

Зырянов Александр Олегович, Павловский Евгений Николаевич,
Дюбанов Владимир Владимирович

РЕШЕНИЕ ЗАДАЧИ АВТОМАТИЧЕСКОГО ПОДБОРА ПАРАМЕТРА r^* АЛГОРИТМА FRIS-CLUSTER

Работа посвящена алгоритму кластеризации FRIS-Cluster, слабым местом которого является параметр r^* . Данный параметр необходимо задавать вручную, при этом судить о правильности выбора можно только по окончании работы алгоритма. В статье исследована зависимость качества решения задачи кластеризации от значения r^* . Сформулирована эвристика для определения заведомо плохого качества кластеризации до окончания работы алгоритма. На базе полученной эвристики предложен алгоритм для автоматического подбора параметра r^* .

Адрес статьи: www.gramota.net/materials/1/2014/8/14.html

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

Альманах современной науки и образования

Тамбов: Грамота, 2014. № 8 (86). С. 74-79. ISSN 1993-5552.

Адрес журнала: www.gramota.net/editions/1.html

Содержание данного номера журнала: www.gramota.net/materials/1/2014/8/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: almanac@gramota.net

Таким образом, развитие английского ландшафтного дизайна в период с XIII по XIX века прошло несколько этапов, основными из которых являются регулярный сад и пейзажный парк, отражающие принципиально важные вехи в жизни и культуре английского социума. Пейзажный сад стал отражением национально-культурного восприятия англичанами природного мира и политических концепций, заложенных в основу ландшафтного парка как средства сопротивления французскому абсолютизму.

Список литературы

1. Иконников А. В. Лондон. Л.: Стройиздат, 1972. 86 с.
2. Пейзажный парк, английский парк, натуральный сад [Электронный ресурс]. URL: <http://creatioart.ru/dictionary/index.php?let=17&id=575> (дата обращения: 25.05.2014).
3. Швидковский Д. О. Происхождение английского пейзажного парка // Пространство и время. 2012. Вып. 1.
4. Fox K. Watching the English. The Hidden Rules of English Behavior. Nicholas Brealey Publishing, 2008.
5. <http://jane Austens world.wordpress.com/2009/03/01/18th-amp-19th-century-pleasure-and-tea-gardens-in-london> (дата обращения: 31.05.2014).
6. <http://tourism-london.ru/spavochnik-po-velikobritanii/istoriya-i-arhitektura-anglii/2181-zamki-plantagenetov-pervye-usadby.html> (дата обращения: 04.06.2014).
7. <http://www.britannica.com> (дата обращения: 04.06.2014).
8. <http://www.gardenmuseum.org.uk> (дата обращения: 31.05.2014).
9. Pevsner N. Studies in Art, Architecture and Design. New York, 1968. Vol. 1.

THE ENGLISH GARDEN ART FORMATION IN THE XIII-XIX CENTURIES

Dzhumagalieva Aliya Satvaldovna

Volgograd State University

asdzhumagalieva@mail.ru

The article is devoted to the formation of garden art in England in the XIII-XIX centuries. The dynamics of the historical development of this phenomenon, which is an essential part of the culture and everyday life of the English society, is shown. The author considers two crucial for studying the history of the English landscape design styles: regular and landscape ones, the main stages of the development of each of the style directions are singled out.

Key words and phrases: nodal garden; garden of medicinal plants; tea garden; pleasure garden; regular garden; landscape park.

УДК 519.688

Физико-математические науки

Работа посвящена алгоритму кластеризации FRiS-Cluster, слабым местом которого является параметр r^ . Данный параметр необходимо задавать вручную, при этом судить о правильности выбора можно только по окончании работы алгоритма. В статье исследована зависимость качества решения задачи кластеризации от значения r^* . Сформулирована эвристика для определения заведомо плохого качества кластеризации до окончания работы алгоритма. На базе полученной эвристики предложен алгоритм для автоматического подбора параметра r^* .*

Ключевые слова и фразы: кластеризация; кластерный анализ; таксономия; FRiS-методология; столп; когнитивный анализ данных; виртуальный противник; расстояние до виртуального противника.

