

Входит в состав базы Russian Science Citation Index (RSCI) на платформе Web of Science

ISSN 0033-8486

# РАДИОТЕХНИКА

Radioengineering

XXI век

**6 2018**

## В номере:

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ  
ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ  
№ 1 (выпуск 7)

и др.



тел./факс: (495) 625-9241  
e-mail: [info@radiotec.ru](mailto:info@radiotec.ru)  
<http://www.radiotec.ru>

ПОДПИСНОЙ ИНДЕКС 70775 В КАТАЛОГЕ АГЕНТСТВА «РОСПЕЧАТЬ»: ГАЗЕТЫ И ЖУРНАЛЫ

Международный научно-технический журнал  
Включен в состав базы Russian Science Citation Index (RSCI) на платформе Web of Science

ISSN 0033-8486

Выходит с февраля 1937 года



# РАДИОТЕХНИКА

Radioengineering

Включен в перечень ВАК

XXI век

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Л.П. Андрианова, д.т.н., проф. П.А.Бакулев, д.ф.-м.н., проф. О.В. Бецкий, д.т.н., проф. А.В. Богословский, д.т.н., проф. А.Л. Бузов, д.т.н., проф. С.А. Букашкин, д.т.н., проф. Р.П. Быстров, чл.-корр. РАН В.С. Верба, д.т.н., проф. В.В. Витязев, д.т.н., проф. Э.А. Засовин, д.т.н., доцент А.В. Иванов, д.т.н., проф. Ю.Л. Козирацкий, д.ф.-м.н., проф. А.Г. Козорезов (Великобритания), к.ф.-м.н., с.н.с. В.В. Колесов (зам. главного редактора), д.ф.-м.н., проф. И.Н. Компанец, д.т.н., проф. Г.С. Кондратенков, д.т.н., проф. А.В. Коренной, д.ф.-м.н., проф. Б.Г. Кутуза, д.т.н., проф. В.И. Меркулов, д.т.н., проф. В.П. Мещанов, д.т.н., проф. В.А. Обуховец, д.т.н., проф. А.И. Перов, д.ф.-м.н. В.П. Плесский (Швейцария), д.ф.-м.н., проф. В.В. Проклов, акад. РАН В.И. Пустовойт, д.т.н., проф. В.Г. Радзиевский, д.т.н., проф. Е.М. Сухарев, д.т.н., проф. Е.Ф. Толстов, д.т.н., проф. В.Н. Ушаков, д.т.н., проф. В.Е. Фарбер, акад. РАН И.Б. Фёдоров, д.т.н., проф. В.А. Цимбал, д.т.н., проф. Ю.С. Шинаков, в.н.с. Энрико Верона (Италия), д.т.н., проф. С.В. Ягольников, д.т.н., проф. М.С. Ярлыков.

## EDITORIAL BOARD:

L.P. Andrianova, Academician RAS I.B. Fedorov, Academician RAS V.I. Pustovoyt, Dr.Sc. (Eng.), Corresponding Member RAS V.S. Verba, Prof. P.A. Bakulev, Dr.Sc. (Phys.-Math.), Prof. O.V. Betskii, Dr.Sc. (Eng.), Prof. A.V. Bogoslovskii, Dr.Sc. (Eng.), Prof. A.L. Buzov, Dr.Sc. (Eng.), Prof. S.A. Bukashkin, Dr.Sc. (Eng.), Prof. R.P. Bystrov, Dr.Sc. (Eng.), Prof. V.Ye. Farber, Dr.Sc. (Phys.-Math.), Prof. I.N. Kompanets, Dr.Sc. (Eng.), Prof. G.S. Kondratenkov, Dr.Sc. (Eng.), Prof. A.V. Korennoi, Dr.Sc. (Eng.), Prof. Yu.L. Koziratskii, Dr. Sc. (Phys.-Math.), Prof. A.G. Kozorezov (Great Britain), Dr.Sc. (Phys.-Math.), Prof. B.G. Kutuza, Dr.Sc. (Eng.), Prof. V.I. Merkulov, Dr.Sc. (Eng.), Prof. V.P. Meshchanov, Dr.Sc. (Eng.), Prof. V.A. Obukhovets, Dr.Sc. (Eng.), Prof. V.P. Plessky (Switzerland), Dr.Sc. (Eng.), Prof. A.I. Perov, Dr.Sc. (Phys.-Math.), Prof. V.V. Proklov, Dr.Sc. (Eng.), Prof. V.G. Radzievskii, Dr.Sc. (Eng.), Prof. Yu.S. Shinakov, Dr.Sc. (Eng.), Prof. Ye.M. Sukharev, Dr.Sc. (Eng.), Prof. Ye.F. Tolstov, Dr.Sc. (Eng.), Prof. V.A. Tsimbal, Dr.Sc. (Eng.), Prof. V.N. Ushakov, Dr.Sc. (Eng.), Dr.Sc. (Eng.), Prof. V.V. Vityazev, Dr.Sc. (Eng.), Prof. S.V. Yagolnikov, Dr.Sc. (Eng.), Prof. M.S. Yarlykov, Dr.Sc. (Eng.), Prof. E.A. Zasovin, Dr.Sc. (Eng.), Associate Prof. A.V. Ivanov, associate researcher Enrico Verona (Italy), Ph.D. (Phys.-Math.), Senior Research Scientist V.V. Kolesov (Deputy Editor).

