УДК 519.5; 511.6+517.1

[1]Минаев Ю.Н., д.т.н.,
[2]Филимонова О.Ю., к.т.н.,
[2]Минаева Ю.И., к.т.н.

ИНДЕКС НЕЧЕТКОСТИ НЕЧЕТКИХ МНОЖЕСТВ В КОНТЕКСТЕ КОНЦЕПЦИИ «*DATA MINING*»

[1]Национальный авиационный университет,

[2] Киевский национальный университет строительства и архитектуры

Рассмотрены вопросы определения индекса нечеткости НМ в контексте "извлечения знаний из нечеткости", в частности, на основании процедур иерархической кластеризации. Показано, что анализ объекта на уровне его структурной модели позволяет получить новые знания об объекте, при этом существует возможность создания автоматизируемых процедур, предназначенных для описания структурных характеристики НМ по описаниям его компонентов. На примере определения индекса нечеткости НМ при помощи методов ИАД показана возможность этого подхода к извлечению новых знаний из нечеткости. Рассматриваемая задача является частной, однако позволяет увидеть новые возможности и перспективность предложенного метода.

Введение

Теория нечетких множеств (ТНМ) в настоящее время подвержена переосмыслению и анализа полученных результатов. Это связанно с внедрением практически во все области знаний активно развивающегося направления исследований, получившего в англоязычной литературе название "*data mining*" и "*knowledge discovery*". В частности, наиболее перспективным в этом контексте является «извлечение знаний из нечеткости», т.е. нечетких данных. Показано [1], что под "*knowledge discovery in databases*" (обнаружение знаний в базах данных) (КДД) понимают нетривиальный процесс идентификации достоверных, новых, потенциально полезных и хорошо понимаемых образцов (структур, *patterns*) в данных.

В общем случае под данными понимают множество фактов, представленных в соответствии с формулируемыми целями их использования, под образцами (*patterns*) – некоторые выражения E (формулы языка L), определенным образом характеризующие подмножество фактов (при этом утверждается, что E не сводится к перечислению подмножества фактов из БД). Предполагается, конечно, что L есть некоторый формальный язык представления знаний.

Особую роль при этом занимают процедуры иерархической кластеризации, т.к. анализ объекта на уровне его структурной модели позволяет получить новые знания об объекте, при этом важное значение имеет возможность создания автоматизируемых процедур, предназначенных для описания структурных характеристики нечеткого множества (НМ) по описаниям его компонентов. Открытие структур, которые невозможно определить с помощью стандартных процедур анализа нечетких данных, играет главную роль в понимании того, какие отношения регулируют поведение реальных систем, в частности, решении задач управления в условиях неопределенности на уровне ТНМ.

Сделана попытка на примере определения индекса нечеткости (ИН) НМ при помощи методов ИАД показать возможности этого подхода к извлечению новых знаний из нечеткости. Рассматриваемая задача является частной, однако позволяет увидеть новые возможности и перспективность предложенного метода.

Современный уровень исследований

Понятие *индекса нечеткости* [2] играет своеобразную роль в ТНМ: с одной стороны, ИН позволяет установить границы приближенных оценок (чем больше ИН, тем менее точной будет оценка), с другой стороны, вопрос сравнительной оценки, например, величины ИН для $\tilde{A} \cap \tilde{B}$ или $\tilde{A} \cup \tilde{B}$, если известны величины ИН для $\tilde{A}$ и/или $\tilde{B}$, не может быть решен в принципе. Формальное определение ИН базируется на использовании понятия относительного расстояния между НМ.

Линейный $\nu(\widetilde{A})$ и квадратичный $\eta(\widetilde{A})$ ИН определяются в виде: $\nu(\widetilde{A})=\frac{2}{n}d(\widetilde{A},A_{crisp}^{nearest})$, $\eta(\widetilde{A})=\frac{2}{n^{1/2}}d(\widetilde{A},A_{crisp}^{nearest})$, где $\widetilde{A}, A_{crisp}^{nearest}$ – исходное НМ и ближайшее к нему ЧМ, $0\le\nu(\widetilde{A})\le1$ и $0\le\eta(\widetilde{A})\le1$,($\widetilde{A}$, $A_{crisp}^{nearest}\in X$). При этом $0\le\delta(\widetilde{A},A_{crisp}^{nearest})\le½$ и $0\le\varepsilon(\widetilde{A},A_{crisp}^{nearest})\le½$, где $\delta(\cdot)$ и $\varepsilon(\cdot)$ – обобщенное относительное расстояние Хэмминга между 2-мя НЧ и относительное Евклидово расстояние между НМ соответственно. В [1] приведен пример, где для НМ $\widetilde{A}=\{x_1/0.2,\ x_2/0.6,\ x_3/0.1\}$ вычислен линейный ИН, $\nu(\widetilde{A})=0.46$.

