

© 2022 г. Т.В. АФАНАСЬЕВА, д-р техн. наук (afanaseva.tv@rea.ru)
(Российский экономический университет им. Г.В. Плеханова, Москва)

ГРАНУЛЯЦИЯ МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ В ЗАДАЧЕ ДЕСКРИПТИВНОГО АНАЛИЗА СОСТОЯНИЯ И ПОВЕДЕНИЯ СЛОЖНЫХ ОБЪЕКТОВ

Многомерные временные ряды, являясь источником скрытых знаний, могут выступать моделями восприятия объектов во многих прикладных областях. Статья посвящена разработке концептуальных положений гранулярных вычислений многомерных временных рядов, на основе которых предложена методика дескриптивного анализа, позволяющая получать информационные гранулы о состоянии и поведении объекта наблюдения, выраженные в текстовой форме с использованием протоформ. Рассмотрено применение грануляции многомерного временного ряда в дескриптивном анализе развития экономики РФ.

Ключевые слова: многомерный временной ряд, грануляция, протоформа, дескриптивный анализ.

DOI: 10.31857/S000523102206006X, **EDN:** ACSAXP

1. Введение

Одним из направлений в развитии прикладных интеллектуальных систем являются моделирование и реализация когнитивного процесса представления информации об объектах окружающего мира, согласованной с экспертными представлениями. Понятие грануляции возникло как естественная потребность в обобщенном представлении информации, ориентированном на человека, для поддержки процессов понимания данных и преобразования информации в семантически значимые сущности. Информационная грануляция и связанный с этим понятием термин “информационная гранула” были введены Лофти Заде [1] как некоторая общая концепция представления и структурирования информации, характеризующая сложные объекты. Под гранулой в смысле Л. Заде [2] понимается группа концептуально значимых сущностей, объединяемых отношениями неразличимости, эквивалентности, сходства, близости, для определения которой требуется задать обобщенные ограничения. Исходная посылка теории грануляции Л. Заде заключалась в описании связи восприятия человеком объектов окружающего мира и их представлениями в терминах естественного языка в виде информационных гранул (ИГ). Принимая во внимание, что для представления терминов характерна неточность, обусловленная ограниченностью языка, восприятием пространства, времени, движения и когнитивными возможностями человека,

Л. Заде предложил математическую конструкцию для описания лингвистически значимых ИГ в виде пропозиций (коротких предложений), формальную основу которых образуют протоформы [2–4]. Как абстракции данных ИГ не только отражают природу данных, но и могут эффективно фиксировать дополнительные знания предметной области, сообщаемые пользователем [5]. Так, понятие гранулярной мета-онтологии введено в публикации [6], и предложен новый подход к представлению знаний о жизненном цикле сложной технической системы, опирающийся на онтологическое моделирование и теорию грануляции информации. Идея грануляции в представлении знаний о свойствах объектов является ключевой во многих прикладных областях, в частности, в таких как медицина, электронная коммерция, транспорт, управление, в исследованиях по кибербезопасности, в обработке и анализе больших данных, при решении задач сентимент-анализа [7–13]. Развитие направления представления знаний о свойствах объектов в виде ИГ привело к грануляции темпоральной информации, представленной в виде одномерных временных рядов [14–21]. Учитывая сложную структуру временного ряда, исследователи предложили использовать грануляцию для представления знаний о его поведении в лингвистической форме [22–25]. Опираясь на концепцию нечеткой грануляции и операцию обобщения на основе лингвистического резюмирования, был предложен подход к грануляции одномерного временного ряда на основе протоформ с нечеткими квантификаторами [26–30]. Авторы в публикации [31] определяют иерархический язык темпоральных правил для выражения сложных шаблонов, представленных в многомерных временных рядах. Семиотическая иерархия временных паттернов, которые не заданы априори, строится из семиотических троек: уникальный символ, грамматическое правило и определяемая пользователем метка, при этом необходим эксперт, чтобы интерпретировать правило.

Отмечая актуальность и достижения в области гранулярных вычислений, отметим, что решение задачи грануляции объектов, информация о которых представлена в виде многомерных временных рядов, не нашло достаточного отражения в научных публикациях. В то же время многомерные временные ряды (МВР) являются объектом анализа в прикладных задачах, решаемых в рамках различных классов систем: телекоммуникационных, финансовых, образовательных, транспортных, производственных, медицинских, коммерческих, социальных, экологических. Приведем основные положения теории грануляции информации, значимые для дескриптивной аналитики сложных объектов.

