat-файлах, а также УИК г. Байконур и имена которых формируются в соответствие со следующим шаблоном:

«Название субъекта РФ, сформировавшего ИК Название ТИК».

В данный файл записывается матрица

размерности  $20 \times$  число УИК. Пример данной матрицы, содержащей электоральные данные, представленные УИК г. Байконур, приведен на рис. 6.

Результатом выполнения первого скрипта являются 2778 mat-файлов, размещенных в папке d:\Data (рис. 6).

Для подтверждения правильности рабо-

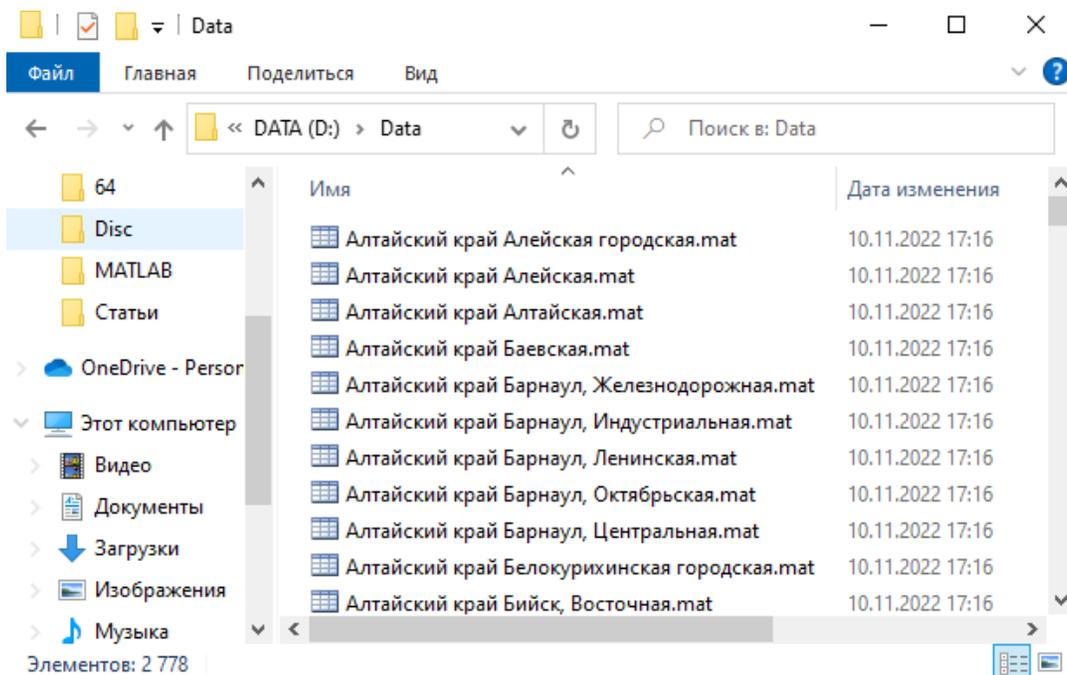


Рис. 6. Фрагмент каталога папки «d:\Data»

	1	2	3	4	5	6	7
1	2132	2207	2249	1769	1880	2047	2291
2	2000	2000	2000	1500	1500	1900	1900
3	0	0	0	0	0	0	0
4	1447	1470	1490	1065	1171	1302	1546
5	11	14	7	48	13	13	62
6	542	516	503	387	316	585	292
7	11	14	7	48	13	13	62
8	1447	1470	1490	1065	1171	1302	1546
9	9	14	27	20	10	14	10
10	1449	1470	1470	1093	1174	1301	1598
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	4	2	5	5	7	1	8
14	176	128	171	98	124	138	191
15	79	87	94	72	105	96	95
16	1136	1214	1162	882	902	1022	1250
17	30	19	17	17	7	18	22
18	9	4	3	8	9	9	9
19	5	7	12	5	10	5	6
20	10	9	6	6	10	12	17

Рис. 7. Данные, загруженные из файла «Город Байконур (Республика Казахстан).mat»

ты скрипта на рисунке 8 представлены аналогичные электоральные данные, находящиеся в первичном файле «Данные ЦИК.xlsx».

