



О СОГЛАСОВАННОСТИ ОТНОШЕНИЙ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

ЛУГАНСКИЙ А.М., МАШТАЛИР В.П.,
ПУТЯТИНА А.Е., ШЛЯХОВ В.В.

Для моделирования информационных систем формализуются свойства отношений эквивалентности, индуцируемых произвольными n -арными отношениями. Вводятся и изучаются матрицы согласованности отношений. Обсуждаются особенности применения результатов в мультиалгебраических системах.

1. Введение

Отношения эквивалентности являются одним из основных и предпочтительных типов связи между объектами и их множествами в информационных системах интерактивного и коммуникативного сервиса. И концептуально, и с практической точки зрения, в явном виде или опосредованно они представляют собой основу для решения классификационных задач, извлечения *ad hoc* информации и знаний из наборов неструктурированных сведений. В условиях многозначности определенный интерес представляют «естественные» отношения эквивалентности, индуцируемые произвольными отношениями между множествами объектов или семействами гетерогенных признаков. Другими словами, при анализе данных различной природы поиск подобных отношений эквивалентности, раскрывающих внутренние временные, причинно-следственные, морфологические, семантические и иные связи, создает предпосылки для синтеза адекватных математических моделей, а следовательно, и для выявления конструктивных подходов к решению прикладных задач. *Цель работы* – анализ свойств отношений эквивалентности, индуцируемых произвольными n -арными отношениями.

2. Состояние методов моделирования информационных систем на основе отношений и задача исследования

Начиная с пионерской работы [1], моделированию информационных систем на основе «аппроксимирующих множеств» (rough sets, далее по тексту – RS) уделяется весьма пристальное внимание [см., например, 2–10]. В первую очередь, эта популярность объясняется тем, что в отличие от статистических и фаззи-подходов, в которых неопределенность и недостаточность информации в той или иной степени связана с некоторыми привлекаемыми

ми извне вероятностными характеристиками, RS-методы базируются только на учете «внутренних» свойств анализируемых данных [2–5]. В качестве основных свойств используются «неразличимость» данных, определяемая отношениями эквивалентности [6–8], и «сходство», задаваемое толерантностями [9, 10]. К настоящему времени в RS-моделировании достаточно полно исследованы методы использования аппроксимаций сверху и снизу, граничных областей [2, 5, 8], значительные успехи достигнуты при решении автоматического построения минимальных, достаточных и исчерпывающих решающих правил [4, 7]. При моделировании информационных систем обычно предполагается, что отклик системы определяется на декартовой степени множества возможно разнородных атрибутов. Однако гетерогенные исходные данные, с одной стороны, представляя собой более общий случай с позиций теоретических исследований, а с другой – описывая целый ряд важных приложений, практически не изучены.

Решаемая в работе задача – изучение согласованности индуцированных произвольными отношениями эквивалентностей, что обеспечит верификацию позиционирования моделируемых систем и применяемого математического инструментария в условиях различного рода неопределенности.

3. Разбиения, индуцируемые n -арными отношениями

Предположим, что семейства значений входных данных заданы в номинальных, дихотомических, порядковых, интервальных, пропорциональных шкалах множествами произвольной природы A_1, \dots, A_n . Ясно, что любой интерпретант информационной системы с той или иной точностью можно представить n -арным отношением $E(x_1, \dots, x_n)$, определенным на декартовом произведении $A_1 \times A_2 \times \dots \times A_n$. Если на множестве аргументов зафиксировать некоторую позицию k и рассмотреть пару элементов x'_k и x''_k , тогда при

$$E(x_1, \dots, x'_k, \dots, x_n) \equiv E(x_1, \dots, x''_k, \dots, x_n), \quad (1)$$

где под знаком « \equiv » понимается выполнение равенства для произвольных наборов

$$x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n \in A_1 \times \dots \times A_{k-1} \times A_{k+1} \times \dots \times A_n,$$

эти элементы неразличимы с использованием введенного отношения, т.е. они находятся в отношении эквивалентности \mathcal{E}_k^E , задаваемом на декартовом квадрате $A_k \times A_k$ условием

$$\langle x'_k, x''_k \rangle \in \mathcal{E}_k^E \Leftrightarrow E(x_1, \dots, x'_k, \dots, x_n) \equiv E(x_1, \dots, x''_k, \dots, x_n).$$