1. Информационные гранулы (ИГ) обладают свойствами компактности, структурности, иерархичности, ограниченности, лингвистической интерпретируемости [2]. Выделяют четкие интервальные (с-гранулы), нечеткие гранулы (f-гранулы) и гранулы в виде протоформ (р-гранулы), последние служат основой создания пропозиций, выражающих знания о свойствах объектов в форме предложений естественного языка [3, 24, 29].

2. В основе гранулярных вычислений (Granular Computing) лежит процесс автоматического создания ИГ, последующая обработка которых требует меньше времени, что важно в контексте больших данных [3–5, 18, 27].
3. Процесс грануляции информации основан на моделировании когнитивных операций абстрагирования и обобщения в процессах восприятия и представления свойств объектов [2, 5, 31], что является фундаментальным для интеллектуальных систем.

Поэтому цель данной статьи — разработка концептуальных положений грануляции МВР для извлечения информационных гранул, характеризующих состояние и поведение объекта исследования в лингвистической форме.

2. Задача грануляции МВР в контексте дескриптивного анализа объектов

Определим объект дескриптивного анализа O в виде совокупности элементов $G = \{g_i, i = 1, 2, \dots, gk\}$, описываемых множеством показателей $M = \{m_j, j = 1, 2, \dots, mk\}$, числовые значения которых изменяются на временном интервале $T = [1, tk]$. Тогда модель восприятия объекта O определим в виде многомерного временного ряда:

$$(1) \quad X = \{x_{ijt}, x_{ijt} \subseteq \mathbb{R}, i = 1, 2, \dots, gk; j = 1, 2, \dots, mk; t = 1, 2, \dots, tk\}.$$

В настоящей статье задача дескриптивного анализа объекта заключается в том, чтобы получить знания о состоянии и поведении объекта исследования в виде ИГ. Учитывая, что назначение ИГ — представлять знания, извлеченные из данных, согласованные с экспертными знаниями, на различных уровнях абстракции с использованием лингвистических терминов [3, 5], введем два класса ИГ: *expert-defined* гранулы и *data-extracted* гранулы.

Expert-defined гранулы (е-гранулы) рассматриваются в виде ИГ, основанных на экспертных знаниях о специфике моделей восприятия данных, и определяются типом решаемой задачи. Формальную основу *expert-defined* гранул могут составлять множества, интервалы, нечеткие множества, грубые множества, лингвистические переменные, правила и функции агрегирования [7, 8]. При этом в качестве е-гранул могут применяться с-гранулы и f-гранулы, которые определяют метод грануляции МВР и точку зрения лица, принимающего решения. Результат грануляции МВР будем рассматривать в виде *data-extracted* гранул (d-гранул), которые сжато представляют свойства объекта, распределенные по его показателям и элементам. Чтобы представить эти свойства в контексте состояния и поведения объекта в лингвистической форме, целесообразно использовать пропозиции, формально задаваемые в виде протоформ вида [4, 28]:

$$(2) \quad m \text{ is } Z,$$

$$(3) \quad Qx' \text{ s are } Z,$$

где m представляет показатель объекта (например, энергопотребление), x обозначает некоторую сущность объекта (например, подсистема), Z определяет data-extracted гранулу, описывающую состояние или поведение объекта (например, эффективное или стабильное), Q обозначает data-extracted гранулу, в виде квантификатора (например, большинство), обобщающего сущности с одинаковыми Z . В формулах (2) и (3) глаголы “is” (является) и “are” (являются) определяют отношение принадлежности data-extracted гранулы Z к объектам левой части этих формул. Текстовое выражение этих глаголов зависит от контекста описания свойств объектов, а их примеры приведены в табл. 3.