Второй скрипт обеспечивает автоматическую выгрузку из переменной Data сводных электоральных данных, представленных в ИК

	A	B	C	D	E	F	G	H
1	I1	I2	I3	c1	c2	c3	c4	c5
12101	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8140	2132	2000	0	1447	11
12102	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8141	2207	2000	0	1470	14
12103	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8142	2249	2000	0	1490	7
12104	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8143	1769	1500	0	1065	48
12105	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8144	1880	1500	0	1171	13
12106	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8145	2047	1900	0	1302	13
12107	Город Байконур (Республика Казахстан)	Без образования ТИК	УИК №8146	2291	1900	0	1546	62
12108	Город Байконур (Республика Казахстан)	Без образования ТИК						
12109	Город Байконур (Республика Казахстан)	Пустое поле		14575	12800	0	9491	168

Рис. 8. Электоральные данные, представленных УИК г. Байконур (Республика Казахстан) (файл «Данные ЦИК.xlsx»)

субъекта РФ соответствующими ТИК<sup>1</sup>, и их сохранение в папке «d:\Data2» в mat-файлах, имена которых формируются в соответствие со следующим шаблоном:

«Название субъекта РФ»,

а также автоматическую выгрузку из переменной Data сводных электоральных, представленных ИК субъектов РФ, УИК г. Байконур (Республика Казахстан) и УИК территорий, на-

ходящихся за пределами РФ и сохранение их в папке d:\Data3 в mat-файле «d:\Data3\All\_Subject.mat», в который записывается таблица ячеек размером 20 × 87 ячеек (рис. 10).

Результатом выполнения данного скрипта являются 87 файлов, размещенных в папке «d:\Data2» (рис. 9) и файл «All\_Subject.mat», размещенный в папке «d:\Data3».

Правильность работы данного скрипта

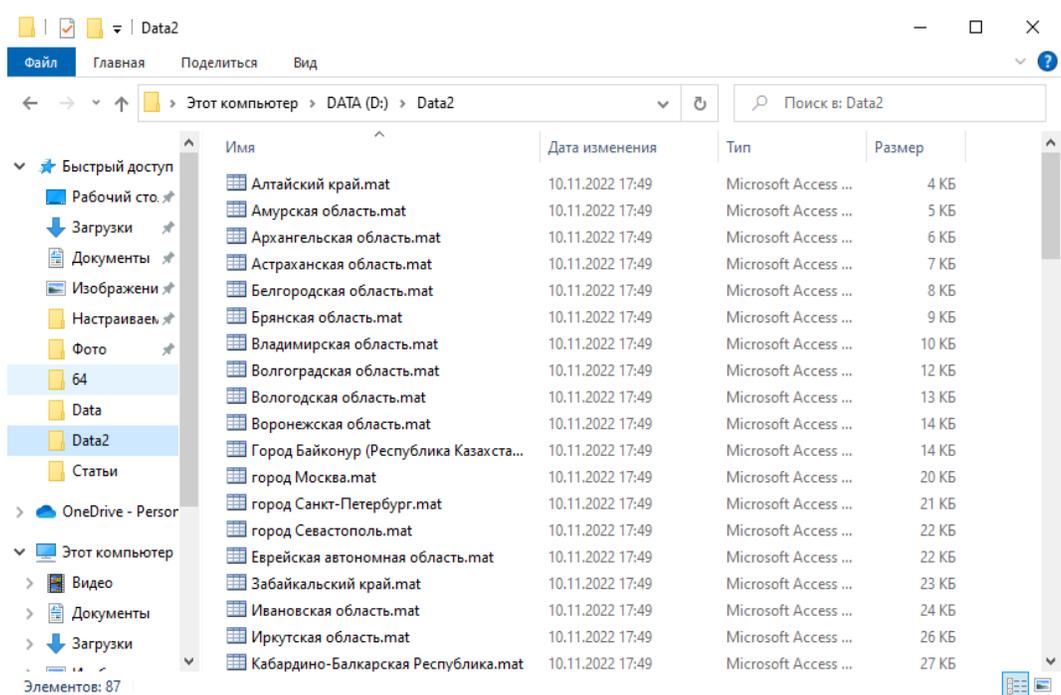


Рис. 9. Фрагмент каталога папки «D:\Data2»

подтверждается результатами сравнения данных, представленных на рис. 10, с соответствующими данными, размещенными в файле «Данные ЦИК.xlsx». Например, на рисунке 10 приведены сводные электоральные данные, представленные ИК Алтайского края.

