УДК 303.425.4:519.25:303.092.8

Г.В. АВЕРИН, А.В. ЗВЯГИНЦЕВА

G.V. AVERIN, A.V. ZVIAGINTSEVA

**О ПОСТРОЕНИИ ФЕНОМЕНОЛОГИЧЕСКИХ МОДЕЛЕЙ**

**ДЛЯ ОПИСАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ**

**ABOUT THE CONSTRUCTION OF PHENOMENOLOGICAL MODELS FOR  DESCRIBING SOCIO-ECONOMIC SYSTEMS**

*Показано, что актуальное направление применения методов естественных наук в информатике связано с развитием феноменологических подходов описания темпоральных массивов данных, характеризующих системы различной природы. Предложен метод построения моделей социально-экономических систем, основанный на представлении дискретных данных в многомерных пространствах состояний и использовании эмпирических мер для комплексной характеристики состояний объектов. Метод основан на применении естественнонаучного принципа соответственных состояний. Возможности метода показаны на примерах построения моделей описания статистических данных при комплексной оценке и прогнозировании развития стран, регионов и городов, изучении человеческого развития и ранжировании социально-экономических объектов по комплексу показателей, анализе загрязнения природной среды городов по показателям экологической безопасности и т.д.*

*Ключевые слова: социально-экономические системы, темпоральные массивы данных, феноменологические методы, модели описания данных.*

*It is shown that the actual direction for the application of methods of natural sciences in computer science is associated with the development of phenomenological approaches to the description of temporal data sets characterizing systems of different nature.The method for constructing models to socio-economic systems is proposed, which is based on the representation of discrete data in multidimensional state spaces and the use of empirical measures for the complex characteristics of the states of objects. The method is based on the application of the natural science principle of the corresponding states. The possibilities of the method are shown in examples of models for describing statistical data in the comprehensive assessment and forecasting of the development of countries, regions and cities, the study of human development and ranking of socio-economic objects in a set of indicators, the analysis of pollution of the natural environment of cities in terms of environmental safety and etc.*

*Keywords: socio-economic systems, temporal data sets, phenomenological approaches, describing statistical data.*

Одно из самых актуальных направлений применения методов естественных наук в информатике и, в частности, феноменологических подходов связано с описанием темпоральных данных, характеризующих системы различной природы [1–7]. В данном вопросе можно выделить два аспекта.

Первый – современная теоретическая информатика не имеет общей теории описания данных, хранящихся в базах данных и характеризующих системы различной природы. Данный аспект касается проблемы получения знаний на основе имеющейся обширной количественной информации. В области анализа процессов развития социально-экономических систем сегодня накоплены обширные темпоральные массивы данных, которые позволяют строить модели таких систем. Например, при изучении социально-экономического состояния регионов Российской Федерации используют базы данных показателей государственных и муниципальных образований, которые имеются в открытом доступе в Федеральной службе государственной статистики [8]. Статистическая база данных социально-экономических показателей субъектов Федерации разделена на 25 групп показателей. Она включает информацию по каждому из 83 субъектов для 438 показателей за последние 16 лет (c 2002 по 2017 гг.). Аналогичным образом для 169 городов Российской Федерации социально-экономическая информация за период c 2003 по 2017 годы разделена на 21 группу показателей, всего 276 показателей.

Второй аспект затрагивает проблему сущности представления об «интеллектуальности», применительно к информационной системе, которое должно выражаться не только в сложности и эффективности применяемых вычислительных методов и алгоритмов обработки данных, но и в понимании процессов функционирования анализируемых объектов. Последнее связано с применением адекватных моделей для описания целых классов систем по отношению к эмпирическим данным, представленным в виде темпоральных массивов информации.

Темпоральные базы данных – это базы, хранящие временные данные. В очень широком смысле – это произвольные данные, которые явно или неявно связаны с определенными датами или промежутками времени. Особенность таких данных в том, что они несут в себе информацию о любых процессах, происходящих в природе и обществе.

В настоящее время актуален поиск новых идей в области изучения сложных систем, которые позволили бы привлечь множество опытных данных по исследованию динамических процессов. В этой области основной проблемный вопрос может быть сформулирован в виде: Можно ли применить общую структурно-логическую схему к модельному описанию темпоральных данных, характеризующих системы различной природы?

Теоретические методы в этой области анализа данных могут предоставить естественные науки, потому, что они обладают развитым методом феноменологического описания систем. Известно, что феноменологический метод обычно рассматривается как способ макроскопического описания объектов, процессов и явлений, основанный на данных опыта или наблюдений и позволяющий получить прикладной результат с необходимой точностью без объяснения истинных причин явления.