Зырянов Александр Олегович

Павловский Евгений Николаевич, к. ф.-м. н.

Новосибирский государственный университет

alexander.zyryanov44@gmail.com; euxsun@gmail.com

Дюбанов Владимир Владимирович

г. Новосибирск

vladimir@expasoft.ru

РЕШЕНИЕ ЗАДАЧИ АВТОМАТИЧЕСКОГО ПОДБОРА
ПАРАМЕТРА R^* АЛГОРИТМА FRIS-CLUSTER[©]

1. Введение

В Лаборатории анализа данных Института математики имени С. Л. Соболева активно ведется разработка новой методологии анализа данных FRiS. Методология основана на применении функции конкурентного

сходства, так называемой *FRiS*-функции 1. В рамках данной методологии был разработан алгоритм кластеризации *FRiS-Tax* 4.

Предметом нашего исследования стал алгоритм *FRiS-Cluster*, который является первым шагом алгоритма *FRiS-Tax* и позволяет находить отдельные кластеры. Каждый кластер описывается столпом – наиболее представительным элементом данного кластера. Основная идея алгоритма заключается в оптимальном выборе столпов для максимизации *FRiS*-компактности. **Ошибка! Источник ссылки не найден..**

Алгоритм таксономии *FRiS-Tax*, первым шагом которого выполняется алгоритм *FRiS-Cluster*, показывает результаты лучше, чем его конкуренты. **Ошибка! Источник ссылки не найден..** В условиях отсутствия классов объектов вводится класс виртуальных противников, относительно которого вычисляется значение *FRiS*-функции для любой пары объектов из исходной выборки. Недостатком данного алгоритма является ручная перебор параметра r^* – расстояния до ближайшего виртуального противника.

Исследователь может судить о правильности выбора параметра r^* лишь после окончания работы алгоритма, выступив в роли эксперта и оценив качество кластеризации. Более того, кластеризация большого числа элементов может занять существенное время: алгоритм имеет сложность порядка $O(n^2)$, и неправильный выбор параметра может привести к необходимости повторно затратить время.

Цель работы – разработка подхода к автоматизации подбора параметра r^* . Задачами работы являются: изучение влияния параметра r^* на качество конечной кластеризации; разработка эвристики, позволяющей определять плохое решение задачи кластеризации до завершения работы алгоритма; предложить алгоритм, определяющий оптимальный параметр r^* автоматически на основе заданной выборки объектов.

2. Влияние параметра r^* на работу алгоритма и качество конечной кластеризации

Для установления степени влияния параметра на качество конечной кластеризации были проведены эксперименты, напрямую измеряющие зависимость качества кластеризации от r^* . За качество конечной кластеризации принималось качество кластеризации, которую алгоритм считал наилучшей для данной выборки объектов и данного набора параметров (максимальное количество кластеров и параметр r^*). Для оценки качества кластеризации использовалась оценка *FRiS*-компактности.

Для различных выборок объектов запускается алгоритм *FRiS-Cluster* с параметром r^* , принимающим значение от 0 до много превосходящего среднее расстояние между объектами в выборке. В соответствие каждому значению параметра ставится значение качества кластеризации. Для демонстрации были выбраны тестовые наборы данных с названиями *orange1* и *landscape*.

Наряду с численной оценкой качества использовалась экспертная оценка для установления факта, что неправильный выбор параметра r^* действительно ухудшает кластеризацию.

Чтобы понять природу влияния параметра r^* на работу алгоритма, были построены тепловые карты для случаев сильно заниженного, правильного и сильно завышенного значения параметра r^* . Для построения тепловой карты каждому объекту в соответствие ставилось число – оценка, насколько хорошо данный объект подходит на роль нового столпа. Затем проводилась интерполяция для удобства визуализации.

