

УДК (519.7+004):332

DOI: 10.31359/2312-3427-2019-4-1-297

В.М. Дьоміна, канд. техн. наук, доцент

vvdemina17@gmail.com

<https://orcid.org/0000-0001-6467-5021>

Харківський національний аграрний університет ім. В.В Докучаєва

І.О. Побіженко, канд. техн. наук, доцент

irina_pob@ukr.net

<http://orcid.org/0000-0002-0723-1878>

Т.Г. Білова, канд. техн. наук, доцент

belovatat@ukr.net

0000-0002-1085-7361

Харківська державна академія культури

Т.О. Дьоміна

t.dyomina241@gmail.com

0000-0003-3375-3749

Харківський національний університет ім. В.Н. Каразіна

**ЧИСЛОВІ МОДЕЛІ ОПИСУ РЕАЛЬНОСТІ,
ЯКІ ДОЗВОЛЯЮТЬ ОБРОБКУ ВЕЛИКИХ ОБСЯГІВ ДАНИХ
ЗА УМОВ ЗРОСТАННЯ ЇХ НЕВИЗНАЧЕНОСТІ
ПРИ УПРАВЛІННІ РЕГІОНОМ**

У статті запропоновано поєднання теорії і емпіричного досвіду методів управління із сучасним пізнавальним інструментарієм для прийняття управлінських рішень в умовах мінливості, невизначеності та багатомірності ринкового середовища. Зроблено спробу застосувати системно-структурний аналіз у системі моделей стратегічного управління розвитком регіону.

***Ключові слова:** моделі опису реальності, великі дані, умови невизначеності, нечітка логіка, територіальний розвиток.*

Постановка проблеми. Сучасний етап соціально-економічного розвитку України характеризується наявністю просторових деформацій на регіональному та внутрішньо-регіональному рівнях. Однак, якщо проблема асиметричності регіональних трендів розвитку широко висвітлюється, а на

її подолання спрямовані механізми державної регіональної політики, но питання внутрішньо-регіональних соціально-економічних диспропорцій залишається недостатньо дослідженим. Водночас саме на внутрішньо-регіональному рівні диференціація за низкою показників економічного і соціального розвитку адміністративно-територіальних утворень значно перевищує міжрегіональні значення. Така ситуація обумовлює існування розірваності і фрагментарності соціально-економічного простору регіонів, розширення ареалів депресивності, що пов'язано з концентруванням фінансових ресурсів у центрах розвитку регіонів (зокрема великих містах), посилення розриву між міськими та сільськими територіями, поглиблення проблем територіальної розбалансованості на внутрішньо-регіональному рівні [1].

Зазначені проблеми особливо загострюються в контексті реалізації реформи місцевого самоврядування та адміністративно-фінансової децентралізації, основним принципом яких є розвиток територіальних громад. Нові економічні умови змусили дослідників і практиків у нашій країні шукати адекватні форми, методи й інструменти стратегічного управління в регіоні, зокрема намагатися по досвіду зарубіжних країн застосовувати принципи і методи стратегічного управління, які реалізовані на рівні підприємств до регіональних систем [2]. Звичайно, існує пряма аналогія між регіональною соціально-економічною системою і господарською системою, якою є корпорація проводити складно, але вивчення зарубіжної теорії й методик стратегічного управління й планування необхідно для розробки вітчизняної концепції стратегічного управління, адаптаційної методологічної роботи на рівні регіональних досліджень особливо в умовах опису реальності, які дозволяють обробку великих обсягів даних за умов зростання їх невизначеності [3]

Аналіз останніх досліджень і публікацій. Вагомий внесок у дослідження питань, пов'язаних з розвитком регіонів та впровадженням стратегічного планування на регіональному рівні, зробили зарубіжні вчені: Ансофф І., Брайсон Ф., Бріан В., Варда Я., Карлоф Б. та вітчизняні – Беседін В.Ф., Кравченко В. І., Леонідов І. Л., Невелєв О. М., Панасюк Б. Я., Склярук Н. І., Симоненко В., Удовиченко В. П., Черевко О. В., Шаров Ю., Щелкунов В. І., Щукін Б. М., Яскала І. та ін. Водночас низка проблем теоретичного і методологічного характеру, пов'язаних із оцінюванням внутрішньорегіональної диференціації соціально-економічного розвитку, вироблення нових підходів до стимулювання економічної активності територіальних громад та ефективного використання власного потенціалу у різних сферах діяльності, потребують подальшого вивчення.