Тогда постановку задачи грануляции МВР рамках дескриптивного анализа объекта O сформулируем в следующем виде: имея многомерный временной ряд X и набор expert-defined гранул E , заданных на $W \supseteq X$, требуется определить множество data-extracted гранул D , т.е. построить отображение

$$(4) \quad F : X \times E \rightarrow D.$$

3. Концептуальная модель expert-defined гранул

В рамках поставленной задачи анализируемые свойства рассматриваются как некоторые качественные характеристики, которые резюмируют состояние объекта O и его поведение по МВР в лингвистических терминах. Исходя из этого для описания состояния объекта могут использоваться лингвистические термины $Y \in Sy$, описывающие качественные уровни значений X , например, из множества $Sy = \{\text{“низкий”}, \text{“средний”}, \text{“высокий”}\}$, а для описания поведения — лингвистические термины $B \in Sb$, характеризующие тенденции изменения качественных уровней на временном интервале $T = [1, tk]$, значения которых содержатся в множестве $Sb = \{\text{“рост”}, \text{“стабильность”}, \text{“падение”}, \text{“колебание”}\}$. Заметим, что каждый термин из множества Sy обобщает некоторый интервал значений на X , ассоциированный с показателем $m \in M$ объекта в заданный момент времени $t \in T$, в то время как термин из множества Sb обобщает последовательность значений показателя $m \in M$ на заданном временном интервале T .

С каждым лингвистическим термином из множеств Sy и Sb согласно теории ИГ [2] необходимо сопоставить обобщенное ограничение, заданное на X , в виде математического описания, например, с использованием понятия интервала, множества, последовательности, функции, группы, нечеткого или грубого множества [3]. Тогда термины $Y \in Sy$ и $B \in Sb$ вместе с их обобщенными ограничениями rY и rB могут быть использованы для гранулярного представления состояния и поведения объекта O . Например, для термина B могут быть использованы ограничения:

$$B = \begin{cases} \text{рост,} & \text{если } a > c_{\max}, \\ \text{стабильность,} & \text{если } a \in [c_{\min}, c_{\max}], \\ \text{падение,} & \text{если } a < c_{\min}, \end{cases}$$

где a определяет оценку коэффициента в уравнении регрессии $x_t = at + b$; интервал $[c_{\min}, c_{\max}]$ включает значения a , имеющие малую вариабельность относительно некоторой константы.

Чтобы агрегировать множество похожих ИГ, целесообразно использовать протоформы с квантификаторами частотности [26, 30], в которых квантификатор $Q \in Sq$ будет резюмировать множество лингвистически эквивалентных ИГ с использованием лингвистических терминов, например, из множества $Sq = \{\text{“все”}, \text{“большинство”}, \text{“половина”}, \text{“меньшинство”}, \text{“ни одного”}\}$. Учитывая вышеприведенное, определим набор е-гранул E в виде протоформ [4] для представления экспертных знаний о состоянии и поведении объекта:

$$\begin{aligned} E1 : Y \text{ is } rY, \quad Y \in Sy, \quad rY \in R, \\ E2 : B \text{ is } rB, \quad B \in Sb, \quad rB \in R, \\ E3 : Q \text{ is } rQ, \quad Q \in Sq, \quad rQ \in R, \end{aligned}$$

где выражение “ A is rA ” обозначает, что лингвистический термин A ограничен математической конструкцией r , глагол “is” (является) используется для текстового выражения этого ограничения. Так, например, нечеткий терм A “большинство” может быть ограничен функцией принадлежности $r_A(z)$, определенной на множестве частот $[0,1]$ в виде правила:

$$r_A(z) = \begin{cases} 1 & \text{при } z \geq 0,85, \\ 2z - 0,6 & \text{при } 0,4 < z < 0,85, \\ 0 & \text{при } z \leq 0,4. \end{cases}$$

Примеры использования лингвистической переменной как обобщенного ограничения для временного ряда приведены в публикациях [18, 32]. Введенные выше термины образуют терминологический словарь $S = \{Sy, Sb, Sq\}$, используемый для представления свойств в рамках дескриптивного анализа объекта. Заметим, что количество и состав е-гранул могут быть изменены в зависимости от контекста решаемой задачи. Множество математически заданных обобщенных ограничений $R = \{rY, rB, rQ\}$ является ключевым компонентом е-гранул, так как определяет способ грануляции МВР. На основе введенных обозначений определим модель expert-defined гранулы в виде

$$E = \langle Z, R, W, Ig \rangle,$$

где Z обозначает лингвистический термин для анализируемого свойства, а его семантика определяется построенными на некотором множестве значений $W \supseteq X$ обобщенными ограничениями R в рамках выбранной теории грануляции Ig [2, 3, 8].