Из рисунков 10, 11, видно, что электро-

ральные данные, размещенные в столбце № 1 в ячейках № 1-20, совпадают с данным размещенными в файле «d:\Data3\All\_Subject.mat» в строке № 1907 в столбцах c1-c20.

Таким образом, приведенные результаты подтверждают правильность функционирования описанных выше скриптов, предназна-

	1	2	3	4	5	6	7
1	1822203	635083	914218	736206	1218896	978509	1141236
2	1650783	578918	884150	700193	1162467	950080	1066871
3	61	3498	3942	432	0	0	0
4	1106745	365862	514021	417312	806610	692036	678717
5	84824	25321	23229	27074	86165	87912	63247
6	459153	184237	342958	255375	269692	170132	324907
7	84869	28810	27168	27487	86165	87903	63241
8	1106390	365743	513794	417239	806269	691575	678160
9	14203	5383	4880	4340	8749	7177	9351
10	1177056	389170	536082	440386	883685	772301	732050
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	7581	2358	4448	2185	5218	4472	5440
14	281978	73485	51868	64047	93102	68375	93649
15	84785	37909	46925	19339	49685	43940	58822
16	770278	264493	407190	342195	711392	636087	546042
17	11788	4428	10588	5060	8474	7463	10777
18	7855	2466	3842	2823	6534	4265	5075
19	5532	2080	4982	2233	4835	4175	6098
20	7259	1951	6239	2504	4445	3524	6147

Рис. 10. Фрагмент таблицы, сохраненной в файле «d:\Data3\All\_Subject.mat»

	A	B	C	D	E	F	G
1	I1	I2	I3	c1	c2	c3	c4
1907	Алтайский край	Пустое поле		1822203	1650783	61	1106745

Рис. 11. Фрагмент строки со сводными электоральными данными, представленными ИК Алтайского края

ченных для извлечения электоральных данных из файла «d:\Data3\All\_Subject.mat» в соответствие со структурой ИК, ТИК и УИК РФ, проводивших выборы президента РФ в 2018 г.

#### 4. Некоторые результаты анализа электоральных данных ЦИК РФ, собранных в ходе выборов Президента РФ в 2018 г.

Одно из традиционных обвинений ЦИК со стороны ОЭС состоит в искусственном увеличении числа ТИК и УИК, что облегчает по их утверждениям фальсификацию на данных уровнях электоральной статистики (см, например, [9]). В этой связи авторы исследовали функциональные зависимости между числом ТИК, организованных в РФ для проведения выборов Президента РФ в 2018 г., и числом избирателей, отнесенных к данной ТИК (функ-

циональная зависимость № 1), а также числом УИК и числом УИК, отнесенных к данной УИК (функциональная зависимость № 2).

Для анализа функциональной зависимости № 1 на основе анализа электоральных данных, размещенных в папке «D:\Data2», для каждой ИК субъекта РФ подсчитывалось число ТИК, созданных данной ИК, а также число избирателей, зарегистрированных данной ИК. При этом число ТИК, созданных на территориях, находящихся за пределами РФ, и в г. Байконур принималось равными единице. Таким образом, общее число значений функциональной зависимости № 1, представленных на рисунке 12, составило 87 пар значений.