Основная идея исследования связана с представлением массивов дискретных выборочных данных в многомерных пространствах состояний относительно переменных состояния, а также с использованием эмпирических мер для комплексного описания состояний систем и построения моделей таких систем.

Будем рассматривать эмпирические данные наблюдений темпоральной структуры. В этом случае речь идет о массивах дискретных данных, которые имеют структуру таблиц в виде «объекты-показатели», причем соответствующее количество таблиц *h* упорядочено по времени с определенным шагом, равным некоторому временному диапазону [5–7, 9]. В таких данных время выступает системообразующим фактором по отношению к изменению и развитию состояний объектов различной природы.

Цель решаемой задачи сформулировать методические принципы обработки и анализа темпоральных массивов опытных данных для получения количественных закономерностей, которые бы имели общий характер и не были бы явно привязаны к той или иной области знаний. Ранее такие принципы были отработаны для физических, биологических и социальных систем [5–7, 9, 11–19], исходя из представлений, что закономерности поведения социально-экономических систем описываются феноменологическими моделями и наблюдения могут быть структурированы в темпоральных (исторических) базах данных.

Для темпоральных данных можно построить общую среду моделирования в виде многомерного пространства состояний объектов. Предположим, что для *m* объектов одного класса в темпоральных массивах данных содержится количественная информация о *n* показателях *zk* (*k=1,2,…,n*), характеризующих свойства изучаемых объектов. Примем эти величины в качестве переменных состояния.

Множество *n* переменных для параметров свойств задает *n*-мерное пространство состояний *En*, где *z=*(*z1,z2,…,zn*) *zϵEn*. Точки этого пространства соответствуют *n*-мерным наборам значений всех переменных *z1,z2,…,zn*. Будем считать, что полную информацию о состоянии объектов несут в себе не только данные, определяемые переменными *z1,z2,…,zn*, но и различные величины и отдельные значимые события.

Особенностью предлагаемого в данном докладе подхода является принятие допущения, что существуют комплексные меры, характеризующие в целом состояния объектов или систем. Будем называть такие величины эмпирическими мерами.

Эмпирическая мера представляет собой системную величину, например, эмпирическую температуру, эмпирическое время, количество теплоты, статистическую вероятность событий и т.д. Эта величина однозначно характеризует состояния объектов в определенном аспекте, зависит от параметров атрибутивных свойств и может быть измерена. Математически эмпирическую меру будем рассматривать в виде функции многих переменных, при этом эта величина должна быть инвариантом в пространстве состояний и комплексно характеризовать состояния объектов. Эмпирическая мера отражает распределения дискретных опытных данных.

Будем считать, что показатели свойств объектов и эмпирические меры могут быть связаны в пространствах состояний между собой функциональными зависимостями вида:

 или , (1)

где *τ* – время.

Также при построении феноменологических моделей используем следующие гипотезы:

* cуществует возможность представления темпоральных массивов как дискретной выборки данных опыта из многомерного пространства состояний (принцип континуальности пространства состояний системы);
* темпоральные данные могут формировать в пространстве состояний некий геометрический «образ», который может быть описан с помощью моделей или алгоритмов (принцип инвариантности);
* феноменологический подход позволяет в многомерном пространстве описать данные с необходимой точностью и представить их в виде моделей в пространстве состояний (принцип изоморфизма, как наличие сходства в состояниях, принцип соответственных состояний).

Для сравнения состояний объектов между собой можно использовать следующие эмпирические меры относительно опорных состояний или всего множества объектов в целом:

* геометрические величины (евклидовое, манхэттенское, степенное, экспертное и др. расстояния);
* эмпирические величины (мера относительных изменений некоторого показателя, температура объекта, количество теплоты, энергия и т.п.);
* вероятностные величины (статистическая вероятность наблюдения состояния объекта в определенном объеме пространства *En*, кумулятивная относительная частота наблюдения состояний всех объектов в определенном объеме пространства *En* и т.д.);
* групповые (кластерные) меры (мера попарного среднего, центроидная мера и т.д.).

Указанные величины определяются алгоритмическим путем на основе имеющихся темпоральных массивов дискретных данных.