4. Грануляция МВР для дескриптивного анализа свойств объекта

Учитывая свойство иерархичности ИГ, грануляцию МВР будем рассматривать на нескольких уровнях, соответствующих значению показателя $m \in M$

и множеству показателей M , моменту времени $t \in T$ и временному интервалу наблюдения T , элементу $g \in G$ и множеству элементов G . В этом случае с каждым уровнем декомпозиции МВР сопоставим способ грануляции $f \in F$, генерирующий на основе е-гранул $E = \{E1, E2, E3\}$ data-extracted гранулы, соответствующие терминам терминологического словаря S , которые затем объединяются формальной конструкцией протоформы p . В табл. 1 приведены состав и параметры гранулярных вычислений, распределенных по уровням грануляции МВР “снизу-вверх”. В первом столбце табл. 1 приведены уровни грануляции, для которых указаны представления модели МВР согласно выражению (1). При этом представление уровня 0 соответствует элементу модели МВР, определенному для конкретного показателя в заданный момент времени, в то время как уровень 3 определяет представление МВР (1) в виде множеств всех элементов, показателей и временных интервалов. Уровни грануляции 2 могут ассоциироваться с проекциями МВР по элементу или по атрибуту, или по моменту времени. Как видно из столбца 2 табл. 1, на каждом уровне грануляции МВР происходит обобщение гранул предыдущего уровня, которое выражает, с одной стороны, свойство иерархичности ИГ, а с другой — соответствует иерархической природе МВР, что важно в дескриптивном анализе. Результат информационной грануляции определяется обобщенными ограничениями из множества R и представлен в унифицированной форме в виде d-гранул. Отметим, что свойство иерархичности ИГ проявляется также в наборе генерируемых d-гранул, так как на основе ИГ состояния Y формируются ИГ тенденции B , а ИГ-квантификаторы Q образуются как для гранул-состояния, так и для гранул-тенденций. В последнем столбце табл. 1 приведены виды формируемых протоформ (2) и (3), связывающих d-гранулы и элементы различных уровней грануляции МВР. С использованием вышеприведенной грануляции МВР разработана методика дескриптивного анализа объекта O , позволяющая получать d-гранулы о его состоянии и поведении согласно выражению (4) и представлять эти знания в текстовой форме, она включает следующие этапы:

1. Представление экспертных знаний о состоянии и поведении объекта O в виде е-гранул и разработка способов их применения в гранулярных вычислениях:
 - а) создание терминологического словаря используемых лингвистических терминов $S = \{Sy, Sb, Sq\}$ для дескриптивного анализа в зависимости от контекста задачи;
 - б) определение множества обобщенных ограничений $R = \{rY, rB, rQ\}$ в виде математических выражений для каждого лингвистического термина, входящего в словарь S в рамках используемой теории грануляции [2, 3, 8];
 - в) разработка алгоритмов гранулярных вычислений F для выбранных уровней грануляции МВР согласно табл. 1.
2. Реализация гранулярных вычислений F на МВР, формирование d-гранул и вывод множества пропозиций о состоянии и поведении объекта O :

а) применение алгоритмов грануляции F для получения d -гранул для МВР, используя состав и структуру параметров вычислений, приведенных в столбце 2 табл. 1;

б) формирование текста пропозиций с использованием протоформ P , приведенных в столбце 3 табл. 1.

Предложенная методика обеспечивает описание свойств, характеризующих состояние и поведение объекта исследования на разных уровнях грануляции по МВР в виде предложений на естественном языке. В данной методике ИГ поведения B описывают изменение значений показателя МВР в

