
ПОСТРОЕНИЕ FUZZY-СИСТЕМ ПРИНЯТИЯ РЕШЕНИЙ

Овакимян Анна Седраковна¹, канд. техн. наук

Саркисян Сирануш Гегамовна², канд. физ.-мат. наук, доц.

Зироян Алвард Альбертовна³, канд. экон. наук, доц.

Мартынова Марина Алексеевна⁴, канд. экон. наук

¹ Российско-Армянский университет, ул. Овсепя Эмина, 123, Ереван, Армения, 0051; e-mail: anna.hovakimyan@rau.am

² Ереванский государственный университет, ул. Алека Манукяна, 1, Ереван, Армения, 0025; e-mail: siranushs@ysu.am

³ Российский государственный социальный университет, ул. Вильгельма Пика, 4, стр. 1, Москва, Россия, 129226; e-mail: ziralla@mail.ru

⁴ Чеченский государственный педагогический университет, пр. Х. Исаева, 62, Грозный, Россия, 364068; e-mail: marina-martynova@list.ru

Цель: исследовать задачу построения систем принятия решений на основе имеющегося большого объема распределенных по узлам вычислительного кластера данных, характеризующих предметную область. *Обсуждение:* построение систем принятия решений, которые, как правило, реализуются в виде экспертных систем, является трудоемкой задачей по двум причинам. Ее решение требует привлечения высококвалифицированных экспертов в данной предметной области для обеспечения системы экспертными знаниями, а эффективная реализация системы логического вывода обеспечит принятие экспертной системой правильного решения. Поэтому актуальной является задача автоматизации процесса генерации экспертных знаний, построения и реализации гибких и интеллектуальных систем принятия решений. *Результаты:* предложен подход к генерации экспертных знаний, основанный на одном из методов интеллектуального анализа данных (Data Mining), а именно: методе ассоциативных правил, позволяющем выявлять скрытые в данных зависимости и закономерности. Данный подход универсален относительно предметной области. Сгенерированная Fuzzy-система может предсказать значения требуемых показателей или помочь принять решение в соответствующей предметной области на основе имеющихся текущих данных.

Ключевые слова: ассоциативные правила, Data Mining, кластерная обработка данных, нечеткие множества, нечеткие ассоциативные правила, нечеткая логика, система принятия решения.

1. Введение

Принятие решений является неотъемлемой частью деятельности человека, часто присутствующее в разных областях исследования таких, как медицина, биология, психология, лингвистика, менеджмент, экономика и т.д. Разработан ряд моделей и инструментов поддержки принятия решений, которые широко используются на практике во многих реальных прикладных контекстах [6]. Обычно для автоматизации решения задач по принятию решений привлекаются экспертные системы, основанные на экспертных знаниях и механизмах логического вывода. Экспертные системы направлены на класс задач в слабоструктурированных областях. Диапазон их применения достаточно широк, но каждая из них предназначена для конкретной области знаний. Попытки расширения предметной области даже в рамках данной области знаний в большинстве случаев не привели к существенному успеху.

Построение систем принятия решений, реализуемых, как правило, в виде экспертных систем, является трудоемкой задачей. Ее решение требует привлечения высококвалифицированных экспертов в данной предметной области для обеспечения системы экспертными знаниями, и, следовательно, все более необходимым и востребованным является наличие опыта и интуиции у экспертов. Эффективная реализация системы логического вывода обеспечит принятие экспертной системой правильного решения, являющегося результатом ее функционирования. По этой причине весьма актуальной является задача автоматизации процессов генерации экспертных знаний, построения и реализации гибких и интеллектуальных систем принятия решений.

В современном информационном обществе хранится огромное количество данных, относящихся к разным предметным областям и размещенных в распределенных базах данных. В этих данных зачастую скрыта важная информация, которая проявляется в виде неочевидных зависимостей и закономерностей и которую можно выявить в результате интеллектуального анализа данных (Data Mining). Эта информация может быть использована в качестве экспертных знаний при разработке систем принятия решений в различных областях человеческой деятельности.

Среди методов Data Mining особое место занимает метод ассоциативных правил. Эти правила отражают закономерности в данных, взаимосвязь отдельных элементов в базе данных и частоту их совместных вхождений в базу данных [3]. Совокупность ассоциативных правил может составить основу для построения систем принятия решений.

Наиболее гибкими нам представляются системы принятия решений, основанные на концепциях нечеткой логики и нечетких множеств [5]. Нечеткие подходы дают возможность выражать понятия окружающего мира не количественными характеристиками, а ближе к процессам человеческого мышления. В этом случае зависимости между понятиями могут быть выра-

жены в виде нечетких ассоциативных правил, а логические умозаключения можно делать по законам нечеткой логики.