Для построения феноменологических моделей описания данных в пространстве состояний *En* примем дополнительно следующие гипотезы:

1. Пусть в пространстве состояний сложной системы *En* каждой многомерной точке *М* поставлено в соответствие действительное число *W*, которое будем называть эмпирической мерой состояния объекта.
2. Величина *W* является функцией переменных состояния и образует скалярное поле, которое считаем непрерывным в пространстве состояний *En*.
3. Предположим, что в изучаемом пространстве можно задать простую непрерывную функцию многих переменных *T=T(z1, z2,...,zn)*, которую будем называть метрикой пространства состояний *En*.
4. Пусть в пространстве состояний *En* скалярные поля величин *W* и *T* однозначно связаны между собой. Если в окрестности точки М осуществляется некий процесс, то для линии процесса справедливо соотношение *dW=CldT*, где феноменологические величины *Cl*являются функциями процесса и определяются по имеющимся данным

На основе данных гипотез приходим к возможности существования разложения эмпирических мер *W* по простым функциям метрики пространства *En* *T* [5–7, 9]. Проверка справедливости таких разложений должна определяться обработкой темпоральных данных и качеством получаемых регрессионных зависимостей. В результате сделанных предположений исходная задача связана с дифференциальным уравнением в частных производных [5–7, 10], решение которого выполняется методом характеристик

. (2)

На основе решения данной задачи получены основные уравнения и соотношения для обработки массивов темпоральных данных [6, 7]. Общая схема построения модели описания данных приведена на рисунке 1.



Рисунок 1 – Схема построения модели описания данных

Таким образом, в общем виде задача сводится к построению моделей описания данных для отдельных проблемно-ориентированных баз данных, имеющих многомерную темпоральную структуру.

Методика построения моделей включает следующие этапы.

Выбираются атрибутивные показатели, характеризующие определенный класс объектов. Образуется *n-*мерное пространство переменных состояния *En* и предполагается непрерывность данного пространства, исходя из предположения существования бесконечного множества состояний для некоторой генеральной совокупности объектов.

В многомерном пространстве *En*задается метрика вида *T=T(z1,z2,...,zn)*, которая определяет модельную среду. Вводится гипотеза, что в пространстве состояний *En* существует некоторая эмпирическая мера, характеризующая процессы изменения состояний объектов *W=W(M)*. Предполагается, что данная мера образует непрерывное скалярное поле. Постулируется феноменологическая связь между величинами вида *dW=Сl∙dT*. Проводится обработка имеющихся данных на основе полученных зависимостей и проверяется адекватность и точность модели.

Основные следствия данной теории сводятся к следующему. Каждая сложная система, для пространства состояний которой существует эмпирическая мера как скалярное поле, обладает:

* характеристической функцией пространства состояний, называемой энтропией, которая является криволинейной естественной координатой пространства состояний;
* характеристической функцией пространства состояний в виде поверхности уровня, ортогональной линиям энтропии, которую можно назвать потенциалом.

Причем данные утверждения не зависят от природы систем, если можно считать, что принятые гипотезы справедливы для определенной проблемно-ориентированной темпоральной базы данных, свойственной исследуемой предметной области.

Предложенный метод был апробирован в следующих областях:

* комплексная оценка и прогнозирование развития стран, регионов и городов по ретроспективным статистическим данным [6, 7, 9, 11–16];
* построение моделей человеческого развития и ранжирование объектов по комплексу показателей [6, 7, 17];
* анализ загрязнения природной среды городов по показателям экологической безопасности [7, 16, 18];
* систематика биологических видов и создание моделей биоразнообразия [19]
* комплексная оценка состояния социальных групп и построение моделей коллективного поведения [7].

Все это позволило получить новые результаты, связанные с построением моделей данных в предметных областях. Это дает возможность развить технологии прогнозной аналитики, которые позволяют получить более точные и достоверные прогнозы развития социально-экономических систем.

**СПИСОК ЛИТЕРАТУРЫ**

1. Корпускулярные и континуальные подходы в естествознании [http://www.studfiles.ru/preview/5754244/page:2/#3](http://www.studfiles.ru/preview/5754244/page%3A2/#3) (10.09.18).
2. Mathematical modeling of collective behavior in socio-economic and life sciences / G. Naldi, L. Pareschi, G. Toskani (eds.). Berlin, Springer, 2010, 438 p. National Governors Association.
3. Econophysics and sociophysics: trends and perspectives / B.K. Chakrabarti, A. Chakraborti, A. Chatterie (eds.). Berlin: Wiley-VCH, 2006, 622 p.
4. Encyclopedia of complexity and systems science / R.A. Meyers (Editor-in-chief). Berlin, Springer, 2009, 10370 p.
5. Аверин Г.В., Константинов И.С., Звягинцева А.В. О континуальном подходе к модельному представлению данных // Вестник компьютерных и информационных технологий, №10. 2016. – С. 47–52. URL: <http://www.vkit.ru/index.php/archive-rus/541-047-052> (15.09.18).
6. Аверин Г.В. Системодинамика. Донецк: Донбасс, 2014. – 405 с.