Таблица 1. Многоуровневая грануляция МВР X на d -гранулы

Уровень грануляции X	Гранулярные вычисления, формирующие data-extracted гранулы D	Вид формируемых протоформ P
0: $m \in M, g \in G, t \in T$	$z0 = f0(x_{i,j,t}, E1, rY(X)),$ $i, j, t = \text{const}, Y = z0$	$p1: m \text{ is } Y \text{ for } g \text{ for } t$
1: $M, g \in G, t \in T$	$z11 = f11(x_j, z0, E3, rQ(m)),$ $x_j = x_{i,j,t}, i, t = \text{const}, j = 1, 2, \dots mk,$ $Qm's = z11, Y = z0$	$p2: Qm's \text{ are } Y \text{ for } g \text{ for } t$
1: $G, m \in M, t \in T$	$z12 = f12(x_i, z0, E3, rQ(g)), x_i = x_{i,j,t},$ $g \in G, Y \in Sy, x_i = x_{i,j,t}, j, t = \text{const},$ $i = 1, 2, \dots gk, Qg's = z12, Y = z0$	$p3: Qg's \text{ are } Y \text{ for } m \text{ for } t$
1: $T, g \in G, m \in M$	$z13 = f13(x_t, E2, rB(T)), x_t = x_{i,j,t},$ $i, j = \text{const}, t = 1, 2, \dots, tk, B = z13$	$p4: m \text{ is } B \text{ for } g \text{ for } T$
2: $G, M, t \in T$	$z21 = f21(x_{ij}, z11, E3, rQ(m)),$ $x_{ij} = x_{i,j,t}, t = \text{const}, i = 1, 2, \dots gk,$ $j = 1, 2, \dots mk, Qm's = z21, Y = z0$	$p5: Qm's \text{ are } Y \text{ for } t \text{ for } G$
2: $G, M, t \in T$	$z22 = f22(x_{ij}, z12, E3, rQ(g)),$ $x_{ij} = x_{i,j,t}, t = \text{const}, i = 1, 2, \dots gk,$ $j = 1, 2, \dots mk, Qg's = z22, Y = z0$	$p6: Qg's \text{ are } Y \text{ for } t \text{ for } M$
2: $G, T, m \in M$	$z23 = f23(x_{it}, z13, E3, rQ(g)),$ $x_{it} = x_{i,j,t}, j = \text{const}, i = 1, 2, \dots ik,$ $t = 1, 2, \dots, tk, Qg's = z23, B = z13$	$p7: Qg's \text{ are } B \text{ for } m \text{ for } T$
2: $G, T, m \in M$	$z24 = f24(x_{it}, z0, E3, rQ(g)),$ $x_{it} = x_{i,j,t}, j = \text{const}, i = 1, 2, \dots ik,$ $t = 1, 2, \dots, tk, Qg's = z24, Y = z0$	$p8: Qg's \text{ are } Y \text{ for } m \text{ for } T$
2: $M, T, g \in G$	$z25 = f25(x_{jt}, z13, E3, rQ(m)),$ $x_{jt} = x_{i,j,t}, i = \text{const}, j = 1, 2, \dots jk,$ $t = 1, 2, \dots, tk, Qm's = z25, B = z13$	$p9: Qm's \text{ are } B \text{ for } g \text{ for } T$
3: G, M, T	$z31 = f31(x_{ijt}, z23, E3, rQ(g)),$ $i = 1, 2, \dots, gk, j = 1, 2, \dots mk;$ $t = 1, 2, \dots, tk, Qg's = z31, B = z13$	$p10: Qg's \text{ are } B \text{ for } M \text{ for } T$
3: G, M, T	$z32 = f32(x_{ijt}, z24, E3, rQ(m)),$ $i = 1, 2, \dots, gk, j = 1, 2, \dots mk;$ $t = 1, 2, \dots, tk, Qm's = z32, B = z13$	$p11: Qm's \text{ are } B \text{ for } G \text{ for } T$

виде глобальной тенденции отдельного временного ряда и не учитывают его локальные тенденции. В то же время локальные тенденции могут служить эффективным средством для извлечения зависимостей, описывающих их пересечение, совпадение, опережение или отставание в МВР. В рамках предложенной методики эту задачу можно решить в три этапа. На первом этапе определить е-гранулы для локальных тенденций и соответствующие протоформы их описания. На втором этапе при грануляции МВР 1-го уровня во временном ряду показателя получить d-гранулы, выполнив темпоральную декомпозицию на временные интервалы с использованием е-гранул локальных тенденций. На третьем этапе извлечь d-гранулы, характеризующие темпоральные зависимости между локальными тенденциями.

5. Применение грануляции МВР для дескриптивного анализа региональных социально-экономических показателей