Выполнение свойств рефлексивности, симметричности и транзитивности \mathcal{E}_k^E доказывается непосредственной проверкой. Итак, произвольное n -арное отношение $E(x_1, \dots, x_n)$ на произвольных наборах носителей $A_1 \times A_2 \times \dots \times A_n$ в общем случае индуцирует n отношений эквивалентности \mathcal{E}_k^E на $A_k \times A_k$, где $k \in \{1, 2, \dots, n\}$.

$$\begin{aligned}
E(x_1, \dots, x'_i, \dots, x_n) &\equiv E(x_1, \dots, x''_i, \dots, x_n) \Leftrightarrow \\
&\Leftrightarrow \tilde{E}(y_1, \dots, y_{j-1}, x'_i, y_{j+1}, \dots, y_m) \equiv \\
&\equiv \tilde{E}(y_1, \dots, y_{j-1}, x''_i, y_{j+1}, \dots, y_m)
\end{aligned}$$

для каких-либо x'_k и x''_k . Тем самым, можно говорить о (i, j) -согласованности отношений E и \tilde{E} , если $A_i = \tilde{A}_j$ и $\mathfrak{E}_i^E = \mathfrak{E}_j^{\tilde{E}}$. Совершенно аналогично предыдущим рассуждениям вводится $(n \times m)$ матрица согласованности $\mathcal{J}(E, \tilde{E})$, в которой на $(i \times j)$ -м стоит единица тогда и только тогда, когда отношения E и \tilde{E} (i, j) -согласованы. Ясно, что, продолжая $m - n$ (или $n - m$) раз область определения отношений пустыми множествами, можно ограничиться анализом лишь квадратных матриц. Более того, если задан набор E_1, \dots, E_q отношений, каждое из которых имеет свою арность, то, выбирая из них максимальное значение N , можно построить $S_q^2(N \times N)$ матриц согласованности $\mathcal{J}(E_\alpha, E_\beta)$ ($\alpha, \beta \in \{1, \dots, q\}$) и q матриц внутренней согласованности $\mathcal{J}(E_\alpha)$. Свойства согласованности семейств отношений окончательно могут быть легко сформулированы в виде свойств полученных матриц.

6. Анализ результатов и перспектив исследований

Рассмотрим сначала ряд примеров согласованных отношений. На множестве натуральных чисел N зафиксируем некоторое число p и введем тернарное отношение $E(n_1, n_2, n_3)$, $n_1, n_2, n_3 \in N$ в виде

$$\langle n_1, n_2, n_3 \rangle \in E \Leftrightarrow \text{mod}_p(n_1 + n_2) = \text{mod}_p(n_3). \quad (3)$$

Нетрудно заметить: если зафиксировать $k = 1$ и рассмотреть эквивалентности

$$\langle n'_1, n''_1 \rangle \in \mathfrak{E}_1^E \Leftrightarrow E(n'_1, n_2, n_3) \equiv E(n''_1, n_2, n_3),$$

то получим разбиение множества натуральных чисел на классы вычетов по модулю p . Если $n'_1 = pq_1 + r_1$ и $n''_1 = pq_1 + r'_1$, то для произвольного натурального числа $n_2 = pq_2 + r_2$ имеем

$$\text{mod}_p(n'_1 + n_2) = \text{mod}_p(r_1 + r_2),$$

$$\text{mod}_p(n''_1 + n_2) = \text{mod}_p(r'_1 + r_2),$$

откуда следует $E(n'_1, n_2, n_3) \equiv E(n''_1, n_2, n_3)$. Справедливо также и обратное утверждение. Действительно, предположим противное: пусть

$$E(n'_1, n_2, n_3) \equiv E(n''_1, n_2, n_3)$$

$$\text{и } n'_1 = pq_1 + r_1, \quad n''_1 = pq_1 + r'_1, \quad r_1 \neq r'_1.$$

Тогда при $n_2 = 0$ и $n_3 = r_1$ получаем $\text{mod}_p(n'_1 + n_2) = \text{mod}_p(n_3)$, т.е. $\langle n'_1, 0, r_1 \rangle \in E$. С другой стороны, $\text{mod}_p(n''_1 + n_2) \neq \text{mod}_p(n_3)$, так как