При этом значение коэффициента R2, характеризующего качество линейной аппрок-

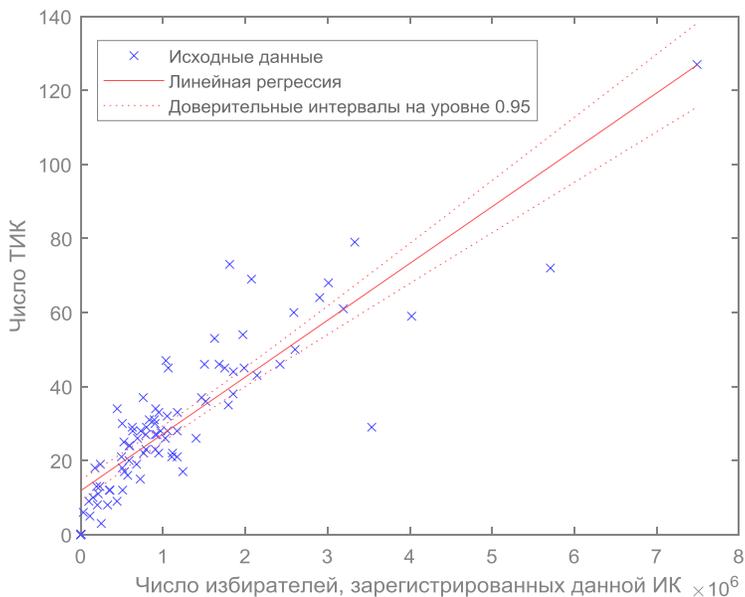


Рис. 12. Функциональная зависимость № 1

симиляции функциональной зависимости № 1, оказалось равным 0.757. Данный результат позволяет сделать обоснованный вывод о том, что ИК субъектов РФ при создании ТИК, стремились следующему следовать принципу: «число ТИК должно быть прямо пропорционально числу избирателей, зарегистрированных ИК данного субъекта РФ». Этот результат, с нашей точки зрения, свидетельствует об отсутствии преднамеренной организации такого числа ТИК, которые непропорционально числу избирателей, с целью последующих фальсификаций электоральных данных.

Для анализа функциональной зависимости № 2 на основе анализа электоральных данных, размещенных в папке «D:\Data», для каждой ТИК, созданной соответствующей ИК субъекта РФ, подсчитывалось число УИК и число избирателей в каждой из УИК. Таким образом, общее число значений функциональной зависимости № 1, представленных на рисунке 13, составило 2778 пар значений.

При этом значение коэффициента  $R^2$ , характеризующего качество линейной аппроксимации функциональной зависимости № 2, оказалось равным 0.654, что также позволяет

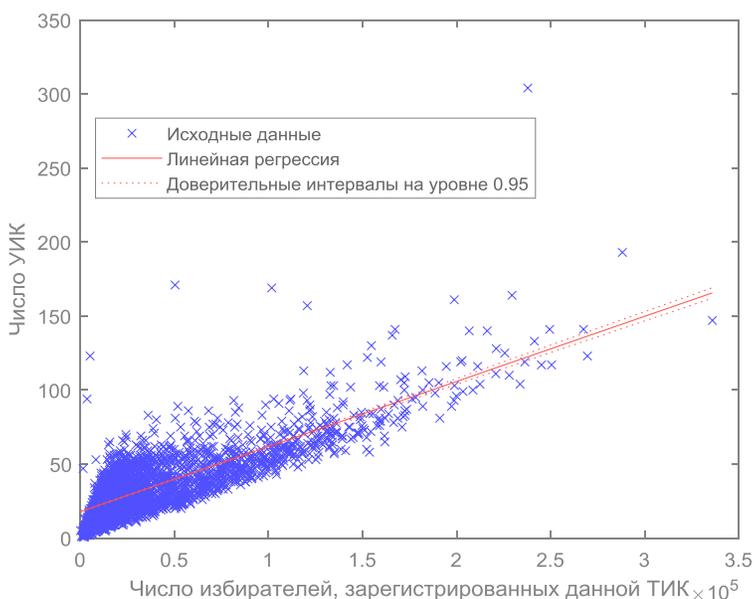


Рис. 13. Функциональная зависимость № 2

говорить о значимой линейной связи между числом УИК и числом избирателей, зарегистрированных данной ТИК. Меньшее, чем в предыдущем случае, значение коэффициента  $R^2$ , с нашей точки зрения, объясняется тем, что при выборе числа УИК, созданных данной ТИК, приходится учитывать не только численность избирателей, но их распределение по территории, относящейся к данной ТИК.

Таким образом, анализ функциональных зависимостей № 1, № 2 позволяет сделать вывод об отсутствии попыток организаторов выборов Президента РФ в 2018 г. произвольного выбора числа ТИК и УИК в субъектах РФ с целью последующей фальсификации электоральных данных.