Изложенная выше методика была апробирована в задаче анализа состояния и тенденций развития экономики РФ по множеству социально-экономических показателей. Цель дескриптивного анализа в контексте решаемой задачи – получить оценку и выявить проблемы в состоянии и динамике развития экономики по субъектам РФ и показателям, объединенных в группы “информационное общество”, “наука и инновации”, “предпринимательство”, “рынок труда” и “эффективность экономики”. Для исследования были выбраны 15 показателей развития экономики за девять лет с 2010 по 2018 г. по 83 субъектам РФ¹, которые образовали МВР $X(83, 15, 9)$. Следуя методике, представленной в разделе 4, на первом этапе были разработаны нечеткие е-гранулы. В словарь терминов состояния были включены оценки уровня развития экономики субъектов РФ $Sy = \{\text{“низкий”}, \text{“средний”}, \text{“высокий”}\}$, а словарь терминов развития (поведения) содержал значения $Sb = \{\text{“рост”}, \text{“стабильность”}, \text{“падение”}\}$. Лингвистические термины Y моделировались равнобедренными треугольными функциями принадлежности нечетких множеств, носителями которых являлись следующие интервалы: “низкий”: $0 \leq x < 0,5$; “средний”: $0,3 \leq x < 0,9$; “высокий”: $0,8 \leq x \leq 1,2$ (x – нормированное значение показателя). Чтобы получить нечеткое значение термина $B \in Sb$, был использован алгоритм, приведенный в [18], идея которого заключается в преобразовании временного ряда отдельного показателя в нечеткий временной ряд [32] и агрегировании интенсивностей изменений нечетких значений показателей. Словарь квантификаторов частотности Q содержал лингвистические термины $Sq = \{\text{“все”}, \text{“большинство”}, \text{“половина”}, \text{“меньшинство”}, \text{“ни одного”}\}$, в качестве обобщенных ограничений для которых использовались нечеткие множества, построенные на универсальном множестве частотности обнаружения гранул Y или B . Нечеткие термы Q моделировались треугольными функциями принадлежности, параметры которых приведены в табл. 2.

¹ Федеральная служба государственной статистики [Электронный ресурс].

URL: http://old.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/
(дата обращения: 25.03.2021).

Таблица 2. Параметры функций принадлежности квантификаторов

Лингвистические термины квантификатора	Параметры носителя функции принадлежности нечеткого квантификатора Q
все	80, 100, 100
более половины	50, 70, 90
половина	40, 50, 60
менее половины	10, 30, 50
ни одного	0, 0, 20

Таблица 3. Протоформы и пропозиции, полученные в результате грануляции МВР

Протоформа	Пропозиция
$p1: m \text{ is } Y \text{ for } g \text{ for } t$ $m = \text{уровень ЗП}, Y = \text{Ниже Нормы},$ $g = \text{Астраханская область}, t = 2018$	В 2018 г. в Астраханской области показатель уровень ЗП был Ниже Нормы
$p2: Qm's \text{ are } Y \text{ for } g \text{ for } t$ $Qm's = \text{более половины},$ $Y = \text{Ниже Нормы},$ $g = \text{Республике Адыгея}, t = 2017$	В 2017 г. в Республике Адыгея более половины показателей были Ниже Нормы
$p3: Qg's \text{ are } Y \text{ for } m \text{ for } t$ $Qg's = \text{более половины},$ $Y = \text{Ниже Нормы},$ $m = \text{Внутр. затраты на научные исследования и разработки}, t = 2018$	В 2018 г. более половины субъектов имели показатель Внутр. затраты на научные исследования и разработки Ниже Нормы
$p5: Qm's \text{ are } Y \text{ for } t \text{ for } G$ $Qm's = \text{меньше половины},$ $Y = \text{Ниже Нормы}, t = 2018$	В 2018 г. меньше половины показателей экономики были Ниже Нормы
$p7: Qg's \text{ are } B \text{ for } m \text{ for } T$ $Qg's = \text{более половины},$ $B = \text{стабильность},$ $m = \text{Эффективность экономики},$ $T = [2010, 2018]$	С 2010 по 2018 г. более половины субъектов имели тенденцию стабильность в показателе Эффективность экономики
$p10: Qg's \text{ are } B \text{ for } M \text{ for } T$ $Qg's = \text{менее половины},$ $B = \text{негативная},$ $T = [2010, 2018]$	С 2010 по 2018 г. менее половины субъектов имели негативную тенденцию развития

Для всех протоформ вычислялась степень истинности, причем для протоформ с нечетким квантификатором Q использовалась формула лингвистического резюмирования, приведенная в [29]. С помощью порогового значения истинности ($\varepsilon \geq 0,7$) были выбраны протоформы для получения пропозиций, характеризующих состояние и тенденции развития экономики РФ с 2010 по 2018 г., некоторые из которых приведены в табл. 3. Отметим, что при переходе от протоформ к пропозициям в контексте задачи выявления проблем

были введены лингвистические оценки “ниже нормы” для состояний “низкий” и “средний”, и “негативная” тенденция для показателей, имеющих тенденцию “падение”. В результате дескриптивного анализа получены набор d-гранул, характеризующих регионы РФ и социально-экономические показатели с точки зрения наличия или отсутствия проблем в состоянии и динамике развития. Иерархия d-гранул позволяет анализировать объект исследования на разных уровнях абстракции, что является востребованным в системах поддержки принятия решений.