Главный редактор  
академик РАН  
Ю. В. Гуляев

Editor-in-Chief,  
Academician RAS,  
Yu.V. Gulyaev

## Содержание

№ 6 июнь 2018 г.

### МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ

(Редакторы выпуска – д.т.н., проф. К.К. Васильев и д.т.н., проф. Н.Г. Ярушкина)

#### Статистический синтез и анализ инфокоммуникационных систем

Дважды стохастические модели цилиндрических изображений

Крашенинников В.Р., Субботин А.Ю.

5

Оптимизация оценивания взаимной информации двух изображений

Сафина Г.Л., Таплинский А.Г., Царев М.Г.

9

Псевдоградиентный алгоритм оценивания параметров привязки изображений на основе взаимной информации

Ташлинский А.Г., Воронов С.В., Жукова А.В.

14

Эффективность фильтрации случайных полей с кратными корнями характеристических уравнений

Васильев К.К., Андриянов Н.А., Абдулкадим Х.А.

20

Обработка пространственно-неоднородных изображений с помощью дважды стохастического фильтра

Дементьев В.Е.

24

Ульяновский  
государственный  
технический университет  
Научная библиотека



Применение сверточных нейронных сетей для тематического картографирования последовательностей спутниковых многозональных изображений Дементьев В.Е., Кондратьев Д.С., Френкель А.Г.	29
Выделение на металлографических изображениях малоуглеродистой стали зерен перлита Магдеев Р.Г., Таплинский А.Г.	33
Минимизация объема памяти когнитивной карты декодера в системе поиска эквивалентных кодов Гладких А.А., Пчелин Н.А., Шаханов С.В.	38
Анализ эффективности бэггинга для бинарной классификации при технической диагностике Кувайскова Ю.Е., Клячкин В.Н.	42
Диагностика технического состояния аппаратуры с использованием агрегированных классификаторов Клячкин В.Н., Кувайскова Ю.Е., Жуков Д.А.	46

#### Системы автоматизации проектирования

Исследование точности алгоритма восстановления данных на основе нечеткой кластеризации Афанасьева Т.В., Сибирев И.В.	50
Использование нейронной сети для выбора методов прогнозирования временного ряда в гибридной комбинированной модели Яшин Д.В., Эгов Е.Н.	54
Подход к трансформации кластерного дерева признаков в векторное пространство признаков Дударин П.В., Ярушкина Н.Г.	63
Разработка программной системы семантического анализа контента социальных медиа Ярушкина Н.Г., Мошкин В.С., Филиппов А.А., Гуськов Г.Ю., Романов А.А., Наместников А.М.	73
Структурно-семантическое представление конструкторских решений в САПР Цыганков Д.Э., Похилько А.Ф.	80
Контроль и анализ денотативных и сигнификативных семантических ошибок диаграмматических моделей потоков работ в проектировании автоматизированных систем Афанасьев А.Н., Войт Н.Н., Уханова М.Е.	84
Виртуальный полигон зенитных комплексов Афанасьев А.Н., Гульшин В.А., Канев Д.С., Войт Н.Н.	93