$$\widetilde{A}=\begin{array}{ccc} x_1 & x_2 & x_3 \\ \hline \multicolumn{1}{|c|}{0.2} & \multicolumn{1}{c|}{0.6} & \multicolumn{1}{c|}{0.1} \\ \hline \end{array}$$

$$\nu(\widetilde{A})=0.46$$

Этот пример будет использоваться в дальнейшем для сравнительной оценки нового метода вычисления ИН, который базируется на структурных свойств множеств.

В работе [3] показано, что индексы сравнения обычно имеют теоретико-множественную интерпретацию. В случае многозначных свойств объекты могут быть представлены как НМ. Существуют два основных подхода к проблеме сравнения нечетких множеств. При первом из них используются скалярные индексы сравнения, тогда как при втором подходе нечетко-значные индексы.

Чтобы сравнить два нечетких мно-жества $\widetilde{A}$ и $\widetilde{B}$ с помощью скалярного индекса, следует оценить подходящие комбинации этих нечетких множеств. Предполагается, что такие комбинации должны определяться с помощью введенных ранее операторов. В результате нормировки область изменения этих оценок будет составлять интервал [0, 1]. Индекс сравнения будет иметь вид: $I(\widetilde{A},\widetilde{B})=f(g_1(\widetilde{A}*\widetilde{B}),\ \ldots,\ g_n(\widetilde{A}*\widetilde{B}))$, где $*$ – теоретико-множественный оператор, g_i скалярная оценочная функция НМ, оператор f обеспечивает нормализацию. Скалярная оценочная функция нечеткого множества есть отображение из 2^X – множества всех Нп/М множества X в [0, 1] такое, что: а) $g(\varnothing)=0$, б) $g(X)=1$, в) если $A\subseteq B$(т.е. $\mu_A(x)\le\mu_B(x)$), то $g(A)\le g(B)$.

В [4] показано, что для НМ $\widetilde{A}$, определенного на элементах некото-рого пространства X, степень нечеткости множества $\widetilde{A}$ вычисляется в виде:

$$\rho\,(A)=\max_{x\in X}\{\mu_A(x)\wedge[1-\mu_A(x)]\},$$

$$0\le\rho(A)\le0.5.$$

В [5] рассматривается аксиоматическая структура для измерения функции принадлежности (ФП) НМ (т.е. теоремы представления и единственности), а также предлагается возможный вариант теста на единственность в свете концепции фундаментальноного измерения нечеткости.

Постановка задачи

В качестве расчетной модели для оценки нечеткости примем параметр, физически аналогичный линейному $\nu(\widetilde{A})$ или квадратичному $\eta(\widetilde{A})$ ИН, имеющим смысл расстояний между исходным НМ и ближайшим к нему четким множеством (ЧМ) (или дополнением НМ) в случае векторного ИН, рассмотрев множества как объекты, наделенные определенной структурой. Пусть исходное НМ $\widetilde{A}=\{a_i/\mu_i^a\}$, ближайшее ЧМ $-A_{crisp}^{nearest}$. Для множеств $\widetilde{A}$ и $A_{crisp}^{nearest}$ сформируем матрицы близости $\mathbf{D}^{A\_fuz}$ и $\mathbf{D}^{A\_crisp}$, для которых вычислим расстояние между ними и разность соответствующих матричных норм. Аналогично определим векторные нормы векторов $\widetilde{A}$ и $A_{crisp}^{nearest}$.