URL: <http://dspace.bsu.edu.ru/handle/123456789/17840> (15.09.18).

1. Звягинцева А.В. Вероятностные методы комплексной оценки природно-антропогенных систем / Под науч. ред. д.т.н., проф. Г.В. Аверина. М.: Спектр, 2016. – 257 c.

URL: <http://dspace.bsu.edu.ru/handle/123456789/17837> (15.09.18).

1. База данных Федеральной службы государственной статистики. 2018. Электр. рес. URL: <http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/>publications/catalog/ doc\_1138623506156 (15.09.18).
2. Averin G.V., Zviagintseva A.V., Shevtsova M.V. and Кurtova L.N. 2017. On Representation of Discrete Information of Temporal Databases in the Continuous Form. Journal of Engineering and Applied Sciences Volume: 12. Issue: 15: 3884–3889, DOI: [10.3923/jeasci.2017.3884.3889](http://dx.doi.org/10.3923/jeasci.2017.3884.3889%22%20%5Ct%20%22_blank%22%20%5Co%20%2210.3923/jeasci.2017.3884.3889).
3. Звягинцева А.В., Аверин Г.В. Интегрирование отдельных многомерных уравнений Пфаффа, имеющих важное прикладное значение // Научные ведомости Белгородского государственного университета. Сер. Математика. Физика. 2016. № 27(248), вып. 45. – С. 102–114.
4. Аверин Г.В., Звягинцева А.В. О справедливости принципа соответственных состояний для систем различной природы // Научные ведомости Белгородского государственного университета. Сер. Экономика. Информатика №16(265), вып. 43. 2017. – С. 104–112.

URL: <http://dspace.bsu.edu.ru/handle/123456789/21056> (15.09.18).

1. Аверин Г.В., Звягинцева А.В., Швецова А.А. О подходах к предсказательному моделированию сложных систем // Научные ведомости Белгородского государственного университета. Сер. Экономика. Информатика. Том 45, №1. 2018. – С. 140–148. URL: <http://nv.bsu.edu.ru/nv/mag/detail.php?ID=460491> (15.09.18).
2. Звягинцева А.В., Иващук О.А., Пилипенко О.В. Изучение тенденций развития городов России на основе методов событийной оценки // Строительство и реконструкция, №6(74). 2017. – С. 85–94.
3. Звягинцева А.В., Константинов И.С. Модели эволюционного развития регионов на основе показателей благоустройства городских территорий // Информационные системы и технологии, №6(104). 2017. – С. 40–50.
4. Звягинцева А.В., Константинов И.С. Прогнозирование социально-экономических показателей городов на основе феноменологических моделей // Информационные системы и технологии, №1(105). 2018. – С. 5–15.
5. Звягинцева А.В. Методика событийной оценки и результаты ранжирования регионов и городов по комплексу показателей // Системный анализ и информационные технологии в науках о природе и обществе, №1(10)–2(11). 2016. − С. 157–185.
6. Averin G.V., Zviagintseva A.V., Shevtsova M.V. and Кurtova L.N. 2016. Probabilistic methods of a complex assessment of quantitative information. Research Journal of Applied Sciences, 11(7): 415–418, DOI: [10.3923/rjasci.2016.415.418](http://dx.doi.org/10.3923/rjasci.2016.415.418).
7. Константинов И.С., Звягинцева А.В. Комплексная оценка состояния урбанизированных территорий // Градостроительство и архитектура. 2018. Том 8, №1. – С. 63–71. DOI: 10.17673/Vestnik.2018.01.12.
8. Аверин Г.В. О некоторых феноменологических закономерностях биологической жизни // Системный анализ и информационные технологии в науках о природе и обществе». Донецк: ДонНТУ, №1(10)–2(11). 2016. − С. 157−186. URL: <http://sait.csm.donntu.org/> (15.09.18).

**Аверин Геннадий Викторович**

Белгородский государственный национальный исследовательский университет, г. Белгород

Д.т.н., проф., заведующий кафедрой общей математики

E-mail: averin@bsu.edu.ua

**Звягинцева Анна Викторовна**

Белгородский государственный национальный исследовательский университет, г. Белгород

К.т.н., доцент, старший научный сотрудник

E-mail: zviagintseva@bsu.edu.ua