#### Измерительные и диагностические методы и комплексы

Измерительный комплекс для контроля характеристик submodule выходных усилителей мощности приемо-передающих модулей АФАР X-диапазона Тарасов Р.Г., Сергеев В.А.	98
Релаксационные излучательные процессы в тонкопленочных электролюминесцентных структурах Самохвалов М.К.	103

#### СИНТЕЗ И АНАЛИЗ РАДИОТЕХНИЧЕСКИХ СИСТЕМ И УСТРОЙСТВ

Современные средства высокоточного измерения разности потенциалов между узлами электрической цепи Аракчеев П.В., Безделов В.Л., Бурый Е.В., Ефремов И.А., Орлов Н.Е.	109
---	-----

#### МАТЕМАТИЧЕСКИЕ МЕТОДЫ В РАДИОТЕХНИКЕ

Оптимальный прием радиосигнала с квадратурной фазовой модуляцией с временным сдвигом Симонгуз В.И.	113
---	-----

#### АЛГОРИТМЫ ФУНКЦИОНИРОВАНИЯ: ПЕРЕДАЧА, ПРИЕМ И ОБРАБОТКА СИГНАЛОВ

Робастный подход при синтезе адаптивных блоков нелинейного преобразования с учетом конечного отношения сигнал/шум на входе Артищенко В.М., Воловач В.И.	129
--	-----



## Исследование точности алгоритма восстановления данных на основе нечеткой кластеризации

© Авторы, 2018

© ООО «Издательство «Радиотехника», 2018

**Т.В. Афанасьева** – д.т.н., профессор, доцент, кафедра «Информационные системы», Ульяновский государственный технический университет

E-mail: tv.afanasjeva@gmail.com

**И.В. Сибирев** – аспирант, кафедра «Информационные системы», Ульяновский государственный технический университет

E-mail: ivan.sibirev@yandex.ru

Описан и экспериментально исследован алгоритм восстановления пропусков в числовых данных на основе нечеткой кластеризации с целью оценивания точности алгоритма на искусственных данных с различным числом пропусков. Показано, что применение нечеткой кластеризации для заполнения пропусков данных имеет преимущество по точности в сравнении с алгоритмом на основе среднеарифметического.

**Ключевые слова:** алгоритм восстановления данных, нечеткая кластеризация, точность восстановления.

The article describes and experimentally investigates the algorithm recovery of omissions in numerical data based on fuzzy clustering. The aim of the study is to obtain the accuracy of the algorithm on artificial data with different number of omissions. It is shown that the use of fuzzy clustering to fill data gaps has an advantage in accuracy in comparison with the algorithm based on the arithmetic mean.

**Keywords:** data recovery algorithm, fuzzy clustering, estimate of data recovery.

Одной из задач в области обработки телеметрической информации об исследуемых объектах, полученной с датчиков, является обработка искажений, приводящая к проблеме пропусков и последующего восстановления данных. Хранимые в базах данных наблюдения за объектами обычно представлены наборами их числовых параметров, которые после первичной обработки искажений могут содержать пропуски, вызванные удалением выбросов или других аномальных значений. Данная проблема встречается и в других предметных областях. Так, А. Mockus в работе [1] отмечает, что при анализе данных из репозитория очистка, проверка и восстановление данных занимают 95% усилий, а непосредственно анализ – от 1% до 5%.

**Ц е л ь р а б о т ы** – оценить точность алгоритма восстановления пропусков в числовых данных на основе нечеткой кластеризации, используя искусственные данные с различным числом пропусков.

Распространенным приемом при обработке данных с пропусками является редукция (исключение строк и столбцов, имеющих пропуски). Также используется генерация новых данных, простейшими приемами которой являются заполнение пропусков среднеарифметическими значениями, а также использование для заполнения методов регрессии или главных компонент. Программный пакет Statistica предлагает именно такие способы [3]. Исследователи отмечают, что эти методы не всегда эффективны и нередко ведут к искажению статистических выводов. Обзор методов восстановления данных и обширная библиография приведены в монографии Р.Дж. А. Литтл, Д.Б. Рубина «Статистический анализ с пропусками» [4]. Авторы выделяют четыре пересекающиеся группы методов, среди которых редукция и «методы с заполнением», а также методы взвешивания и методы, основанные на моделировании [4].