В настоящей статье предлагается методика генерации Fuzzy-системы принятия решений, основанная на извлечении нечетких ассоциативных правил из накопленных в результате наблюдений больших объемов данных, с использованием нечетких множеств, нечеткой логики и лингвистической модели принятия решений [7, 8].

Разработанный подход универсален относительно предметной области. Метод апробирован на данных из CDC WONDER – базы данных для эпидемиологических исследований.

2. Ассоциативные правила как инструмент Data Mining

Выявление ассоциативных правил на основе данных из хранилищ представляет собой очень важную задачу в процессе интеллектуального анализа данных. Ассоциативные правила являются важным классом закономерностей в данных. Цель их построения состоит в том, чтобы найти частые сопутствующие вхождения элементов в рамках набора транзакций.

Ассоциативные правила принято разбивать на классические, количественные, нечеткие [2]. Предполагается, что данные хранятся в базах данных, например, реляционных, где каждый набор значений признаков называется транзакцией.

Классические ассоциативные правила характеризуются тем, что для их построения необходимы бинарные данные, показывающие присутствует ли данный элемент в транзакции или нет. Задача вывода классических двоичных ассоциативных правил на основе имеющейся базы данных формулируется следующим образом.

Пусть $I = \{i_1, i_2, \dots, i_m\}$ является множеством признаков, $T = \{t_1, t_2, \dots, t_n\}$ представляет собой множество кортежей (транзакций) из значений признаков из множества I , а значениями признаков i_1, i_2, \dots, i_m в транзакции являются нуль или единица.

Ассоциативным правилом называется выражение вида

$$X \rightarrow Y, \quad (1)$$

где $X \subset I$, $Y \subset I$ и $X \cup Y = \emptyset$.

Говорят, что правило $X \rightarrow Y$ поддерживается кортежем $t \in T$, если каждый раз, когда все значения признаков из множества X равны единице, значения всех признаков из множества Y также равны единице.

В задаче построения набора правил, как и в других задачах машинного обучения, одной из центральных проблем является проблема критерия качества, в соответствии с которым можно было бы выбирать лучшее решение. Ассоциативные правила принято оценивать такими качественными характеристиками, как поддержка и достоверность.

Поддержкой правила $X \rightarrow Y$ является доля кортежей из множества T , которые поддерживают данное правило.

Поддержка правила представляется формулой:

$$\text{supp}(X \rightarrow Y) = \frac{|X \cap Y|}{n},$$

где $|X \cap Y|$ – это количество кортежей, поддерживающих правило, n – общее число кортежей.

Достоверность правила описывает процент кортежей, содержащих единичные значения всех признаков из множества X , которые также содержат единичные значения всех признаков из множества Y .

$$\text{conf}(X \rightarrow Y) = \frac{|X \cap Y|}{|X|},$$

где $|X \cap Y|$ – это количество кортежей, поддерживающих правило, $|X|$ – мощность множества X .

Простое ассоциативное правило, связанное с рыночной корзиной, может быть следующим: Сыр \rightarrow Пиво [поддержка = 0,1, достоверность = 0,8]. Это правило выражает отношение между пивом и сыром. Мера поддержки показывает, что пиво и сыр появлялись вместе в 10% всех зарегистрированных транзакций. Достоверность описывает вероятность того, что если в транзакции присутствует сыр, то в ней присутствует также пиво. Данное правило с достоверностью 0.8 показывает, что 80% всех транзакций, включающих сыр, также включали пиво.

При поиске ассоциативных правил учитываются заданные заранее минимальные значения их поддержки и достоверности. Сам процесс поиска ассоциативных правил состоит из двух частей: определение частых множеств признаков на основе заданной минимальной поддержки и собственно вывод ассоциативных правил с учетом минимальной достоверности [4].

Поскольку на практике часто встречаются количественные данные, для их обработки строятся количественные ассоциативные правила, основанные на бинарных ассоциативных правилах [3]. В этом случае диапазон допустимых значений некоторого атрибута разбивается на интервалы экспертом в данной предметной области, и база данных преобразуется в новую базу, где присутствуют только двоичные данные. В табл. 1 представлен фрагмент базы данных по результатам анализа больных. В таблице приведены количественные данные, отражающие уровень содержания глюкозы в крови больных.

Таблица 1

Количественные данные

ID	Уровень глюкозы в крови
111	3.2
112	3.6
113	5.3
114	10
115	9.3

В табл. 2 представлены данные, преобразованные на основе интервалов, которые заданы экспертом и содержат возможные значения показателя.

Таблица 2

Данные, разбитые на интервалы

ID	Уровень глюкозы: <3.3	Уровень глюкозы: i	Уровень глюкозы: <5.5
111	1	0	0
112	0	1	0
113	0	1	0
114	0	0	1
115	0	0	1

После построения двоичной базы данных используется бинарный алгоритм для вывода ассоциативных правил, например, алгоритм Априори.