### Заключение

Разработана автоматизированная методика извлечения электоральных данных из файла «D:\Data3\All\_Subject.mat» в соответствии со структурой ИК, ТИК и УИК РФ, проводивших выборы президента РФ в 2018 г., и ПО, обеспечивающее его работоспособность.

С помощью разработанного ПО реализована выгрузка:

1) электоральных данных, представленных УИК в соответствующие ТИК, созданные ИК каждого из субъектов РФ, а также УИК г. Байконура и УИК территорий расположенных за пределами РФ (2778 файлов);

2) сводных электоральных данных по каждой ТИК, созданные ИК каждого из субъектов РФ, а также УИК г. Байконура и УИК территорий расположенных за пределами РФ (87 файлов);

3) сводных электоральных данных по каждой ИК, созданные ИК каждого из субъектов РФ, а также УИК г. Байконура и УИК территорий расположенных за пределами РФ (1 файл).

Можно ожидать, что наличие электоральных данных, декомпозированных по уровням ИК, ТИК и УИК РФ, обеспечит проведение статистического анализа результатов выборов Президента РФ в 2018 г. и получение объективных оценок адекватности методов выявления аномалий, активно используемых оппозиционными электоральными статистиками.

---

## Литература

1. Доклад Российского общественного института избирательного права (РОИИП) «Математические инструменты делигитимации выборов»//И.Б. Борисов, И.В. Задорин, А.В. Игнатов, В.Н. Марачевский, В.И. Федоров/ – М.: РОИИП, 2020. 76 с. URL: [http://www.roiip.ru/images/data/gallery/0\\_299\\_Matematicheskie\\_instrumenti\\_delegitimatsii\\_viborov.pdf](http://www.roiip.ru/images/data/gallery/0_299_Matematicheskie_instrumenti_delegitimatsii_viborov.pdf) (дата обращения 19.02.2022).
2. Шень А. Как доклад РОИИП делегитимирует выборы. URL: <https://trv-science.ru/2020/09/kak-doklad-roiiip-delegitimiruuet-vybory/> (дата обращения 19.02.2022).
3. URL: <http://electoralpolitics.org/ru/articles/vozmozhnosti-matematicheskikh-metodov-povyavleniiu-elektoralnykh-falsifikatsii/> (дата обращения 19.02.2022).
4. Поршнев С.В., Рябко Н.Ю., Уксусников Н.А. Обеспечение достоверности экзит-поллов, как задача информационной безопасности//Вестник УрФО. Безопасность в информационной сфере, 2021, № 3(41) / 2021, с. 49–56.
5. URL: <http://www.vybory.izbirkom.ru> (дата обращения 19.02.2022).
6. Мирвода С.Г., Поршнев С.В., Рябко Н.Ю. Автоматизация процедуры доступа к электоральным данным, размещенным на сайте Центральной избирательной комиссии Российской Федерации// Вестник УрФО. Безопасность в информационной сфере, 2022, № 1(43). С. 28–34. DOI: 10.14529/secur220104.
7. URL:<http://www.vybory.izbirkom.ru/region/izbirkom?action=show&global=1&vrn=100100084849062&region=0&prver=0&pronetvd=null> (дата обращения 09.11.2022).
8. Поршнев С.В., Рябко Н.Ю. Опыт применения метода Берра и Скакко для проверки достоверности электоральных данных// Вестник УрФО. Безопасность в информационной сфере, 2022, № 2(44), С. 12–24.
9. <https://www.yaplakal.com/forum7/topic2079742.html>.

## References

1. Doklad Rossijskogo obshchestvennogo instituta izbiratel'nogo prava (ROIIP) «Matematicheskie instrumenty deligitimacii vyborov»//I.B. Borisov, I.V. Zadorin, A.V. Ignatov, V.N. Marachevskij, V.I. Fedorov/ – М.: ROIIP, 2020. 76 s. URL: [http://www.roiip.ru/images/data/gallery/0\\_299\\_Matematicheskie\\_instrumenti\\_delegitimatsii\\_viborov.pdf](http://www.roiip.ru/images/data/gallery/0_299_Matematicheskie_instrumenti_delegitimatsii_viborov.pdf) (data obrashcheniya 19.02.2022).