В последнее время интерес к теме восстановления данных переместился в сферу машинного обучения и приложения известных методов [5]. В работе M. Busse, P. Orbanz, M. Buhmann [6] гетерогенные данные переводятся в числовой вид с присвоением рангов, недостающие значения рангов восстанавливаются с помощью EM-алгоритма. В работе А.В. Масальских [7] в задачах компьютерной графики и картографии для восстановления данных используется сплайновая интерполяция. В работе M.M. Rahman, D.N. Davis [8] восстановление недостающих данных производится с использованием не-



четкого ввода данных, методов машинного обучения, метода к-средних для заключительной классификации.

В работах [9–11] предложен алгоритм восстановления данных на основе нечеткого кластерного анализа. Основная идея алгоритма – итеративное восстановление пропущенных значений параметров обрабатываемых объектов, как линейной комбинации средних значений параметров по кластерам с весовыми коэффициентами принадлежности объектов к кластерам, которые вычисляются с помощью нечеткой кластеризации FCM-методом [5]. Предложенный в [9] алгоритм восстановления данных не зависит от вида статистических распределений и сохраняет работоспособность при большом числе пропущенных данных (при репрезентативности оставшихся). Ниже кратко рассмотрим основные шаги этого алгоритма.

#### Алгоритм восстановления пропущенных данных на основе нечеткой кластеризации

Обозначим множество исследуемых объектов с их параметрами в виде числовой матрицы  $X_0$ , в которой  $X_0(i, j)$  – ячейка данных, соответствующая объекту  $A_i$  и параметру  $P_j$ ,  $i = 1, \dots, I$ ,  $j = 1, \dots, J$ ,  $I$  – число объектов,  $J$  – число параметров.

Перечислим основные шаги алгоритма восстановления данных на основе нечеткой кластеризации [9].

Шаг 0. Введем матрицу исходных данных  $X_0$ , предварительно удалив полностью не заполненные объекты и параметры.

Шаг 1. Получим числовую матрицу  $X_1$ , в которой все пропущенные ячейки заполнены начальными данными – средним арифметическим по параметру среди заполненных значений.

$$X_1(i, j) = \text{GetMatrixWithValues}(X_0).$$

Шаг 2. Произведем первичную кластеризацию одним из четких методов (центроидным, Варда или др.) с использованием критерия нахождения числа кластеров  $K$  – число кластеров должно быть максимально возможным; в каждом кластере для каждого параметра должно быть хотя бы одно заполненное значение  $K = \text{Get } K(X_1)$ . Обозначим через  $C2$  результат кластеризации на  $K$  кластеров:

$$C2 = \text{ClusteringCentroid}(X_1, K).$$

Шаг 3. Выполним нечеткую кластеризацию FCM-методом на  $K$  кластеров для определения степеней принадлежности объектов  $A_i$  ( $i = 1, \dots, I$ ) кластерам (см. табл. 1):

$$\text{FCM} = \text{FCM}(C2, X_1).$$

Шаг 4. Построим таблицу среднеарифметических значений параметра по каждому кластеру (см. табл. 2):

$$\text{CP} = \text{CP}(C2, X_1),$$

и получим матрицу  $\text{CP} = \{\text{CP}(i, j)\}$  размерности  $K \times J$ . Элемент матрицы  $\text{CP}(i, j)$  – это среднеарифметическое значений параметра  $P_j$  в кластере  $C2_i$ .

Шаг 5. Восстановим данные. Новое значение параметра восстановим в виде взвешенной суммы значений параметра по каждому кластеру с весовыми коэффициентами из табл. 2.

$$\tilde{X}_1 = \text{RecoveryMissingQuestionnairesData}(X_1, \text{FCM}, \text{CP}).$$

После расчета  $\tilde{X}_1$  для поддержания итеративности процесса полагаем  $X_1 = \tilde{X}_1$ . Затем переходим к шагу 2. Метод итеративный, число итераций составляет обычно от 2 до 5. На каждой итерации восстанавливаемые значения уточняются, значения  $K$ ,  $\tilde{X}_1$  пересматриваются. Восстанавливаем значения с индексами  $(i, j)$  в матрице  $X_1$  как среднее значение по всем итерациям.