Матрицы близости на основании кластерного анализа позволяют определить *структурные* характеристики множеств $\widetilde{A}$ и $A_{crisp}^{nearest}$, которые также позволяют, в свою очередь,получить новую дополнительную информацию о рассматриваемой в виде НМ неопределенности. В частности, такой информацией может быть фрактальное число и фрактальная размерность множества, бинарная матрица и соответственно бинарное дерево и др. Отметим, что в общем случае матрица расстояний может быть как метрической, так и ультраметрической. Учитывая многообразие способов определения расстояний и формирования кластеров, таких структур может быть несколько, что позволяет существенно расширить объем знаний, получаемых из неопределенности. ЧМ – ($X=\{x_i\}, i=1,n$; $\mathbf{D}_x$), где x_i – элементы множества, $\mathbf{D}_x$ – структура множества, характеризуемая матрицами расстояний в исходной метрике и ультраметрике, $\mathbf{D}_x(d, d_u)$, d_u – ультраметрика, определяемая по вновь созданной структуре или на основании *p*-адических значений параметров модели; НМ $\tilde{X}=\{x/\mu^X\} \Rightarrow (X=\{x_i\}, i=1, n; {}^f\mathbf{D}_x)$, где ${}^f\mathbf{D}_x$ – структура множества, определенная с учетом заданных ФП, наложение ФП меняет внутреннюю структуру множества.

Методы решения поставленной задачи

Метрические свойства множес-тв $\tilde{A}$, $A^{nearest}_{crisp}$ рассмотрим, приняв, что $\tilde{A}$ представляет собой НМ, $\tilde{A}=\{x^a_1 = x_1/\mu^X_1, x^a_2 = x_2/\mu^X_2, \ldots, x^a_n = x_n/\mu^X_n\}$, ближайшее ЧМ будем принимать в виде: $A^{nearest}_{crisp}= \{x_k/0, x_j/1, x_m/0\}$, $n>k>1$, $k<j<m$, $j<m\le n$. Для множеств $\tilde{A}$ и $A^{nearest}_{crisp}$ сформируем матрицы близости, определив попарно расстояния между элементами указанных множеств, получаем симметричные (с нулевой диагональю) матрицы $\mathbf{D}^{A\_FUZ}$ и $\mathbf{D}^{A\_crisp}$ размерностью $n\times n$ и 3×3. Вычислим следующие величины:

(а) разность норм матриц $\mathbf{D}^{A\_FUZ}$ и $\mathbf{D}^{A\_crisp}$, рассмотрев 3 типа матричных норм: 1-*norm*, *fro-norm* и *inf-norm*;

(б) норма разности матриц $\mathbf{D}^{A\_FUZ}$ и $\mathbf{D}^{A\_crisp}$ с теми же типами матричных норм: 1-*norm*, *fro-norm* и *inf-norm*;

(в) расстояние между матрицами $\mathbf{D}^{A\_FUZ}$ и $\mathbf{D}^{A\_crisp}$ в виде

$$d(\mathbf{D}^{A\_FUZ}, \mathbf{D}^{A\_crisp})=\mathrm{trace}\,((\mathbf{D}^{A\_FUZ} -_{\oplus} \mathbf{D}^{A\_crisp})* (\mathbf{D}^{A\_FUZ} -_{\oplus} \mathbf{D}^{A\_crisp})^T),$$

где символ $-_{\oplus}$ – Кронекерова разность, понимаемая следующим образом: если **A, B** – матрицы размерностями $m\times m$ и $n\times n$, то $\mathbf{A} -_{\oplus} \mathbf{B} = \mathbf{I}_A \otimes \mathbf{A} - \mathbf{B} \otimes \mathbf{I}_B$, где $\mathbf{I}_A, \mathbf{I}_B$ – единичные матрицы размерностью $n\times n$ и $m\times m$ соответственно.

Аналогично определим векторные нормы:

(а) разность норм векторов $\tilde{A}$ и $A^{nearest}_{crisp}$, рассмотрев 3 типа матричных норм: 1-*norm*, *fro-norm* и *inf-norm*;

(б) норма разности векторов $\tilde{A}$ и $A^{nearest}_{crisp}$ с теми же типами матричных норм: 1-*norm*, *fro-norm* и *inf-norm*.

В соответствии с нотацией МатЛаб 1-*norm*, *fro-norm* и *inf-norm* имеют следующий смысл:

1-*norm* – наибольшая столбцовая сумма = max(sum(abs(X)));

fro-norm – Фробениусовская норма = *sqrt*(*sum*(*diag*(X'*X)));

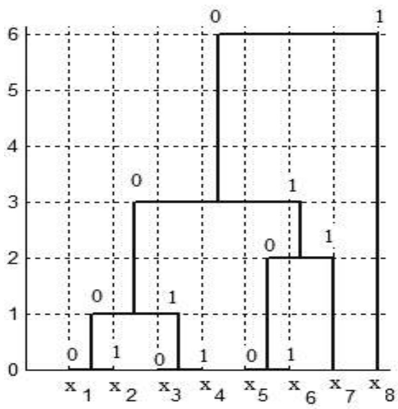
inf-norm – наибольшая строчная сумма = *max*(*sum*(*abs*(X'))).