2.1. Результаты

Результаты экспериментов с экспертной оценкой качества кластеризации, приведенные на Рисунках 1а и 1б, указывают на то, что правильный выбор параметра важен для построения качественной кластеризации. На рисунках видно, что в первом случае отдельные кластеры разделены, тогда как во втором случае отдельные кластеры объединены в более крупные, что ухудшает как качественную оценку по *FRiS*-компактности, так и экспертную оценку.

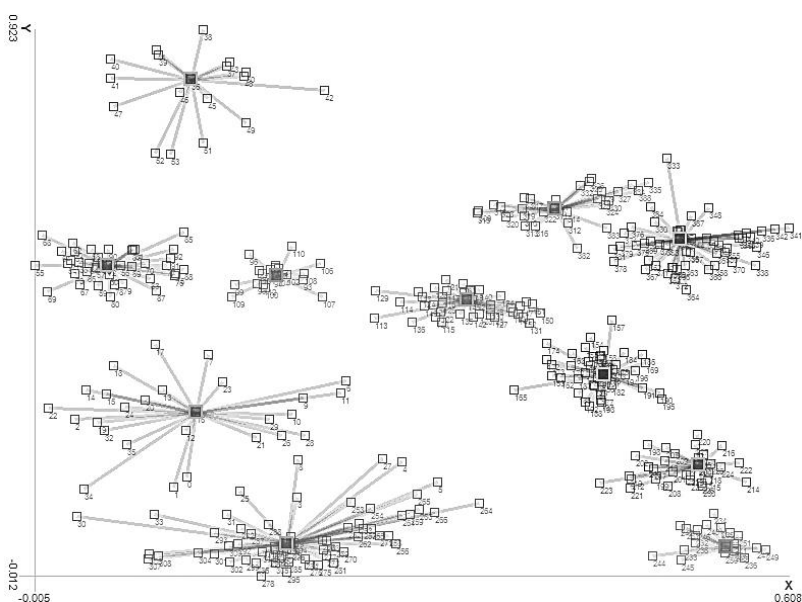


Рис. 1а. Качественная кластеризация тестового набора данных *orange1*, по мнению эксперта

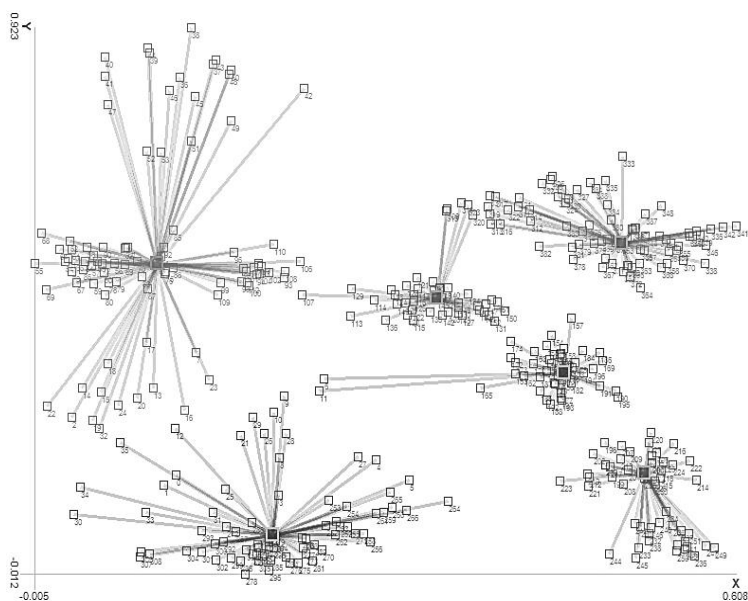


Рис. 16. Некачественная кластеризация тестового набора данных *orange1*, по мнению эксперта

Наиболее характерные графики зависимости качества кластеризации от параметра r^* представлены на Рисунках 2а и 2б. Из графиков видно, что среднее расстояние между объектами (отмечено вертикальной линией) ограничивает область удовлетворительных решений.