Процес побудови узагальненої (комплексної) моделі регіону ускладнюється різноманітністю моделей даних, а також через наявність різних рівнів агрегації даних. Однією з популярних технологій для розроблення

Формулювання цілей статті: набуття теоретичних висновків та практичних укладань про тенденції закономірностей процесів стратегічного управління виробничо-економічної системи регіонального рівня на базі комплексу математично-статистичних методів і моделей.

Виклад основного матеріалу досліджень. У певних географічних сегментах, таких як місто, регіон, країна прикладне застосування систем територіального розвитку сприяє швидкому поширенню знань, навичок та оптимальних практик [4]. Сьогодні стратегічне управління розглядають, передусім, як функцію, що набуває нового змісту, збагачується коштом синтезу з гуманізованим підходом, використовує творчий, науковий, а іноді і інтуїтивний підхід.

Для комплексного аналізу інформації на рівні регіону необхідно спочатку опрацювати її з реляційних, багатовимірних баз даних, баз геоданих, структурованих і слабоструктурованих текстових файлів; проаналізувати її, використовуючи як консолідований, так і декатенаційний підхід отримання [5].

Великі дані завжди обмежені терміном, що використовується для ідентифікації наборів даних, тому існують методики і програмні засоби як для передачі даних, так і для набуття інформаційних гранул з Великих даних (колекції об'єктів, які зазвичай формуються для атрибутів з числовими даними і які розташовані поряд через їх схожість, функціональну або фізичну суміжність). Проте нерозв'язаною задачею залишається задача побудови опису між моделями відображених даних різних джерел.

Аналізуючи можливість використання Великих даних в системах територіального управління, є:

– великий набір сутностей: особистості, організації (фізичні, юридичні), природні ресурси (річки, ліси, озера), рекреаційний фонд (історичні пам'ятки, санаторії), законодавчі акти та звіти, дані;

– база даних особливостей: документи для інтелектуального аналізу даних, онтологічні терміни, словники даних, що зв'язують деякі об'єкти.

Використовуючи дану інформацію, з метою її подальшого аналізу, потрібно вирішити питання формалізації: перелік сутностей і зв'язки послідовності між собою [6].

Одним із способів представлення Великих даних з точки зору математичних методів, що використовують багатовимірну модель даних є гіперкуб даних [7]. Нехай основними поняттями багатовимірної моделі даних є: гіперкуб даних (rel), вимір Ω , атрибут Λ , комірка Ξ , значення $rel(\Omega, \lambda) \in rel(V, A)$.

Гіперкуб даних містить один або більше вимірів і є впорядкованим набором комірок. Кожна комірка визначається одним і лише одним набором значень вимірів – атрибутів. Комірка може містити дані - значення або бути порожньою.

Під виміром у [8] розуміють множину атрибутів, що утворюють одну з граней гіперкуба. Прикладом часового виміру є список днів, місяців, кварталів. Прикладом географічного виміру може бути перелік територіальних об'єктів: населених пунктів, районів, регіонів, країн та ін.