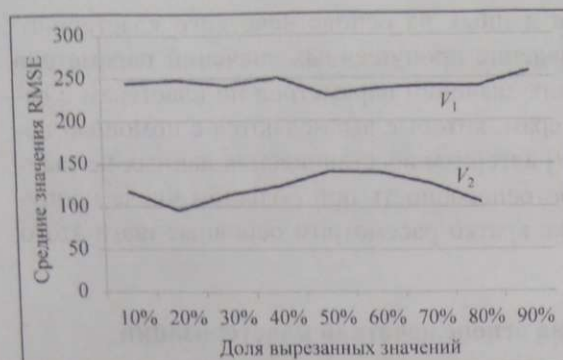
Таблица 1. FCM-таблица

FCM	$C2_1$	$C2_2$	...	$C2_K$
$A_1$	1	0	...	0
$A_2$	0,2	0,3	...	0,1
...	...	...	...	...
$A_I$	0,5	0,1	...	0,2

Таблица 2. Таблица среднеарифметических значений параметра по каждому кластеру

CP	$P_1$	$P_2$	...	$P_J$
$C2_1$	1,5	0,3	...	72,4
$C2_2$	3,2	0,7	...	63,1
...	...	...	...	...
$C2_K$	16,1	0,5	...	54,2

## Анализ точности алгоритма восстановления данных



Графическое сравнение значений RMSE для различных способов восстановления данных

Для оценки точности восстановления пропущенных данных в числовом массиве используем квадратный корень из среднеквадратичной ошибки RMSE:

$$RMSE = \sqrt{MSE}, \quad MSE = \frac{1}{N} \sum_{i=1}^N |Z - \hat{Z}|^2,$$

где  $Z$  — фактическое значение рассматриваемой величины;  $\hat{Z}$  — восстановленное значение рассматриваемой величины;  $N$  — число восстановленных значений.

Цель эксперимента сравнить оценки восстановления данных предлагаемым методом с оценками при заполнении пропусков данных средними арифметическими значениями по параметрам.

Входные данные эксперимента: число объектов 50, каждый объект характеризуется десятью числовыми параметрами, представленными числовой матрицей размерности  $50 \times 10$ . Значения матрицы заполнены искусственно сгенерированными числами (от 0 до 1000) с выраженным разделением на кластеры с помощью авторской программы [12]. Число кластеров для первоначальной кластеризации  $K = 5$ . В серии экспериментов из матрицы произвольным образом вырезались 10%, 20%, ..., 90% данных. Каждый эксперимент, например, с вырезанием 10%, повторялся шесть раз. По шести повторам эксперимента вычислялось среднее арифметическое значение оценок точности RMSE.

На рисунке приведены изменения значений RMSE для различных решений задачи восстановления пропусков в данных. По оси  $OX$  отложены проценты вырезанных числовых значений от 10% до 90%, по оси  $OY$  — средние значения RMSE по различным экспериментам. График  $V_1$  характеризует заполнение пропусков средними арифметическими по столбцам, а  $V_2$  — изменения RMSE для предложенного алгоритма восстановления на основе нечеткой кластеризации, при этом в качестве восстановленного значения берется среднее арифметическое по всем итерациям. Анализ приведенных графиков показывает, что точность алгоритма восстановления числовых параметров исследуемых объектов на основе нечеткой кластеризации дает результат, лучший, чем заполнение пропусков средним арифметическим по параметру. Полученные оценки точности исследуемого алгоритма показывают заметное улучшение по критерию RMSE в 1,6–2 раза в диапазоне от 10% до 80% вырезанных данных по сравнению с методом на основе среднеарифметического по параметру.