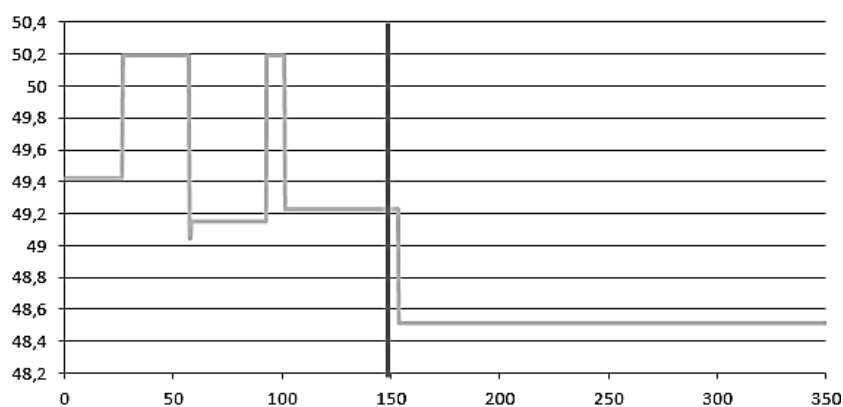


Рис. 2а. Зависимость качества кластеризации от параметра r^* для тестовой выборки *landscape*

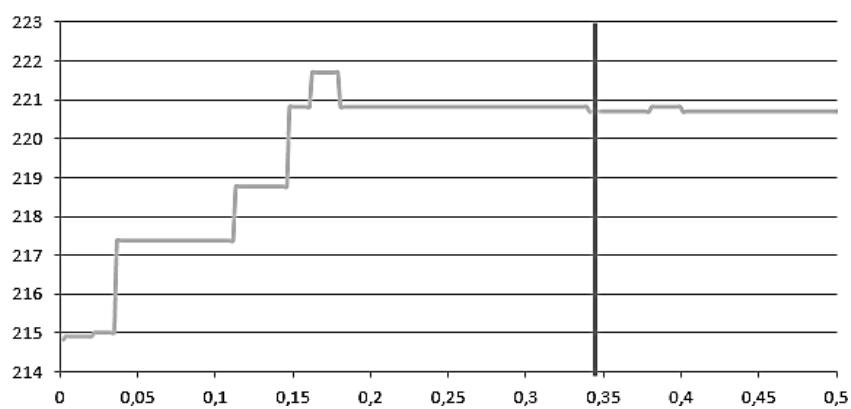


Рис. 2б. Зависимость качества кластеризации от параметра r^* для тестовой выборки *orange1*

На Рисунках 3а, 3б, 3в приведены наиболее характерные тепловые карты. Чем светлее точка, тем лучше объект подходит на роль нового столпа. Видно, что при неправильном выборе параметра r^* и плохом решении задачи кластеризации число светлых точек значительно больше, чем при правильном выборе.

