

Секция «Вычислительная математика и кибернетика»

Декомпозиционный алгоритм кластеризации нечетких реляционных данных

Орлов Алексей Андреевич

Студент

Национальный технический университет Украины КПИ, Институт прикладного системного анализа, Киев, Украина

E-mail: orlov.alexey@gmail.com

В данной работе рассматривается задача кластеризации нечетких реляционных данных, представленных нечеткими матрицами различий. Оперирование такими данными является вполне обоснованным, среди прочего, в задачах поддержки принятия решений: нечеткие попарные различия соответствуют агрегированным экспертным оценкам попарных различий между объектами.

В литературе существует небольшое количество работ, посвященных проблеме кластеризации нечетких объективных данных. Однако рассматриваемая задача является принципиально новой задачей кластеризации, так как заключается в кластеризации именно реляционных данных, которые являются более релевантными в случае их экспертной природы.

В рамках данной работы рассматривался вариативный подход к кластеризации. Для заданного множества альтернатив, совокупности критериев, при заданном количестве кластеров были предложены процедуры получения непустых четких кластеров и нечетких кластеров.

Предложенный декомпозиционный алгоритм кластеризации нечетких реляционных данных (FRDD – fuzzy relational data decomposition) основывается на переходе от задачи кластеризации нечетких реляционных данных, которая не может быть решена непосредственно, к последовательности задач четкой кластеризации реляционных данных за счет декомпозиционного представления нечетких чисел, которое дает возможность перейти к работе с последовательностью интервалов, или пар чисел – границ интервалов.

В работе предложено оригинальное определение функции доверия к кластеризации (характеризует полноту учета входящей нечеткой информации), а также меру неоднозначности кластеризации (характеризует противоречивость кластеризаций при учете различных точек выбранных интервалов в качестве входящих данных). Кроме того, в работе вводится понятие неустойчивых относительно кластеризации элементов, которое позволяет выделить множество объектов, которые занимают относительно однозначное положение в том или ином кластере, и таких, относительно принадлежности которых имеет место неопределенность.

Результаты решения модельных задач продемонстрировали эффективность предложенного алгоритма и адекватность введенной системы показателей оценки кластеризации. Алгоритм кластеризации может быть использован в широком спектре прикладных задач, которые связаны с обработкой нечетких реляционных данных; в частности, представляется целесообразным его применение в качестве информационного инструмента принятия решений в широком классе системных задач, которые используют групповые экспертные оценки в качестве первичной информации.

Литература

1. Krishnapuram R., Joshi A., Nasraoui O., Yi L. Low-complexity fuzzy relational clustering algorithms for web mining. // IEEE Transactions on Fuzzy Systems, 2001, 1(2), PP. 98-110.
2. Oliveira J. V., Pedrycz W. Advances in Fuzzy Clustering and Its Applications. John Willey & Sons Ltd, 2007.
3. Zarandi M.H.F., Razaee Z.S. A Fuzzy Clustering Model for Fuzzy Data with Outliers: <http://arxiv.org/abs/1011.4321v1>.
4. Вятченин Д.А. Нечеткие методы автоматической кластеризации. Минск, 2004.