Очевидно, что при таком преобразовании количественных данных теряется весьма полезная информация, связанная со значением признака, а это значение может нести существенную смысловую нагрузку. Чтобы избежать подобной потери, предлагается воспользоваться понятиями нечеткой логики и нечетких множеств, которые более тонко характеризуют результаты измерений, нежели простая разбивка данных на интервалы.

3. Нечеткие ассоциативные правила и их генерация

Для поиска нечетких ассоциативных правил предлагается каждый атрибут в базе данных представлять лингвистической переменной, а каждый интервал значений данного атрибута – лингвистическим термом и нечетким множеством с соответствующей функцией принадлежности [5]. В результате таких преобразований атрибуты в базе данных формулируются в виде лингвистических термов, а сами элементы в кортежах базы данных представляются двоичными данными. В табл. 3 представлены данные, полученные на основе данных из табл. 2.

Таблица 3

Данные, представленные в терминах лингвистической переменной

ID	Уровень глюкозы: «ниже нормы»	Уровень глюкозы: «в норме»	Уровень глюкозы: «выше нормы»
111	1	0	0
112	0	1	0
113	0	1	0
114	0	0	1
115	0	0	1

Заметим, что и здесь, как и в случае количественных данных, для поиска ассоциаций можно применить алгоритм Априори. Так как исходные данные приведены к бинарному виду, то ассоциативные правила будут строиться таким же образом как в случае наличия бинарных исходных данных. Таким образом, и в этом случае ассоциативное правило будет иметь вид $X \rightarrow Y$, где $X \subset I$, $Y \subset I$, X и Y – названия лингвистических термов. Например, ассоциативное правило может быть следующего вида: «Уровень глюкозы ниже нормы» \rightarrow «Назначить сахароповышающее средство». Такое правило будем записывать в виде:

*если x_1 – это «Уровень сахара ниже нормы»,
то y_1 – это « Назначить сахароповышающее средство».*

Для оценки нечетких ассоциативных правил вновь используются характеристики поддержки и достоверности, но с использованием операций над нечеткими множествами [2].

4. Заключение

Сгенерированную систему нечетких ассоциативных правил можем рассматривать как совокупность экспертных знаний. Система нечетких ассоциативных правил вместе с алгоритмом нечеткого логического вывода представляет из себя нечеткую экспертную систему, которая принимает решение по некоторой совокупности входных четких данных. Принятие решения осуществляется на основе логического вывода, реализуемого по методу Мамдани [1]. Входными данными Fuzzy-системы являются вектора, содержащие «четкую» информацию об исходных характеристиках объекта, на выходе системы в результате Fuzzy-вычислений выдается четкое заключение или решение, наиболее подходящее для объекта, представленного входными данными. Fuzzy-вычисления выполняются по методу Мамдани и реализуются в три этапа: фашификация входных данных, нечеткий логический вывод, дефашификация.

Разработанная методика была апробирована на данных из CDC WONDER-базы данных для эпидемиологических исследований, предоставляющая доступ к информационным ресурсам Центров по контролю и профилактике болезней (CDC), доступных работникам здравоохранения и ответственности в целом. Были рассмотрены данные за период 2003-2011 гг. по смертности от хронической обструктивной болезни лёгких по городам США среди населения старше 35 лет. Для данного периода были взяты также климатические данные и данные загрязненности воздуха по городам с целью выяснения взаимосвязей между рассматриваемыми признаками.

Список источников

1. Cordon O., Herrera F., Hoffmann F., Magdalena L. Genetic fuzzy systems evolutionary tuning and learning of fuzzy knowledge bases // *Advances in Fuzzy Systems-Applications and Theory*, 2001, vol. 19. World Scientific Publishing Co. Pte. Ltd.
2. Farzanyar Z., Kangavari M. *Efficient mining of fuzzy association rules from the pre-processed dataset*. Tehran, 2012.
3. Jean-Marc A. *Data Mining for Association Rules and Sequential Patterns: Sequential and Parallel Algorithms*. Springer, 2000.
4. Sheibani R., Ebrahimzadeh A. *An Algorithm For Mining Fuzzy Association Rules*. Tehran, 2008.
5. Zadeh L. *The concept of linguistic variables and its application to approximate reasoning: Part1-3*. Information Sciences, 1975.
6. Блюмин С., Шуйкова И. *Модели принятия решений в условиях неопределенности*. Липецк, ЛЭГИ, 2001.
7. Борисов А.Н., Крумберг О.А., Федоров И.П. *Принятие решения на основе нечетких моделей: примеры использования*. Рига, Знание, 1990.
8. Овакимян А.С., Саркисян С.Г., Зироян М.А., Тинякова В.И. Проектирование систем принятия решений, функционирующих в условиях неопределенности // *Современная экономика: проблемы и решения*, 2014, no. 9 (57), с. 20-28.
9. Полковникова Н.А., Курейчик В.М. Разработка модели экспертной системы на основе нечеткой логики // *Известия ЮФУ. Технические науки*, 2014, no. 1 (150), с. 83-92.
10. Щеглов С.Н. Анализ моделей принятия решений в новых информационных технологиях в условиях нечеткости и неопределенности // *Известия ЮФУ. Технические науки*, 2012, no. 7, с. 126-131.