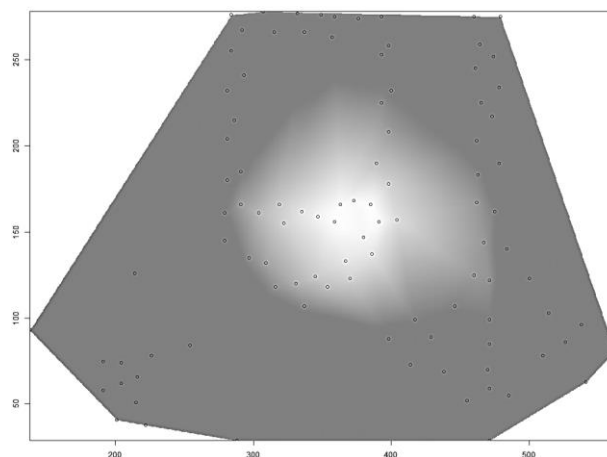


Рис. 3а. Тепловая карта при правильном выборе параметра r^*

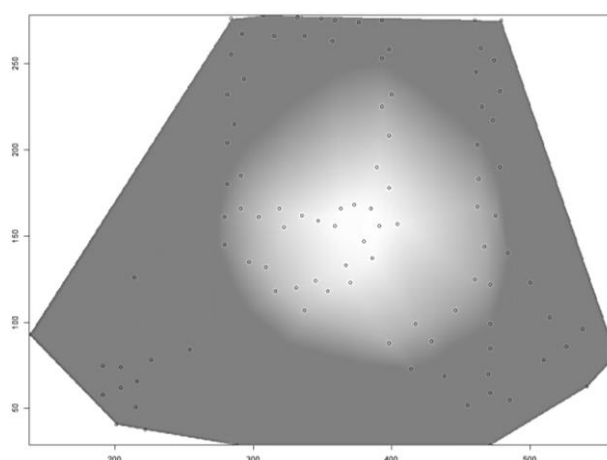


Рис. 3б. Тепловая карта при завышенном значении параметра r^*

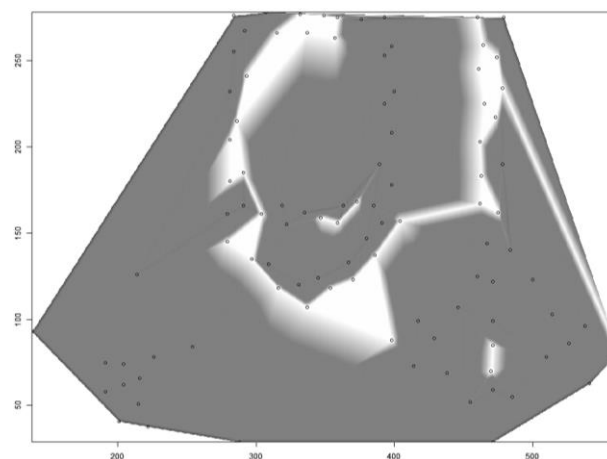


Рис. 3в. Тепловая карта при заниженном значении параметра r^*

Было высказано предположение, что количество светлых точек — точек с высокой температурой — является критерием устойчивости решения задачи кластеризации. Это означает, что даже если мы получили качественное решение при заданном значении параметра r^* , то при большом количестве претендентов на роль нового столпа даже малое изменение параметра не гарантирует, что на роль столпов будут выбраны те же самые объекты, и, следовательно, решение может сильно измениться.

Для проверки гипотезы были проведены эксперименты, выявляющие зависимость качества конечной кластеризации от числа горячих точек. Горячей точка считается, если ее нормированная температура превышает значение 0,9. На данный момент это значение подбирается интуитивно. На Рисунках 4а и 4б хорошо видно, что минимум графика количества горячих точек совпадает с максимумом графика качества кластеризации. Этот факт указывает на возможность применения нашей гипотезы на практике.

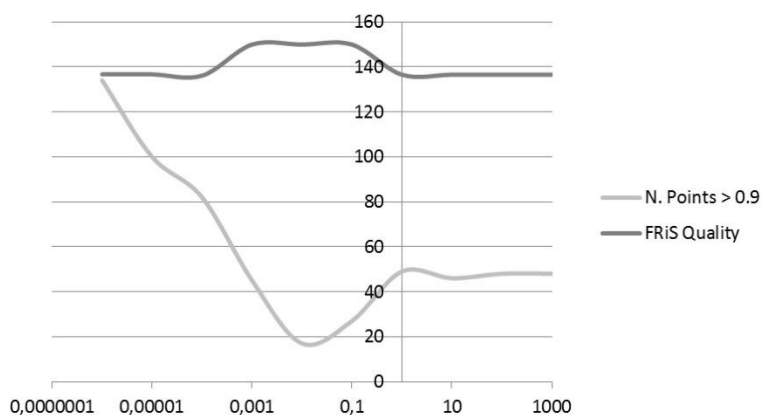


Рис. 4а. Совместные графики числа горячих точек и качества кластеризации тестовой выборки *orange1*: *N.Points>0,9* – количество точек с высокой температурой; *FRiS Quality* – оценка *FRiS*-компактности