6. Заключение

В статье разработаны концептуальные основы грануляции МВР, расширяющие возможности представления свойств МВР в виде информационных гранул состояния и поведения сложных объектов. Предложена новая методика дескриптивного анализа объектов, основанная на многоуровневой грануляции МВР с использованием введенных expert-defined и data-extracted гранул.

Отличиями предложенной методики являются человекоцентричность, ориентация на поддержку принятия решений, формирование текстовых описаний о состоянии и поведении объектов в виде информационных гранул, которые в дальнейшем могут быть использованы для исследования зависимостей в свойствах объектов. Также отметим возможность сегментирования элементов, входящих в состав объекта, по лингвистически значимым для анализа оценкам из терминологического словаря. Результативность методики дескриптивного анализа на основе многоуровневой грануляции показана при анализе развития экономики в контексте субъектов РФ. Будущие исследования будут направлены на разработку подходов к решению задачи сходства и выявления зависимостей data-extracted гранул для последующего применения в диагностическом предиктивном анализе сложных объектов.

СПИСОК ЛИТЕРАТУРЫ

1. *Zadeh L.* Fuzzy Sets and Information Granularity // *Advances in Fuzzy Set Theory and Appl.*, World Science Publishing, Amsterdam. 1979. P. 3–18.
2. *Zadeh L.* Toward a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic // *Fuzzy Sets and Syst.* 1997. V. 90. P. 111–127.
3. *Zadeh L.* Generalized Theory of Uncertainty (GTU) – Principal Concepts and Ideas // *Computational statistic & Data analysis.* 2006. V. 51. P. 15–46.
4. *Zadeh L.* A Prototype-Centered Approach to Adding Deduction Capabilities to Search Engines – the Concept of a Protoform // *Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2002).* 2002. P. 523–525.
5. *Pedrycz W.* Granular Computing for Data Analytics: A Manifesto of Human-centric Computing // *IEEE/CAA J. Autom. Sinica.* 2018. V. 5. No. 6. P. 1025–1034.
6. *Федотова А.В., Ветров А.Н., Тарасов В.Б.* Грануляция информации при моделировании жизненного цикла сложных технических систем // *Науковедение.* 2013. № 5 (18).