Отже Λ – множина вимірів гіперкуба, $\Lambda_{\Omega} = \{\Lambda_1, \Lambda_2, \dots, \Lambda_{m_1}\}$, $i = 1, \dots, n$ – множина атрибутів виміру Ω_i : $\Lambda = \Lambda_{\Omega_1} \cup \Lambda_{\Omega_2} \cup \dots \cup \Lambda_{\Omega_n}$ – множина атрибутів гіперкуба, $\Omega' \subseteq \Omega$ – множина фіксованих вимірів, $\Lambda' \subseteq \Lambda$ – множина фіксованих атрибутів. Гіперкуб даних позначимо як множину комірок, що відповідає множинам $\Omega_i, \Lambda: rel(\Omega, \Lambda)$. Підмножина гіперкуба даних, що відповідає множині фіксованих значень, позначається як: $rel'(\Omega', \Lambda')$.

Кожній комірці гіперкуба даних $rel' \in rel$ відповідає єдино можлива множина атрибутів вимірів $\Lambda_{rel} \subset \Lambda$. Комірка може бути порожня (не містити даних) або містити значення показника [8]. Множину комірок гіперкуба $rel(\Omega, \Lambda)$ позначимо як $\Omega(rel)$. Для отримання доступу до даних користувачу необхідно вказати множину необхідних вимірів $\Omega' \subseteq \Omega$ і значень атрибутів $\Lambda' \subseteq \Lambda$ (фіксувати атрибути). Множину комірок, що

відповідають відповідним атрибутам та вимірам, позначимо як $rel(\Omega', \Lambda') | rel' \subseteq rel$.

Ключем виміру є атрибут, який однозначно визначає кортеж (рядок) виміру гіперкуба. Гіперкуби підтримують ієрархію вимірів і формул без дублювання їх визначень. Набір відповідних гіперкубів складає багатовимірну базу даних (або сховище даних). Комбінації значень вимірів визначають комірки гіперкуба. Залежно від конкретного застосування (прикладної програми), комірки в гіперкубі можуть розташовуватися як розріджено, так і щільно. Гіперкуби, як правило, стають розрідженими у міру збільшення числа розмірностей і ступеня деталізації значень вимірів. Багатовимірне представлення даних добре використовувати для задач візуалізації даних та їх аналізу, але у зв'язку з розрідженістю гіперкуба [8] обсяг даних у такому випадку є більший порівняно з реляційним представленням, що є неприпустимим до Великих даних.

Інформація про модельовані об'єкти подається множиною складних значень [9], що зберігаються в ідентифікованих змінних (у подальшому – об'єкти даних). Можливе використання нечітко-множинних моделей, що можуть допомагати приймати економічно обґрунтовані рішення. Об'єкт даних θ опишемо як $\{id, meta, \theta\}$, де id – унікальний ідентифікатор зазначеної змінної, $meta$ – інформація, що описує структуру зазначеної змінної (метадані), де перелічуються атрибути об'єкта даних, θ – складне значення, що описує стан об'єкта.

Об'єкт θ можна описати як

$$\{id, meta, \{\psi_1, \psi_2, \dots, \psi_n\}\},$$

де ψ_i – деяке значення відношень Ψ_i , що виникає у результаті нормалізації значення θ [10]. Можна розглянути кожне зі значень r_i як значення іменованого атрибута об'єкта θ . Тоді атрибут має розглядатися як змінна відношень, а схема об'єкта даних повинна перелічувати імена атрибутів. Відповідно, об'єкт θ буде описуватися як

$$\theta = \{id, \{\lambda_1, \lambda_2, \dots, \lambda_n\}, \{\psi_1, \psi_2, \dots, \psi_n\}\},$$

де ψ_i – значення відношень Ψ_i із схемою $(\lambda_1^\psi : D_1, \lambda_2^\psi : D_2, \dots)$, $\psi_i = relval(\Psi_i)$, D_j – домени з множини D доменів, не обов'язково різні.

Відзначимо, що в системі Θ може одночасно існувати множина не

обов'язково різних значень Ψ того самого відношення Ψ , що є значеннями атрибутів різних об'єктів або значеннями різних атрибутів того самого об'єкта.