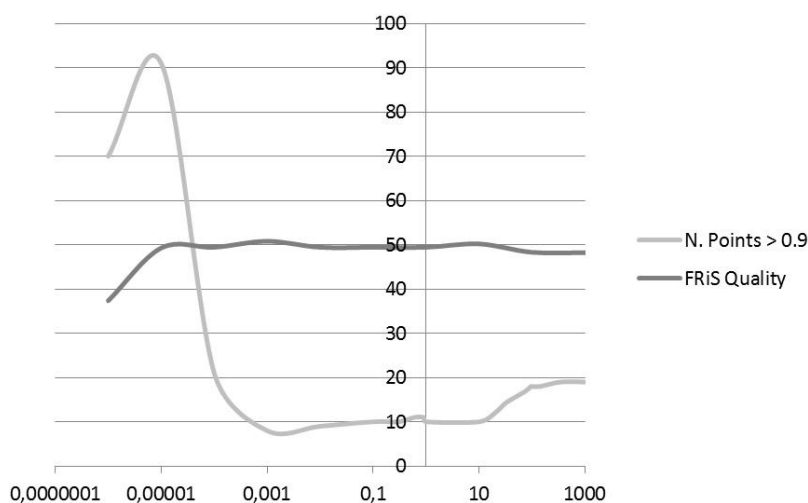


Рис. 4б. Совместные графики числа горячих точек и качества кластеризации тестовой выборки *landscape*: *N.Points>0,9* – количество точек с высокой температурой; *FRiS Quality* – оценка *FRiS*-компактности

2.2. Предсказание качества кластеризации

Для предсказания качества конечной кластеризации была сформулирована эвристика: если процентное содержание горячих точек начинает превосходить определенный порог, который задается вручную, то вероятность получить низкое качество кластеризации высока. Стоит заметить, что речь не идет о численных оценках качества. Эвристика отвечает на вопрос, плохая или хорошая с точки зрения эксперта получится кластеризация.

3. Автоматизация подбора параметра r^*

Для автоматизации подбора параметра r^* предлагается перебирать значения параметра в диапазоне от 0 до среднего расстояния между объектами. Для экономии времени, требуемого на многократные запуски алгоритма, работу алгоритма следует прерывать, если в ходе решения появилось большое количество точек с высокой температурой. Если при определенном значении параметра r^* не возникало большого числа точек с высокой температурой, то перебор значений параметра следует остановить. Таким образом, будет найдено значение параметра r^* , при котором получается удовлетворительное качество кластеризации.

Для экономии времени можно модифицировать алгоритм. Не следует прерывать работу алгоритма, если при постановке нового столпа возникло большое количество точек с высокой температурой, нужно произвести перерасчет текущего шага с использованием нового значения r^* .

4. Заключение

В ходе работы была установлена степень влияния параметра r^* на качество конечной кластеризации, выявлены причины зависимости качества кластеризации от параметра r^* , сформулирована эвристика, с помощью которой можно определить заведомо плохое решение до завершения работы алгоритма, на базе данной эвристики предложен алгоритм, позволяющий автоматически подобрать значение r^* .

Фактически один параметр алгоритма был заменен двумя другими: порог температуры для определения горячей точки, порог количества горячих точек. Однако эти два параметра могут быть установлены из общих соображений без учета характерных особенностей конкретной выборки объектов и выражаются нормированными величинами. Тогда как параметр расстояния до ближайшего виртуального противника r^* не может быть выставлен интуитивно, хотя он и несет в себе смысл масштаба задачи, он, как показали

эксперименты, значительно отличается от среднего расстояния между объектами. Более того, новые параметры могут применяться для тонкой настройки работы алгоритма и в большинстве случаев не требуют изменения своих значений.