Предлагаемые параметры физически имеют смысл расстояний и таким образом несут достаточно полную информацию о величине ИН, особенно это относится к расстоянию между матрицами $\mathbf{D}^{A\_FUZ}$ и $\mathbf{D}^{A\_crisp}$. Кроме того, предложенные параметры могут нести *дополнительную* информацию, извлекаемую из нечеткости, и таким образом являются результатом *интеллектуального* анализа. Весьма важным, на наш взгляд, является то, что предлагаемые параметры могут быть использованы для анализа *гранулированных* НМ, если гранулы представляются матрицами.

Следует отметить, что в настоящее время возможности использования структурных свойств НМ практически не исследованы.

Прикладная задача, решаемая в данной работе, сосотоит в оценке эффективности использования НМ (под/множества упорядоченных пар) с точки зрения *раскрытия* неопределенности, в частности, насколько меняется упорядоченность системы (объекта – НМ) при условии, что некоторое множество, являясь п/множеством УМ, наделено *субъек-тивной* ФП. НМ рассматривается наделенным иерархической структурой, определенной с метрикой ближайшего – *min dist* (*x, y*) (или наиболее удаленного – *max dist*(*x, y*)) соседей. Предложенная методология приобретает следующее содержание. В связи с тем, что теóрия категóрий, как раздел математики, изучает свойства отношений между математическими объектами, не зависящие от внутренней структуры объектов, в нашем случае рассматриваются объекты именно с точки зрения внутренней структуры, более конкретно по степени *близости*, определяемой метрическими (ультраметрическими) методами.

Результатом кластеризации мно-жества является иерархическая структура, представленная в виде бинарного дерева –дендрограммы [6]. Кодирование дендрограммы бинарным кодом [0, 1] приведено на рис. 1 и приводит к использованию т.н. *p-адических чисел*, 2-адическое кодирование определено для любого объекта *x*, связанного с терминальным узлом, в виде: $x = \sum_{j=1}^{n-1} c_j p^j,\ c_j \in \{0,1\}, p = 2$. Здесь *j* – уровень или ранг (корень: *n*-1; терминал: 1) и *I* – объектный индекс. Пусть *x* – вектор-столбец $\{x_1, x_2,\ldots,x_n\}^T$, *p* – вектор столбец $\{p^1, p^2, \ldots, p^{n-1}\}^T$, тогда матричная форма *p*-адического кодирования, использованная выше, имеет вид: $x=\mathbf{Cp}$, где x – десятичный код, *C* – матрица дендрограммных ветвей, *p* – вектор степени целого. Показано [6], что существует эквивалентность между р-адическим числом, *p*-адическим расширением и элементом $\mathbf{Z}_p$ (*p*-адические целые).



$x_1 : 0;$
$x_2 : 1\cdot 2^0;$
$x_3 : 1\cdot 2^1;$
$x_4 : 1\cdot 2^0 + 1\cdot 2^1;$
$x_5 : 1\cdot 2^3;$
$x_6 : 1\cdot 2^0 + 1\cdot 2^3;$
$x_7 : 1\cdot 2^2 + 1\cdot 2^3;$
$x_8 : 1\cdot 2^6;$

$$\begin{pmatrix} 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&1&0\\ 0&0&0&0&0&0&1&0\\ 0&0&1&1&0&0&0&0\\ 0&1&0&1&0&0&0&0 \end{pmatrix}$$

а б в

Рис. 1. а) закодированное бинарное дерево; б) 2-адическое кодирование терминальных узлов; в) матричная форма 2-адического кодирования

Элементы p-адического исчисления. [7]. Пусть *p* – будет простым числом в целых **Z**, используя *p* можно определить метрику на рациональных числах $\mathbb{Q}$ через ограничение на **Z** следующим образом. Для каждого $a \in \mathbf{Z}$ определим p-величину $\nu_p(a)$ как наибольшую степень k, для которой p^k делит a (с $\nu_p(0):=\infty$). Такая функция есть расширенной p-ади-ческой оценкой ν_p: $\mathbf{Q} \to \mathbf{Z}$ через $\nu_p(a/b)= \nu_p(a)-\nu_p(b)$ для всех рациональных чисел a/b.