CREATING OF DECISION MAKING FUZZY-SYSTEMS

Hovakimian Anna Sedrakovna¹, Cand. Sc. (Eng.)

Sarkisian Siranoush Gegamovna², Cand. Sc. (Phys. and Math.)

Ziroyan Alvard Albertovna³, Cand. Sc. (Econ.)

Martinova Marina Alekseevna⁴, Cand. Sc. (Econ.)

¹ Russian-Armenian University, Hovsep Emin st., 123, Yerevan, Armenia, 0051; e-mail:anna.hovakimyan@rau.am

² Yerevan State University, Manoogian st., 1, Yerevan, Armenia, 0025; e-mail: siranushs@ysu.am

³ Russian State Social University, st. Wilhelm Pieck, 4, building 1, Moscow, Russia, 129226; e-mail: ziralla@mail.ru

⁴ Chechen State Pedagogical University, H. Isaeva av., 62, Grozny, Russia, 364068; e-mail: marina-martynova@list.ru

Purpose: to investigate the problem of building decision-making systems which basis of the large amount of data distributed across nodes of the computing cluster and characterizing the subject area. *Discussion:* building decision-making systems that are implemented as expert systems is a laborious task for two reasons. Its solution requires the involvement of highly qualified experts in this subject area to provide the system with expert knowledge. Effective implementation of the logical inference system ensures that the expert system makes the right decision. Therefore, the task of automating the process of generating expert knowledge, constructing and implementing flexible and intelligent decision-making systems is topical. *Results:* authors proposed an approach to generating expert knowledge based on one of the methods of data mining (Data Mining). This approach is universal relative to the subject area. The generated Fuzzy system can predict the values of the required indicators or help to make a decision in the relevant subject area on the basis of available current data.

Keywords: associative rules, Data Mining, cluster data processing, fuzzy sets, fuzzy associative rules, fuzzy logic, decision-making system.

References

1. Cordon O., Herrera F., Hoffmann F., Magdalena L. Genetic fuzzy systems evolutionary tuning and learning of fuzzy knowledge bases. *Advances in Fuzzy Systems-Applications and Theory*, 2001, vol. 19. World Scientific Publishing Co. Pte. Ltd.
2. Farzanyar Z., Kangavari M. *Efficient mining of fuzzy association rules from the pre-processed dataset*, Tehran, 2012.
3. Jean-Marc A. *Data Mining for Association Rules and Sequential Patterns: Sequential and Parallel Algorithms*. Springer, 2000.
4. Sheibani R., Ebrahimzadeh A. *An Algo-*

rithm For Mining Fuzzy Association Rules. Tehran, 2008.

5. Zadeh L. *The concept of linguistic variables and its application to approximate reasoning: Part1–3*. Information Sciences, 1975.

6. Blumin S., Shuikova I. *Modeli prinyatiia reshenii v usloviakh neopredelennosti* [Models of decision-making under conditions of uncertainty]. Lipetsk, LEGI, 2001. (In Russ.)

7. Borisov A.N., Krumberg O.A., Fedorov I.P. *Prinyatie resheniia na osnove nechetkikh modelei: primeri ispol'zovaniia* [Decision-making based on fuzzy models: examples of use]. Riga, Znanie, 1990. (In Russ.)

8. Ovakimyan A.S., Sarkisyan S.G., Ziroyan M.A., Tiniikova V.I. *Proektirovanie system prinyatiia reshenii, funktsioniruiuschikh v usloviakh neopredelennosti* [Designing

decision-making systems that function in the conditions of uncertainty], *Sovremennaya ekonomika: problemi i resheniia*, 2014, no. 9 (57), pp. 20-28. (In Russ.)

9. Polkovnikova N.A., Kureychik V.M. *Razrabotka modeli ekspertnoy sistemy na osnove nechetkoy logici* [Development of a model of an expert system based on fuzzy logic]. *Izvestiya YUFU. Tekhnicheskiye nauki*, 2014, no. 1 (150), pp. 83-92. (In Russ.)

10. Shcheglov S.N. *Analiz modeley prinyatiia resheniy v novykh informatsionnykh tekhnologiyakh v usloviyakh nechetkosti i neopredelennosti* [Analysis of decision-making models in new information technologies in conditions of indistinctness and uncertainty]. *Izvestiya YUFU. Tekhnicheskiye nauki*, 2012, no. 7, pp. 126-131. (In Russ.)