2. SHen' A. Kak doklad ROIIP delegitimiruuet vybory. URL: <https://trv-science.ru/2020/09/kak-doklad-roiiip-delegitimiruuet-vybory/> (data obrashcheniya 19.02.2022).

3. URL: <http://electoralpolitics.ru/articles/vozmozhnosti-matematicheskikh-metodov-po-vyivleniiu-elektoralnykh-falsifikatsii/>(data obrashcheniya 19.02.2022).

4. Porshnev S.V. Ryabko N.YU., Uksusnikov N.A. Obespechenie dostovernosti ekzit-pollov, kak zadacha informacionnoj bezopasnosti//Vestnik UrFO. Bezopasnost' v informacionnoj sfere, 2021, № 3(41) / 2021, s. 49–56.

5. URL: <http://www.vybory.izbirkom.ru> (data obrashcheniya 19.02.2022).

6. Mirvoda S.G., Porshnev S.V., Ryabko N.YU. Avtomatizaciya procedury dostupa k elektoral'nyh dannym, razmeshchennym na sajte Central'noj izbiratel'noj komissii Rossijskoj Federacii// Vestnik UrFO. Bezopasnost' v informacionnoj sfere, 2022, № 1(43). S. 28–34. DOI: 10.14529/secur220104

7. URL:<http://www.vybory.izbirkom.ru/region/izbirkom?action=show&global=1&vrn=100100084849062°ion=0&prver=0&pronetvd=null> (data obrashcheniya 09.11.2022).

8. Porshnev S.V., Ryabko N.YU. Opyt primeneniya metoda Berra i Skakko dlya proverki dostovernosti elektoral'nyh dannyh// Vestnik UrFO. Bezopasnost' v informacionnoj sfere, 2022, № 2(44), S. 12–24.

9. <https://www.yaplakal.com/forum7/topic2079742.html>.

## Приложение

```
% Листинг скрипта № 1
% назначение скрипта:
% выгрузка электоральных данных УИК,
% созданных данной ТИК
% выгрузка электоральных данных
% по всем ТИК, созданных
% ИК субъектов РФ
clear all
load Data.mat Data
N=size(Data,1);
Name_Subject=Data(1,1);
Name_Region=Data(1,2);
Null_Str="";
k=2;
n_min=1;
K=1;
clear('P');
while k<=N-1
    while Name_Region==Data(k,2)
        k=k+1;
        K=K+1;
    end
    K=K-1;
    k=k-1;
    if Data(k,1)=="Город Байконур (Республика Казахстан)" || Data(k,1)=="Территория
за пределами РФ"
        K=K-1;
    end
    P=zeros(20,K);
    for m=1:20
        for n=n_min:n_min+K-1
            P(m,n-n_min+1)=Data(n,m+3);
        end
    end
    File_Name="d:\Data\"+string(Name_Subject)+" "+string(Name_Region)+".mat";
    save(File_Name,'P');
    tmp1=Data(k,3);
    tmp2=Data(k+1,2);
    if (tmp1==Null_Str)&&~(tmp2==Null_Str)
        k=k+1;
        K=0;
    end
    if (tmp1==Null_Str)&&(tmp2==Null_Str)
        Name_Subject=Data;
        Name_Region=Data;
```

```

        k=k+2;
        K=0;
    end
    n_min=k;
    if k<=N
        Name_Subject=Data(k,1);
        Name_Region=Data(k,2);
    end
end

% Листинг скрипта № 2
% назначение скрипта:
% выгрузка в файлы ИК субъектов РФ
% электоральных данных,
% представленной каждой ТИК,
% а также сводных данных
% по всем ИК субъектов РФ
clear all
load Data.mat Data
N=size(Data,1);
Null_Str="";
k=2;
n_min=1;
K1=1;
K2=1;
for i=2:N-1
    if (Data(i,3)==Null_Str)&&~(Data(i,2)==Null_Str)&&~(Data(i,1)=='Город
Байконур (Республика Казахстан)')&&~(Data(i,1)=='Территория за пределами
РФ')
        P1(K1,:)=Data(i,:);
        K1=K1+1;
    end
    if (Data(i,1)=='Город Байконур (Республика
Казахстан)')&&(Data(i,3)==Null_Str)
        if ~(Data(i,4)==Null_Str)
            P1(K1,:)=Data(i,:);
            K1=K1+1;
        end
    end
    if (Data(i,1)=='Территория за пределами РФ')&&(Data(i,3)==Null_Str)
        if ~(Data(i,4)==Null_Str)
            P1(K1,:)=Data(i,:);
            K1=K1+1;
        end
    end
    if (Data(i,3)==Null_Str)&&(Data(i,2)==Null_Str)&&~(Data(i,1)=='Город
Байконур (Республика Казахстан)')&&~(Data(i,1)=='Территория за пределами
РФ')
        P2(K2,:)=Data(i,:);
        K2=K2+1;
    end
    if (Data(i,1)=='Город Байконур (Республика
Казахстан)')&&(Data(i,3)==Null_Str)
        if ~(Data(i,4)==Null_Str)
            P2(K2,:)=Data(i,:);
            K2=K2+1;
        end
    end
    if (Data(i,1)=='Территория за пределами РФ')&&(Data(i,3)==Null_Str)
        if ~(Data(i,4)==Null_Str)
            P2(K2,:)=Data(i,:);
            K2=K2+1;
        end
    end
end
end