Список литературы

1. **Загоруйко Н. Г.** Когнитивный анализ данных. Новосибирск: Гео, 2013.
2. **Borisova I. A., Dyubanov V. V., Kutnenko O. A., Zagoruiko N. G.** Use of the FRiS-Function for Taxonomy, Attribute Selection and Decision Rule Construction // Lecture Notes in Artificial Intelligence. 2011. V. 6581. P. 256-270.
3. **Zagoruiko N. G., Borisova I. A., Dyubanov V. V., Kutnenko O. A.** A Quantitative Measure of Compactness and Similarity in a Competitive Space // Journal of Applied and Industrial Mathematics. 2011. Vol. 5. № 1. P. 144-154. DOI: 10.1134/S1990478911010157.
4. **Zagoruiko N. G., Borisova I. A., Dyubanov V. V., Kutnenko O. A.** Methods of Recognition Based on the Function of Rival Similarity // Pattern Recognition and Image Analysis. 2008. Vol. 18. № 1. P. 1-6. DOI: 10.1007/s11493-008-1001-8.

SOLUTION OF AUTOMATIC SELECTION OF ALGORITHM FRIS-CLUSTER PARAMETER R^*

Zyryanov Aleksandr Olegovich

Pavlovskii Evgenii Nikolaevich, Ph. D. in Physical-Mathematical Sciences

Novosibirsk State University

alexander.zyryanov44@gmail.com; euxsun@gmail.com

Dyubanov Vladimir Vladimirovich

Novosibirsk

vladimir@expasoft.ru

The work is devoted to the clustering algorithm FRiS-Cluster, the weak point of which is the parameter r^* . This parameter must be specified manually, thus the correctness of choice can be judged only at the end of the algorithm. The article explores the dependence of the quality of clustering task solution on the value of r^* . The heuristics for determining the obviously poor quality of clustering before the end of the algorithm working is formulated. On the basis of the derived heuristics an algorithm for automatic selecting the parameter r^* is proposed.

Key words and phrases: clustering; cluster analysis; taxonomy; FRiS-methodology; pillar; cognitive analysis of data; virtual opponent; distance to virtual opponent.

УДК 378.14

Педагогические науки

В статье рассматриваются вопросы индивидуализации обучения математике курсантов военного института. Низкий уровень математических знаний выпускников школы, а также отсутствие требований к математической подготовке абитуриентов осложняют преподавание дисциплины в военном институте ВВ МВД. Обоснована необходимость индивидуального подхода с целью повышения качества обучения курсантов, описаны методические приемы, реализуемые на занятиях по математике в военном институте.

Ключевые слова и фразы: индивидуальный подход; индивидуализация обучения; математическое образование курсантов; самостоятельная работа; алгоритмизация; учебная мотивация.

Игошина Надежда Михайловна, к. пед. н., доцент

Новосибирский военный институт внутренних войск имени генерала армии И. К. Яковлева МВД России
igoshina_mail@mail.ru

ИНДИВИДУАЛИЗАЦИЯ ОБУЧЕНИЯ МАТЕМАТИКЕ КУРСАНТОВ ВОЕННОГО ИНСТИТУТА[©]

Система военного образования переживает сложный и ответственный период реформирования. Образовательные реформы высшей военной школы актуализировали проблему самообразования и самостоятельной работы в обучении, наглядно показавшую значение индивидуальных особенностей курсантов и необходимость их учета. Особая система взаимоотношений участников образовательного процесса, построенная на равноправном сотрудничестве, самостоятельности, активности, индивидуальном подходе, должна помочь в формировании личности офицера, обладающего не только определенным объемом знаний, но и способностью принимать нестандартные решения, самостоятельно мыслить и критически оценивать ситуацию, постоянно пополнять свои знания в непрерывном процессе самообразования.