Определим *«расстояние»* между 2-мя рациональными числами *x, y* как $d(x,y)= p^{-\nu_p(x-y)}$. Это определение метрики на $\mathbb{Q}$, в котором близость к нулю эквивалентна высшей делимости на простое *p*. Пополнение рациональных в этой топологии – это поле *p*-адических чисел $\mathbf{Q}_p$ и пополнение **Z** – это подкольцо p-адических целых – $\mathbf{Z}_p$. Любое p-адическое целое имеет уникальное расширение (как сходящийся ряд) в форме $\sum_{n=0}^{\infty} a_p p^n$, где a_p – целые в диапазоне $0 \le a_p < p$. Топология, определенная такой метрикой имеет предпочтительные вычислительно-интуитивные свойства.

Например, для каждого неотрицательного целого *k* множество $p^k\mathbf{Z}_p$ – множество элементов, делимых на p^k – это шары с радиусами p^{-k}, они одновременно открыты и закрыты. $\mathbf{Z}_p$ рассматривается как топологическое пространство, которое одновременно компактно и тотально несвязно. В частности, $\mathbf{Q}_p$ и $\mathbf{Z}_p$ – локальные компактные Абелевы группы для сложения. Можно определить *min*- и *max*-структуры четкого множества ${}^{min}\mathbf{D}_x$, ${}^{max}\mathbf{D}_x$ такие, что в определенном смысле (для некоторых параметров) может быть $r({}^{min}\mathbf{D}_x) \le r({}^{f}\mathbf{D}_x) \le r({}^{max}\mathbf{D}_x)$, в частности, м.б. *r* – ранги. В системе координат, предназначенной для представления структур, в частности, системе бинарных деревьев четкие и нечеткие множества описываются практически единообразно: $X=(\{x_i\}, i=1,n;\ \mathbf{D}_x)$; $\tilde{X}=\{x/\mu^x\} \Rightarrow (\{x_i\}, i=1,n;\ {}^{f}\mathbf{D}_x)$. Отметим, что возможно определение ${}^{f}\mathbf{D}_x$ на основе $\mathbf{D}_x$.

Экспериментальное исследование структурных параметров множеств, аналогичных (адекватных) индексу нечеткости

В работе [1] показано, что Нп/М и его дополнение имеют один и тот же ИН, что приводит к тому, что ни одна из операций $(\bigcap, \bigcup, \bar{\ })$ не дает какого-либо эффекта систематического изменения нечеткости. Это важное условие подтверждено экспериментально на примере НМ $\{0.2;0.6;0.1\}$ и его дополнения $\{0.8;\ 0.4;\ 0.9\}$. На рис. 2 приведены матрицы близости для метрик Хэмминга и Евклида и бинарные деревья, построенные для альтернативных методов ИК. Как видно из рис. 2, матрицы близости и бинарные деревья совпадают.