- Экспериментальные исследования алгоритма восстановления значений числовых параметров объектов на основе нечеткой кластеризации показали приемлемую точность. Алгоритм работоспособен при больших процентах вырезанных данных и имеет ошибку восстановления по критерию RMSE в 1,6–2 раза меньшую, чем при использовании среднего арифметического для каждого параметра.

Исследование выполнено при финансовой поддержке РФФИ в рамках проекта № 16-07-00535.

## Литература

1. Mockus A. How to run empirical studies using project repositories. — Avaya Labs, 2006. URL = <http://www.research.avayalabs.com/user/audris> (от 24.02.17).
2. Алгоритм восстановления данных по файловым «сигнатурам». URL = <https://spark.ru/startup/hetmansoftware/blog/9391/algoritm-vosstanovleniya-dannih-po-fajlovim-signaturam> (от 4.04.18).
3. Халафян А.А. Statistica 6. Статистический анализ данных: Учебник. М.: ООО «Бином-Пресс», 2007. 512 с.
4. Литтл Р.Дж.А., Рубин Д.Б. Статистический анализ данных с пропусками: Пер. с англ. М.: Финансы и статистика, 1990. 336 с.
5. Вятченин Д.А., Нечеткие методы автоматической классификации. Минск: УП «Технопринт», 2004. 219 с.
6. Busse M., Orbanz P., Buhmann M. Cluster Analysis of Heterogeneous Rank Data. URL = <http://www.machinelearning.org/proceedings/icml2007/papers/341.pdf> (от 24.02.17).
7. Масальских А.В. Параллельный алгоритм одного метода восстановления табличных данных // Известия Тульского государственного университета. Естественные науки. 2014. № 3. С. 67–77.



8. *Rahman M.M., Davis D.N.* Fuzzy Unordered Rules Induction Algorithm Used as Missing Value Imputation Methods for K-Mean Clustering on Real Cardiovascular Data. URL = [http://www2.dcs.hull.ac.uk/NEAT/dnd/papers/Fuzzy\\_InductionAlgorithm\\_MVI.pdf](http://www2.dcs.hull.ac.uk/NEAT/dnd/papers/Fuzzy_InductionAlgorithm_MVI.pdf) (от 4.05.18).
9. *Сибирев И.В., Афанасьева Т.В.* Алгоритм предобработки и восстановления анкетных данных // Материалы VI Междунар. науч.-техн. конф. «Открытые семантические технологии проектирования интеллектуальных систем» (Open Semantic Technologies for Intelligent Systems, OSTIS-2016). Минск: БГУИР, 2016. С. 271–274.
10. *Сибирев И.В.* Предобработка данных в интеллектуальном анализе на основе восстановления пропущенных анкетных значений // Труды Пятнадцатой национальной конф. по искусственному интеллекту с Междунар. участием КИИ-2016. В 3-х томах. Смоленск: Универсум, 2016. Т. 1. С. 378–386.
11. *Сибирев И.В., Афанасьева Т.В.* Анализ эффективности алгоритма восстановления анкетных данных // Сб. научных трудов III Междунар. научно-практическая конф. «Электронное обучение в непрерывном образовании (ЭОНО-2016)». Ульяновск: УлГТУ, 2016. 367–373 с.
12. *Сибирев И.В.* Программа генератор исходных данных для кластеризации / Нечеткие системы и мягкие вычисления. Промышленные применения // Сб. науч. трудов IV Всерос. научно-практич. мультikonф. с Междунар. участием «Прикладные информационные системы (ПИС-2017)». Ульяновск (Россия). 29–31 мая 2017. Ульяновск: УлГТУ, 2017. С. 171–174.

Поступила 24 мая 2018 г.

## Accuracy study of data recovery technique based on fuzzy clustering

© Authors, 2018  
© Radiotekhnika, 2018

**T.V. Afanasyeva** – Dr.Sc.(Eng.), Professor, Associate Professor, Department «Information Systems», Ulyanovsk State Technical University  
E-mail: [tv.afanasjeva@gmail.com](mailto:tv.afanasjeva@gmail.com)

**I.V. Sibirev** – Post-graduate Student, Department «Information Systems», Ulyanovsk State Technical University  
E-mail: [ivan.sibirev@yandex.ru](mailto:ivan.sibirev@yandex.ru)