Розглянемо множину Θ усіх вхідних у систему об'єктів. Кожному об'єктові $\theta_i (\theta_i \in \Theta)$ ставиться у відповідність схема об'єкта H : $H_i = \lambda Schema(\theta_i)$, що є власною підмножиною множини $\lambda^0 H \subseteq \lambda^0$. Та сама схема H може бути поставлена у відповідність багатьом об'єктам. Класом P , $P \subseteq \Theta$ називають множину об'єктів з однаковою схемою: об'єкти $\theta_i \{id, meta, \Omega_i^0\}$ і $\theta_j \{id, meta, \Omega_j^0\}$ належать до того самого класу P тільки тоді, коли $\lambda Hschema(\theta_i) = \lambda Hschema(\theta_j)$. Отже, можна стверджувати, що між множиною схем і множиною класів існує функціональна відповідність $H_i = Hschema(P)$.

Також ми можемо говорити про існування множини Ψ усіх визначених на множині доменів D відношень Ψ , значення яких є значеннями атрибутів λ^0 об'єктів із множини Λ . Кожному імені λ^0 ставиться у відповідність одне і тільки одне відношення Ψ з множини Ψ , тим самим обмежуючи множину можливих значень атрибута з іменем oa_i значеннями цього відношення Ψ . Оскільки мова йде про пару «множина - визначене на цій множині значення», відношення Ψ можна позначити як домен атрибута λ^0 . Однак, оскільки значення, визначене на цьому домені, являє собою значення відношень (relval) і також є множиною, будемо називати відношення Ψ *реляційним доменом* атрибута

$$\lambda^0, \Psi = \psi dom(\lambda^0).$$

Можливий випадок, що домени різних атрибутів збігаються $\psi dom(\lambda_i^0) = \psi dom(\lambda_j^0)$, $\lambda_i \neq \lambda_j$.

Отже, ми можемо говорити про функціональну відповідність між множиною атрибутів λ^0 і множиною відношень Ψ ,

$$\Psi = \psi dom(\lambda^0).$$

Ця відповідність повинна вказуватися в схемі об'єкта, що буде перелічувати пари $\lambda^0 : \Psi$, де $\Psi = \Psi dom(\lambda)$.

Звідси випливає, що

– об'єкт має унікальний ідентифікатор, і, отже, у системі можуть існувати два об'єкти з абсолютно однаковим власним значенням,

– значення атрибутів об'єктів є значеннями відношень, що є множиною кортежів - іншими словами атрибут можна розглядати як групу повторення.

Висновки. Для опису об'єкта, щоб розкрити ідентифікаційні параметри регіону, можна використовувати багатовимірну модель даних та об'єктне подання даних. В ситуаціях, пов'язаних із наявністю різного роду невизначеностей, коли ці невизначеності не можуть бути чітко формалізовані за допомогою методів торії ймовірності і математичної статистики можна використовувати методи нечіткої логіки.

Бібліографічний список.

1. Тибінка І.Я. Концептуальні засади регулювання внутрішньо-регіональної соціально-економічної диференціації в умовах активізації місцевого самоврядування // Економічний форум. 2016. №1. С.77–84.

2. Оглих В.В., Ефанова Т.И. Моделирование процессов развития региона в контексте стратегии планирования // Математичні методи, моделі та інформаційні технології в економіці. № 12(174), 2015. С. 430-436.

3. Дьоміна В.М., Ачкасова А.М., Дьоміна Т.О. Розвиток туристичного потенціалу регіону: інструменти електронного маркетингу // Вісник Волинського інституту економіки та менеджменту. № 25. 2019. С. 47-55.

4. Дьоміна В.М. Застосування системно-структурного аналізу у системі моделей стратегічного управління розвитком регіону // Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: тези доповідей XXVII міжнародної наук.-практич. конф. MicroCAD-2019, 15-17 травня 2019 р.: у 4 ч. Ч. IV. / за ред. проф. Сокола Є.І. Харків: НТУ «ХП». С. 147.

5. Болюбаш Ю.Я. Методи та засоби опрацювання великих даних у системах територіального управління // Науковий вісник Національного лісотехнічного університету: збірник наукових праць. Львів : РВВ НЛТУ України. 2016. Вип. 26.4. С. 341-354.