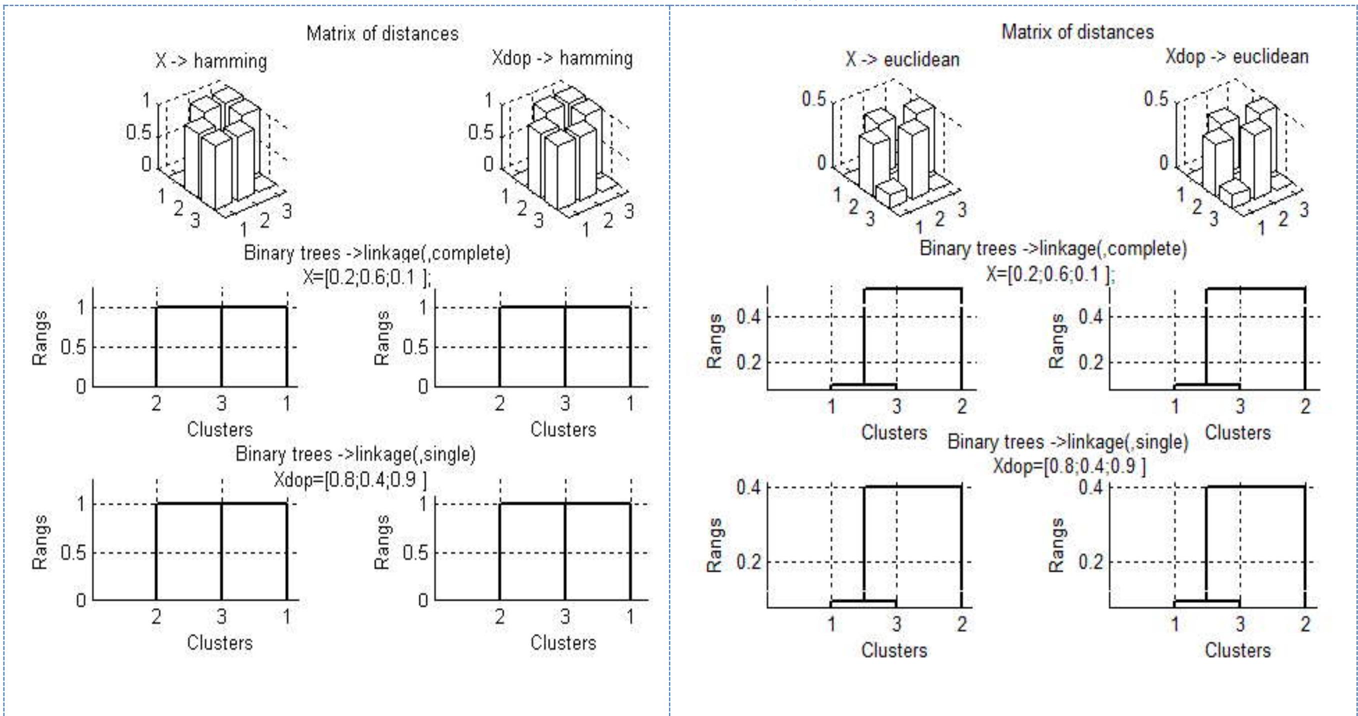


Рис.2. Иерархическая кластерная структура исходного Нп/М и его дополнения

Рассматриваются объекты: НМ $\tilde{A}=\{0.2;\ 0.6;\ 0.1\}$,ближайшее к $\tilde{A}$ ЧМ $A_{crisp}^{nearest}=\{0;1;0\}$. Параметры множеств (исходное НМ и БЧМ) и матрицы близости в метрике Хэмминга представлены на рис. 3.

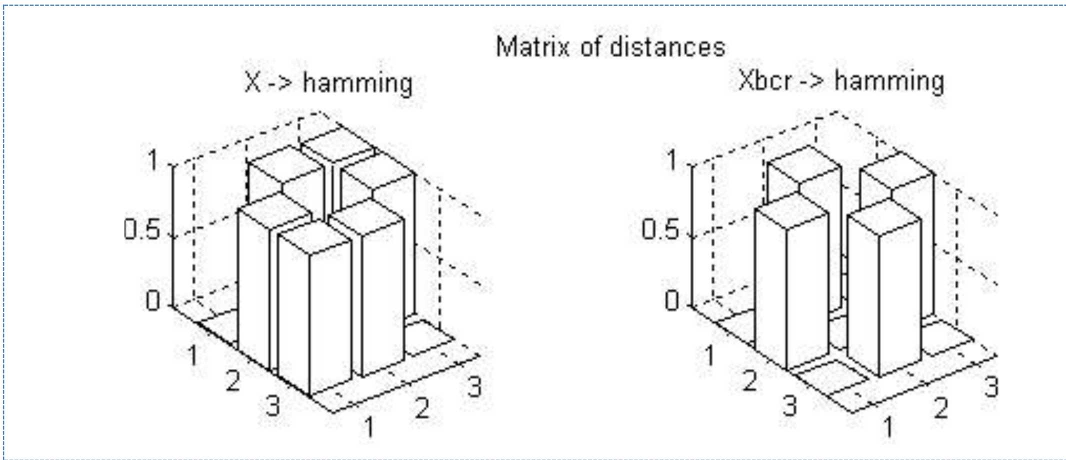


Рис. 3. Матрицы близости для множеств $\tilde{A}=\{0.2;\ 0.6;\ 0.1\}$ и $A_{crisp}^{nearest}=\{0;1;0\}$

В табл. 1 представлены матричные нормы матриц близости. Во всех случаях для норм разности матриц близости имеем достаточно близкое совпадение контрольного параметра (линейный ИН – $\nu(\tilde{A})=0.46$) с расчетными: $0.46 \leftrightarrow [0.5 \div 0.7]$ и $0.46 \leftrightarrow [0.25 \div 0.3]$.