```

```

end
end
if Data(i+1,1)=="Ярославская область"
    P2=[P2;Data(N,:)];
    for k=1:size(P2,1)
        for n=4:size(P2,2)
            P2_Num(k,n-3)=str2num(P2(k,n));
        end
    end
    File_Name="D:\Data3\\"+"All_Subject"+'.mat';
    save(File_Name,'P2_Num')
end
M=size(P1,1);
K1=1;
m=1;
while m<=M-1
    if P1(m,1)=="Город Байконур (Республика
Казахстан)"||P1(m,1)=="Территория за пределами РФ"
        Line=P1(m,4:23);
        for i=1:20
            P1_Num(i,K1)=Line(1,i);
        end
        K1=K1+1;
        File_Name="D:\Data2\\"+' '+P1(m,1)+'.mat';
        save(File_Name,'P1_Num')
    end
    if (P1(m,1)=="Ненецкий автономный округ")
        Line=P1(m,4:23);
        for i=1:20
            P1_Num(i,K1)=str2num(Line(1,i));
        end
        K1=K1+1;
        m=m+1;
        Line=P1(m,4:23);
        for i=1:20
            P1_Num(i,K1)=str2num(Line(1,i));
        end
    end
    if P1(m,1)=="Псковская область"
        Line=P1(m,4:23);
        for i=1:20
            P1_Num(i,K1)=str2num(Line(1,i));
        end
        K1=K1+1;
    end
    if ~ (P1(m,1)=="Псковская область")
        File_Name="D:\Data2\\"+' '+P1(m,1)+'.mat';
        save(File_Name,'P1_Num')
        clear('P1_Num')
        K1=1;
    end
    if P1(m+1,1)=="Ярославская область"&&P1(m+1,2)=="Ярославская"
        Line=P1(m+1,4:23);
        for i=1:20
            P1_Num(i,K1)=str2num(Line(1,i));
        end
        File_Name="D:\Data2\\"+' '+P1(m,1)+'.mat';
        save(File_Name,'P1_Num')
    end
    m=m+1;
end

```

---

**ПОРШНЕВ Сергей Владимирович**, доктор технических наук, профессор, директор Учебно-научного центра «Информационная безопасность» федерального государственного авто-

номного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: s.v.porshnev@urfu.ru

**РЯБКО Николай Юрьевич**, аспирант федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: N.Yu.Ryabko@urfu.ru

**PORSHNEV Sergey Vladimirovich**, Doctor of Technical Sciences, Professor, Director of the Educational and Scientific Center «Information Security» of the Federal State Autonomous Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N. Yeltsin». 620002, Yekaterinburg, st. Mira, 32. E-mail: s.v.porshnev@urfu.ru

**RYABKO Nikolay Yurievich**, post-graduate student of the Federal State Autonomous Educational Institution of Higher Education "Ural Federal University named after the first President of Russia B.N. Yeltsin". 620002, Yekaterinburg, st. Mira, 32. E-mail: N.Yu.Ryabko@urfu.ru