One of the problem of data mining is pre-processing of data, which are often presented as numerical arrays with omissions (not all parameters are available, some parameters are lost or not removed; not all parameters are credible). The article describes an algorithm for data recovery. It is based on the method of cluster analysis, in particular, fuzzy clustering (FCM-technique). The main idea of the algorithm is the iterative recovery of missing numerical values of parameters as a linear combination of mean values of the parameter in clusters with the weighting coefficients of objects belonging to clusters. Weighting coefficients obtained by using FCM algorithm of fuzzy clustering. Experimental studies have shown that the algorithm does not depend on the type of statistical distributions. It works, among other cases, for large data sets. The algorithm has an acceptable estimate of data recovery. It is shown that the use of fuzzy clustering to fill data gaps has an advantage in accuracy in comparison with the algorithm based on the arithmetic mean.

### References

1. *Mockus A.* How to run empirical studies using project repositories. – Avaya Labs, 2006. URL = <http://www.research.avayalabs.com/user/audris> (от 24.02.17).
2. Алгоритм восстановления данных по файловым «подписям». URL = <https://spark.ru/startup/hetmansoftware/blog/9391/algorithm-vosstanovleniya-dannyh-po-fajlovim-signaturam> (от 4.04.18).
3. *Xalafyan A.A.* Statistica 6. Statisticheskij analiz dannyx: Uchebnik. M.: OOO «Binom-Press». 2007. 512 s.
4. *Littl R.Dzh.A., Rubin D.B.* Statisticheskij analiz dannyx s propuskami: Per. s angl. M.: Finansy i statistika. 1990. 336 s.
5. *Vyatchenin D.A.* Nечеткие методы автоматической классификации. Минск: UP «Texnoprint». 2004. 219 s.
6. *Busse M., Orbanz P., Buhmann M.* Cluster Analysis of Heterogeneous Rank Data. URL = <http://www.machinlearning.org/proceedings/icml2007/papers/341.pdf> (от 24.02.17).
7. *Masalskix A.V.* Parallel'nyj algoritm odnogo metoda vosstanovleniya tablitsnyx dannyx // Izvestiya Tul'skogo gosudarstvennogo universiteta. Estestvenny'e nauki. 2014. № 3. S. 67–77.
8. *Rahman M.M., Davis D.N.* Fuzzy Unordered Rules Induction Algorithm Used as Missing Value Imputation Methods for K-Mean Clustering on Real Cardiovascular Data. URL = [http://www2.dcs.hull.ac.uk/NEAT/dnd/papers/Fuzzy\\_InductionAlgorithm\\_MVI.pdf](http://www2.dcs.hull.ac.uk/NEAT/dnd/papers/Fuzzy_InductionAlgorithm_MVI.pdf) (от 4.05.18).
9. *Sibirev I.V., Afanas'eva T.V.* Алгоритм предобработки и восстановления анкетных данных // Материалы VI Междунар. науч.-техн. конф. «Открытые семантические технологии проектирования интеллектуальных систем» (Open Semantic Technologies for Intelligent Systems, OSTIS-2016). Минск: БГУИР, 2016. С. 271–274.
10. *Sibirev I.V.* Предобработка данных в интеллектуальном анализе на основе восстановления пропущенных анкетных значений // Труды Пятнадцатой национальной конф. по искусственному интеллекту с Междунар. участием КИИ-2016. В 3-х томах. Смоленск: Универсум, 2016. Т. 1. С. 378–386.
11. *Sibirev I.V., Afanas'eva T.V.* Анализ эффективности алгоритма восстановления анкетных данных // Сб. научных трудов III Междунар. научно-практической конф. «Электронное обучение в непрерывном образовании (ЭОНО-2016)». Ульяновск: УлГТУ, 2016. 367–373 с.
12. *Sibirev I.V.* Программа генератор исходных данных для кластеризации / Нечеткие системы и мягкие вычисления. Промышленные применения // Сб. научных трудов IV Всерос. научно-практической мультikonференции с Междунар. участием «Прикладные информационные системы (ПИС-2017)». Ульяновск (Россия). 29–31 мая 2017. Ульяновск: УлГТУ, 2017. С. 171–174.