Отметим, что параметр – расстояние между матрицами близости, определенный по выражению:

$$d = \left(\mathrm{trace}((\mathbf{X} - \mathbf{X}_{bcr}) * (\mathbf{X} - \mathbf{X}_{bcr})^{T})/n^{2} \right)^{1/2},$$

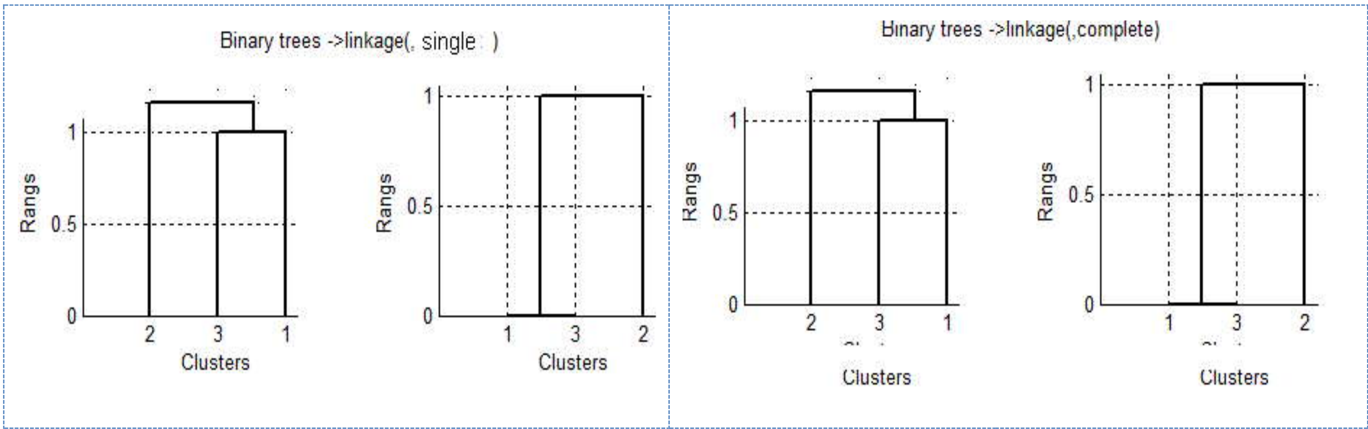
имеет значение 0.46, что полностью совпадает с контрольным параметром.

Построение структур множеств $\tilde{A}$ и $A^{nearest}_{crisp}$ с использованием ИК по методам ближайшего (*min* (d_{ji})) и наиболеее удаленного соседей (*max* (d_{ji})) представлено на рис. 4. Отметим, что в обоих случаях структуры для $\tilde{A}$ и $A^{nearest}_{crisp}$ различны: $S_1=(3\cup 1)_{rang=1}\cup 2$ и $S_2=(3\cup 1)_{rang=0}\cup 2$.

Таблица 1.

Параметры матриц близости

	Linkage(, *single*)		*Linkage*(, *complete*)	
Тип матричной нормы	½ Разности норм матриц	½ Нормы разности матриц	½ Разности норм матриц	½ Нормы разности матриц
Fro norm	0.25	0.7	0.25	0.7
Inf norm	0	0.5	0	0.5
Norm-1	0.3	0.5	0.3	0.5



а) б) в) г)

Рис. 4. Бинарные деревья, построенные на основании матриц близости для $\tilde{A}$ (а, в) и $A^{nearest}_{crisp}$ (б, г) для методов ИК (а, в – МНБС, в, г – МНуС)

Таблица 2.

Параметры бинарных деревьев

Отношение ½ рангов и ½ фрактальных чисел	
Linkage(, *single*)	*Linkage*(, *complete*)
½ $(r^{(1)}/r^{(2)})$=½ ·1/1=0.5; ½ $(f_r^{(1)}/f_r^{(2)})$=½ $(2^1+2^1)/2*(2^0+2^1) \approx 0.33$	½ $(r^{(1)}/r^{(2)})$=½ ·1/1=0.5; ½ $(f_r^{(1)}/f_r^{(2)})$=½ $(2^1+2^1)/2*(2^0+2^1) \approx 0.33$

Анализ табл. 2 показывает, что предложенные параметры – отношение ½ рангов и ½ фрактальных чисел – дают достаточно близкое совпадение с контрольным параметром $0.46\in[0.33\div 0.5]$. Отметим, что с целью экономии места и тривиальностью бинарных матриц кодирование бинарных деревьев в работе не приведено в табл. 3 приведены векторные нормы НМ $\tilde{A}$ и ближайшего четкого $A^{nearest}_{crisp}$.