$r_1 \neq r'_1$ и $r_1, r'_1 \leq p$ или $\langle n''_1, 0, r_1 \rangle \notin E$, что противоречит исходной посылке. Нетрудно заметить, что \mathfrak{E}_2^E и \mathfrak{E}_3^E также продуцируют разбиения множества N , совпадающие с \mathfrak{E}_2^E , т.е. матрица внутренней согласованности имеет вид

$$\mathcal{J}(E) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Отметим, что на множестве $\{0, 1, \dots, p-1\}$ отношение E индуцирует групповую операцию

$$r_1 \oplus r_2 = r_3 \Leftrightarrow \langle n'_1, n_2, n_3 \rangle \in E,$$

где $r_1, r_2, r_3 \in \{0, 1, \dots, p-1\}$, $n_1 = pq_1 + r_1$, $n_2 = pq_2 + r_2$, $n_3 = pq_3 + r_3$. Таким образом, данное отношение индуцирует алгебраическую структуру, представляющую собой циклическую абелеву группу с носителем в виде классов вычетов.

Если модифицировать (3) к виду

$$\langle n_1, n_2, n_3 \rangle \in E_1 \Leftrightarrow \text{mod}_{p_1}(n_1 + n_2) = \text{mod}_{p_2}(n_3),$$

где $p_1 \neq p_2$, то $\mathfrak{E}_1^E = \mathfrak{E}_2^E \neq \mathfrak{E}_3^E$, а матрица внутренней согласованности представляется в виде

$$\mathcal{J}(E_1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

При этом E_1 индуцирует бинарную (но не групповую) операцию из $\{0, 1, \dots, p_1-1\} \times \{0, 1, \dots, p_1-1\}$ в $\{0, 1, \dots, p_2-1\}$.

Наконец, если в (3) продолжить область определения аргументов отношений от N до множества целых чисел Z , то порождаемая структура (циклическая группа) сохраняется, однако классы эквивалентности «укрупняются» за счет отрицательных чисел. Рассмотрение аналитических примеров согласованности отношений мотивируется лишь простотой представления и достаточной наглядностью. На практике конструктивность подхода сохраняется при любом способе задания информационных систем.

В [11] исследовались мультиалгебраические системы, являющиеся некоторым расширением понятия фактор-алгебраических систем. Ключевым моментом при определении мультиалгебраических систем является введение отношения эквивалентности на множестве входных данных. Как представляется авторам, мультиалгебраические системы ни в коей мере не выступают в качестве какой-либо альтернативы RS-моделированию информационных систем, а лишь могут дополнять их средствами управления степенью детализации данных на основе оперирования классами эквивалентности, т.е. грануляцией данных.

Научная новизна полученных результатов состоит в том, что найденные внутренние отношения эквивалентности позволяют учитывать требуемую степень и детализации, и огрубления данных (либо инструментальную точность их регистрации, либо обобщение или конкретизацию их информационных свойств). В итоге это обеспечивает рациональный выбор и конструктивные пути применения математических методов в зависимости от доступных или достаточных для анализа свойств входных данных. В *прикладном* аспекте манипулирование носителями входной информации, в том числе и классами их эквивалентности направлено на рациональный выбор форм их представления с точки зрения поиска разумного компромисса между характеристиками быстродействия, точности, надежности программных средств и их сложности. Таким образом, изучение свойств согласованности произвольных отношений является в некотором смысле обоснованием, точнее говоря, позиционированием использования входной информации при моделировании элементов информационных систем мультиалгебраическими системами, что ранее детально не рассматривалось. В качестве направления развития полученных результатов следует указать целесообразность проведения исследований конкретных математических структур.