7. *Pedrycz W., Skowron A., Kreinovich V.* Handbook of Granular Computing. Wiley, 2008.
8. *Dubois D., Prade H.* Bridging Gaps Between Several Forms of Granular Computing // *Granul. Comput.* 2016. V. 1. P. 115–126.
<https://doi.org/10.1007/s41066-015-0008-8>
9. *Yen G., Beliakov G., Triguero I., Pratama M., Zhang X., Li H.* Data Mining and Granular Computing in Big Data and Knowledge Processing. 2019.
<https://doi.org/10.1109/ACCESS.2019.2908776>
10. *Pedrycz W.* Information Granules and Their Use in Schemes of Knowledge Management // *Scientia Iranica.* 2011. V. 18. No. 3. P. 602–610.
11. *Han Liu, Mihaela Cocca.* Fuzzy Information Granulation Towards Interpretable Sentiment Analysis // *Granul. Comput.* 2017. V. 2. No. 4. P. 289–302.
<https://doi.org/10.1007/s41066-017-0043-8>
12. *Бутенков С.А.* Структурная организация гранулированных вычислений при обработке данных на реконфигурируемых вычислительных системах // *Изв. ЮФУ. Технич. науки.* 2018. № 8. С. 250–262.
13. *Бутакова М.А., Климанская Е.В., Чернов А.В.* Формальные структуры и представления для гранулярных вычислений // *Современные наукоемкие технологии.* 2018. № 5. С. 36–40.
14. *Bargiela A., Pedrycz W.* Granulation of Temporal Data: a Global View on Time Series // 22nd Int. Conf. of the North American Fuzzy Information Processing Society. 2003. P. 191–196. <https://doi.org/10.1109/NAFIPS.2003.1226780>
15. *Donga R., Pedrycz W.* A Granular Time Series Approach to Long-term Forecasting and Trend Forecasting // *Physica A.* 2008. V. 387. P. 3253–3270.
16. *Al-hmouz R., Pedrycz W.* Models of Time Series with Time Granulation // *Knowledge and Inform. Syst.* 2016. V. 48. No. 3. P. 561–580.
<https://doi.org/10.1007/s10115-015-0868-x>
17. *Ярушкина Н.Г. и др.* Интеграция нечетко-гранулярных и онтологических методов в задаче анализа временных рядов // *Автоматизация процессов управления.* 2015. № 2 (40). С. 72–79.
18. *Afanasyeva T., Moshkina I.* Descriptive Model of Temporal Features of Multivariate Time Series Based on Granulation // *CEUR Workshop Proc.* 2020. V. 2667. P. 287–292.
19. *Ярушкина Н.Г., Афанасьева Т.В., Тимина И.А.* Нечеткая грануляция в моделировании и прогнозировании объема телекоммуникационного трафика // *Наукоемкие технологии.* 2013. Т. 14. № 5. С. 67–72.
20. *Ярушкина Н.Г., Афанасьева Т.В.* Гранулярное моделирование временных рядов // *Тринадцатая национальная конф. по искусственному интеллекту КИИ-2012.* 2012. С. 143–148.
21. *Онтологический и нечеткий анализ слабоструктурированных информационных ресурсов / под науч. ред. Н.Г. Ярушкиной.* Ульяновск: УлГТУ, 2016.
22. *Jun M., LiXia W., XiuKun W., TsauYoung L.* Granulation-based Symbolic Representation of Time Series and Semi-supervised Classification // *Computers & Math. with Appl.* 2011. V. 62. No. 9. P. 3581–3590.
<https://doi.org/10.1016/j.camwa.2011.09.006>
23. *Pedrycz W., Homenda W., Jastrzebska A., Yu F.* Information Granules and Granular Models: Selected Design Investigations // 2020 IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE). 2020. P. 1–8. <https://doi.org/10.1109/FUZZ48607.2020.9177696>

24. *Novak V.* Linguistic Characterization of Time Series // *Fuzzy Sets and Syst.* 2016. V. 285. P. 52–72.
25. *Glockner I., Knoll A.* Fuzzy Quantifiers for Data Summarization and Their Role in Granular Computing // *Proc. Joint 9th IFSA World Congr. and 20th NAFIPS Int. Conf.* 2001. V. 4. P. 2029–2034. <https://doi.org/10.1109/NAFIPS.2001.944380>
26. *Kacprzyk J., Wilbik A., Zadroiny S.* Linguistic Summarization of Time Series Under Different Granulation of Describing Features // *RSEISP 2007.* 2007. V. 4585. P. 230–240. https://doi.org/10.1007/978-3-540-73451-2_25
27. *Kacprzyk J., Zadrozny S.* Linguistic Summaries of Time Series: A Powerful Tool for Discovering Knowledge on Time Varying Processes and Systems // *Informatyka Stosowana.* 2014. V. 1. P. 149–160.
28. *Kacprzyk J., Wilbik A., Zadroiny S.* Linguistic Summarization of Time Series Using a Fuzzy Quantifier Driven Aggregation // *Fuzzy Sets and Syst.* V. 159. No. 12. P. 1485–1499.
29. *Afanasieva T.V., Rodionova T.E.* Methodology of Patient-oriented Assessment of Cardiovascular Health of Men Using Fuzzy Sets and Formal Conceptual Analysis // *World Scientific Proc. Series on Computer Engineering and Information Science Developments of Artificial Intelligence Technologies in Computation and Robotics.* 2020. P. 857–865.
30. *Zadeh L.* A Computational Approach to Fuzzy Quantifiers in Natural Languages // *Computers and Math. with Appl.* 1983. V. 9. P. 149–184.
31. *Mörchen F., Ultsch A.* Mining Hierarchical Temporal Patterns in Multivariate Time Series. 2004. V. 3238. P. 127–140.
32. *Song Q., Chissom B.* Fuzzy Time Series and Its Models // *Fuzzy Sets and Syst.* 1993. V. 54. P. 269–277.

Статья представлена к публикации членом редколлегии О.П. Кузнецовым.

Поступила в редакцию 26.11.2021

После доработки 11.01.2022

Принята к публикации 26.01.2022