Таблица 3.

Векторные нормы НМ $\widetilde{A}$ и ближайшего четкого $A^{nearest}_{crisp}$

Тип нормы	Разность норм векторов *norm*($\widetilde{A}$,тип)-*norm*($A^{nearest}_{crisp}$, тип)	Норма разности векторов *norm*($\widetilde{A}$ - $A^{nearest}_{crisp}$,тип)
Fro norm	0.36	0.46
Inf norm	0.40	0.40
norm-1	0.36	0.46

Как видно из табл. 3, в 1-ом случае – разность норм векторов имеем: 0.46 $\leftrightarrow$ [0.36÷0.4], во 2-ом случае – норма разности векторов: 0.46$\in$[0.40÷0.46].Отметим, что разность норм векторов имеет косвенное отношение к ИН, норма разности векторов – прямое.

Проведенные эксперименты с различными типами метрик: Евклидовой, Чебышева, городских кварталов и др. подтвердили практическое совпадение полученных результатов с контрольными.

Выводы

1. Современный уровень исследований в ТНМ является наиболее перспективным в контексте «извлечения знаний из нечеткости», т.е. нечетких данных. Особую роль при этом отводят процедурам иерархической кластеризации, т.к. анализ объекта на уровне его структурной модели позволяет получить новые знания об объекте, при этом существует возможность создания автоматизируемых процедур, предназначенных для описания структурных характеристики НМ по описаниям его компонентов. В работе на примере определения индекса нечеткости НМ при помощи методов ИАД показана возможность этого подхода к извлечению новых знаний из нечеткости. Рассматриваемая задача является частной, однако позволяет увидеть новые возможности и перспективность предложенного метода.

2. В качестве расчетной модели для оценки нечеткости естественно принимать параметр, физически аналогичный линейному $\nu(\widetilde{A})$ или квадратичному $\eta(\widetilde{A})$ ИН, имеющим смысл расстояний между исходным НМ и ближайшим к нему ЧМ (или дополнением НМ в случае векторного ИН), рассмотрев множества как объекты, наделенные определенной структурой.

3. На основании матриц близости путем использования иерархического кластерного анализа определены *структурные* характеристики множеств $\widetilde{A}$ и $A^{nearest}_{crisp}$, которые позволяют, в свою очередь,получить новую дополнительную информацию о рассматриваемой в виде НМ неопределенности. В частности, такой информацией может быть фрактальное число и фрактальная размерность множества, бинарная матрица и соответственно бинарное дерево и др. Отметим, что в общем случае матрица расстояний может быть как метрической, так и ультраметрической.

Список литературы

1. Финн В.К. Об интеллектуальном анализе данных // Новости Искусственного интеллекта. – № 3. – 2004. – С. 12-19.
2. Кофман А. Введение в теорию нечеткиз множеств: Пер. с франц. – М.: Радио и связь, 1982. – 432 с.
3. Дюбуа Д.,Прад А. Общий подход к определению индексов сравнения в теории нечетких множеств. – В кн. Нечеткие множества и теория возможностей, Последние достижения: Пер. с англ. / Под ред. Р.Р. Ягера. – М.: Радио и связь, 1986. – 408 с.
4. Гожальчаны М.Б., Кишка Е.Б., Стахович М.С. Некоторые проблемы изучения адекватности нечетких моделей. – В кн. Нечеткие множества и теория возможностей, Последние достижения: Пер. с англ./Под ред. Р.Р. Ягера. – М.: Радио и связь. – 1986. – 408 с.
5. Норвич А.М., Турксен И.Б. Фундаментальное измерение нечеткости. В кн. Нечеткие множества и теория возможностей, Последние достижения: Пер. с англ. / Под ред. Р.Р. Ягера. – М.: Радио и связь. 1986. – 408 с.
6. Murtagh F. Symmetry in Data Mining and Analysis: A Unifying View based on Hierarchy. – 18 May 2008. – P. 34.
7. Baaket M., Moody Г.V., Schlottmann. Limit-(quasi)-periodic points set as quasicrys-tals with p-adic internal spaces. – 17 Jan 1999. – P. 17.