Литература: 1. Pawlak Z. Rough sets // International Journal of Information and Computer Science. Vol. 11, No. 5. 1982. P.341-356. 2. Yin X., Zhou Z., Meyer F., Li N., Chen S. An approach for data filtering based on rough set theory // Lecture Notes in Computer Science. Vol. 2118. – Berlin: Springer-Verlag. 2001. P. 367-374. 3. Yao Y.Y., Wang T. On rough relations: an alternative formulation // Lecture Notes on Artificial Intelligence. Vol. 1711. Berlin-Heidelberg: Springer-Verlag. 1999. P.82-90. 4. Yao Y.Y., Wong S.K.M., Lin T.Y. A review of rough set models // Rough Sets and Data Mining: Analysis of Imprecise Data / T.Y.Lin, N. Cercone (eds.). Boston: Kluwer Academic Publisher. 1997. P. 47-75. 5. Gattaneo G. Abstract approximation spaces // Rough Sets in Knowledge Discovery / L. Polkowski, A. Skowron (eds.). Heidelberg:

Physica-Verlag. 1998. P.59-98. 6. Inuiguchi M., Tanino T. On rough sets under generalized equivalence relations // Bulletin of International Rough Set Society. Vol. 5, No. 1-2. 2001. P.167-171. 7. Slowicki R., Vanderpooten D. A generalized definition of rough approximations based on similarity // IEEE Transactions on Data and Knowledge Engineering. Vol. 12, No. 2. 2000. P. 331-336. 8. Konikowska B. A logic for reasoning about similarity // Rough Set Analysis / E. Orłowska (ed.). Heidelberg–New York: Physica-Verlag. 1998. P.462-491. 9. Skowron A., Stepaniuk J. Tolerance approximation spaces // Fundamenta Informaticae. Vol. 27, No. 2-3. 1996. P. 245-253. 10. Polkowski L., Skowron A., Jytkow J. Tolerance based rough sets // Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery / T.Y. Lin, A.M. Wildberger (eds.). San Diego: Simulation Council, Inc. – 1995. P. 55-58. 11. Машталир В.П., Шляхов В.В. Свойства мультиалгебраических систем в задачах компаративного распознавания // Кибернетика и системный анализ. 2003. №6. С. 12-32.

Поступила в редколлегию 20.10.2004

Рецензент: д-р техн. наук, проф. Бодянский Е.В.

Луганский Александр Михайлович, начальник научно-исследовательского информационно-компьютерного центра Национального университета внутренних дел. Научные интересы: разработка информационно-управляющих систем. Адрес: Украина, 61080, Харьков, пр. 50-летия СССР, 27, тел. 503–143, e-mail: alex@adm.univd.kharkov.ua.

Машталир Владимир Петрович, д-р техн. наук, ст. научн. сотр., профессор каф. информатики ХНУРЭ. Научные интересы: распознавание образов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 7021–419, e-mail: mashtalir@kture.kharkov.ua.

Путятин Александра Евгеньевна, студентка ХНУРЭ. Научные интересы: проектирование и анализ информационных систем. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-419.

Шляхов Владислав Викторович, канд. техн. наук, доцент, ведущий научный сотрудник кафедры информатики ХНУРЭ. Научные интересы: алгебраические структуры, искусственный интеллект. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-419.

УДК 004.4:621.311

ИССЛЕДОВАНИЕ ВСТРАИВАНИЯ ЦВЗ В ФОРМАТ JPEG

ДУДАРЬ З.В., ЗБИТНЕВА М.В.

Анализируется робастность и пропускная способность формата JPEG. Рассматриваются алгоритмы встраивания в коэффициенты дискретного косинусоидного преобразования. Предлагается метод встраивания в формат JPEG цифрового водяного знака, ориентированного на идентификацию факта проведения атаки.

Состояние проблемы и анализ алгоритмов сжатия

Одна из актуальных проблем сегодняшнего дня – защита прав интеллектуальной собственности, представленной в цифровом виде [1], которая является одним из аспектов защиты информации в целом. Причина популярности стеганографии – ограниче-

ние на использование криптосредств в ряде стран мира [2].

Проблема защиты интеллектуальной собственности существует уже давно, но применительно к цифровым контейнерам информации она появилась сравнительно недавно [1]. Еще 5 лет назад авторские права на электронную информацию (аудио, видео, статические изображения, электронные книги и т.д.) не соблюдались в связи с тем, что такие файлы изначально создавались как нелегальные копии цифровых аудиодисков или коллекций профессиональных фотографий. Сейчас студиям звукозаписи, фотостудиям и другим организациям, зарабатывающим на продаже такого рода информации, необходима возможность оградить свои произведения от нелегального копирования или незаконного присвоения авторских прав. Для этого используются цифровые водяные знаки (ЦВЗ) – скрытая информация, встраиваемая в контейнер и идентифицирующая автора. На данный момент эта