6. Big Data от А до Я. Часть 1: Принципы работы с большими данными, парадигма MapReduce URL: <https://habrahabr.ru/company/dca/blog/267361>. (Дата обращения: 27.11.2019).

7. Jalal-Kamali A. How to Understand Connections Based on Big Data / Jalal-Kamali A., Hossain M.S., Kreinovich V. // Cliques to Flexible Granules / Department of Computer Science, University of Texas at El Paso, El Paso. - 2014. Т.10. Р. 63-87.

8. Maté, A. A Novel Multidimensional Approach to Integrate Big Data in Business Intelligence // Journal of Database Management (JDM). 2015. Vol. 26.2. P.14-31.

9. Chang Fay, Dean Jeffrey, Ghemawat Sanjay, Hsieh Wilson C., Wallach, Deborah A., Burrows Michael, Chandra Tushar, Fikes Andrew, Gruber Robert E. Bigtable: A Distributed Storage System for Structured Data URL: <http://static.googleusercontent.com/media/research.google.com/ru//archive/bigtable-osdi06.pdf>. (Last accessed : 27.12.2019)

10. Кузнецов С. Пространства данных: исследовательский полигон или путь к новому поколению систем управления данными? URL: <http://synthesis.ipi.ac.ru/sigmod/seminar/s20060420>. (Дата обращения: 27.11.2019).

В.М. Дёмина, И.А. Побиженко, Т.Г. Белова, Т.А. Дёмина. Числовые модели описания реальности, позволяющие обрабатывать большие объемы данных в условиях роста их неопределенности при управлении регионом. В статье изложена попытка объединить теоретические подходы и эмпирический опыт методов управления с современным познавательным инструментарием для принятия управленческих решений в условиях изменчивости, неопределенности и многомерности рыночной среды. Сделана попытка применить системно-структурный анализ в системе моделей стратегического управления развитием региона.

Ключевые слова: модели описания реальности, большие данные, условия неопределенности, нечеткая логика, территориальное развитие.

Dyomina V.M., Pobizhenko I.O., Belova T.G., Domina T. Numerical models of the description of the reality, allowing to process large volumes of data in the conditions of growing their uncertainty in the management of

the region. The article describes the problem of asymmetry of regional development trends. There are mechanisms of state regional policy, but the issue of internal regional socio-economic imbalances needs to be investigated.

The fragmentation and fragmentation of the socio-economic space of the regions is observed, the areas of depression are expanding. This is due to the concentration of financial resources in the regional development centers (in particular large cities) and entails the widening of the gap between urban and rural territories, the deepening of the problems of territorial imbalance at the intra-regional level.

For a comprehensive analysis of information at the regional level, you must first process it from relational, multidimensional databases, geodatabases, structured and poorly structured text files. Then analyze it using both a consolidated and a decanation approach. The task of constructing a description between the reflected data models of various sources remains an unsolved problem.

The article uses mathematical methods. BigData is presented as a hypercube of data using a multidimensional model.

A dimension is shown as a set of attributes that forms one of the faces of a hypercube. An example of a geographical measurement is given, this is a list of territorial objects: settlements, districts, regions, etc. Information about the modeled objects is provided by a set of complex values that are stored in data objects. It is possible to use fuzzy-multiple models that can help make economically sound decisions.

It is proved that the object has a unique identifier, and therefore, two objects with exactly the same eigenvalue can exist in the system, the values of the attributes of the objects are the values of relations, is a set of tuples that is, the attribute can be considered as a group of repetitions.

To describe the object, using the identification parameters of the region, you can use a multidimensional data model and object data representation. In situations involving various kinds of uncertainties, when these uncertainties cannot be clearly formalized using the methods of probability theory and mathematical statistics, fuzzy logic methods can be used.

Key words: models for describing reality, BigData, uncertainty conditions, fuzzy logic, territorial development.

Стаття надійшла до редакції: 15.09.2019 р.