

© А. С. БЕЗРУКОВ, М. С. ВОРОБЬЕВА

Тюменский государственный университет
mail@asbezrukov.ru, vms_08@mail.ru

УДК 519.24

СЕГМЕНТАЦИЯ НЕОДНОРОДНЫХ ОБЪЕКТОВ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ ЭНЕРГОПОТРЕБЛЕНИЯ

SEGMENTATION OF HETEROGENEOUS OBJECTS IN FORECASTING PROBLEMS OF ENERGY CONSUMPTION

В работе рассмотрены особенности задач прогнозирования энергопотребления, приведены основные методы и модели их решения. В таких задачах важна точность полученного прогноза, зависящая от качественной предварительной обработки значений исследуемых показателей. Авторами предлагается новый метод предварительной обработки данных объектов энергопотребления для задач ранжирования и сегментации при построении прогнозов энергопотребления. Рекомендуется математическая модель системы неоднородных объектов, а также подход к моделированию данной системы с использованием ассоциативных правил. Генерация ассоциативных правил производится с использованием аппарата приближенных множеств. Полученная модель позволяет получить адекватное представление сложного объекта для дальнейшей обработки, анализа, прогнозирования. Результаты работы использованы в исследовании зависимостей при построении прогнозных моделей энергопотребления, предварительной подготовке показателей для задач прогнозирования энергопотребления сложных объектов.

The peculiarities of the forecasting problem of energy consumption are considered in the paper; the main methods and solution models for forecasting problems are presented. The accuracy of the obtained forecast depending on the qualitative pre-processing of the indicators under investigation is important for the forecasting problems of energy consumption. A new method for data pre-processing of energy consumption objects is introduced for ranking and segmentation tasks in the construction of energy consumption forecasts. A mathematical model of heterogeneous objects and the approach to the modeling of heterogeneous objects using association rules are suggested. The generation of association rules is performed using the apparatus of rough sets. The resulting model allows the adequate representation of a complex object to be obtained for further processing, analysis and forecasting. The results of the study are used while investigating the dependencies in the construction of forecast models of energy consumption, preliminary preparation of indicators for forecasting problem of energy consumption of complex objects.

КЛЮЧЕВЫЕ СЛОВА. Прогнозирование энергопотребления, моделирование оценки показателей, неоднородные системы, моделирование сложных объектов.

KEY WORDS. Energy consumption forecasting, performance evaluation modeling, heterogeneous systems, modeling of complex objects.

Введение

Задаче прогнозирования энергопотребления посвящено множество работ на стыке областей экономического планирования, математического моделирования и разработки новых программных продуктов. Построение прогноза энергопотребления является приоритетной задачей для энергосбытовых компаний, являющихся посредниками между потребителем и оптовым рынком энергопотребления [1]. Важной особенностью прогнозирования в данной области является требование высокой точности прогноза, поскольку от этого напрямую зависит экономическая эффективность генерирующих мощностей [6, 10].

Согласно обзору [9], математические методы, используемые в построении прогнозов, можно разделить на две категории — статистические и методы искусственного интеллекта (ИИ). Для первой категории выработка зависимостей и точность прогноза зависит от масштаба агрегации данных, а также однородности объектов анализа. Одним из перспективных статистических методов является метод максимального подобия [9]. Для второй категории используется имитация процессов человеческого мышления и выбора, и в общем случае построение прогнозных моделей средствами ИИ сводится к использованию искусственных нейронных сетей (ИНС) различной архитектуры [8]. Методы ИИ позволяют работать с менее однородными данными при сохранении высокой точности, однако при моделировании сложных систем точность прогноза не гарантирована [3].

Важным фактором высокой точности прогноза энергопотребления является качественная предварительная подготовка входных параметров (замеров энергопотребления или энергогенерации) [5], что само по себе является наукоемкой задачей, поскольку объект энергопотребления является сложным по своей природе и обладает различными фазами потребления энергии.

В данной статье предлагается новый подход предварительной обработки значений показателей энергопотребления с использованием модели неоднородной системы, позволяющей привести сложный по структуре объект к набору однородных сравнимых показателей.

Метод предварительной обработки данных

Пусть w_i — некоторый объект энергопотребления, который можно описать набором характеристик $p_i = \{v_1, v_2, \dots, v_n\}$, $p_i \in P$, заданных в порядковой, абсолютной или номинальной шкале. Характеристиками объекта могут быть объемно-пространственные характеристики зданий, количество сотрудников, параметры энергоэффективности и другие свойства реального объекта, прямо или косвенно влияющие на энергопотребление объекта. Множество таких объектов описывает неоднородную систему W .

Пусть P является множеством всех характеристик системы. Их набор для каждого объекта может быть различным по количеству и значениям, неизменным является наличие параметра энергопотребления. Пусть X — множество общих свойств, имеющих у всех объектов, Y — множество характеристик, недоступных всем объектам, значение которых может быть выражено экспертом, а Z — множество прочих характеристик, через которые могут выражаться объекты множества Y , $P = X \cup Y \cup Z$.

Используя метод экспертной оценки [2] для извлечения сравнимых характеристик в системе неоднородных объектов, возможно получить систему (1) W_{comp} к сравнимых показателей $X \cup Y$ для m объектов, где зависимости между показателями выражены зависимостями f для каждого объекта:

$$W_{comp} = \begin{pmatrix} p_{11} & f_{12} & \cdots & f_{1k} \\ f_{21} & 0 & \cdots & f_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mk} \end{pmatrix} \quad (1)$$

Метод генерации ассоциативных зависимостей

Поскольку часть характеристик, выраженных в номинальной шкале, несет в себе дополнительную семантику предметной области, по нашему мнению, рациональным является комбинированный подход, сочетающий в себе как кластеризацию числовых характеристик объектов, так и их классификацию по ряду значимых показателей.

Для оценки состояния каждого объекта системы по всему набору характеристик необходимо выполнить сегментацию объектов по динамике значений характеристик в момент замера. Пусть $v_{ij} = \{\eta_1, \eta_2, \dots, \eta_n\}$ — значения замеров состояния системы объектов для n известных замеров. Учитывая, что количество характеристик статично и известно заранее, а начальное значение центра кластера возможно установить, используя шкалы измерений, наиболее правильным будет применение улучшенной версии алгоритма кластеризации k средних (k -means++).

Проведем сегментацию объектов системы по P_i , исключая показатели, выраженные в номинальной шкале. Начальное значение центра кластера определяется как нормативное значение по шкале измерений, согласно условиям предметной области. Для каждого измерения η определим положение объекта w_i относительно центров кластеров P_{wi} , для которых известны значения измерений объекта. Зафиксируем S_p как евклидово расстояние от положения объекта в многомерном пространстве до центра кластера P . Таким образом, в каждом измерении η объект w_i характеризуется вектором расстояний $S_{p_1}, S_{p_2}, \dots, S_{p_g} = |P_{wi}|$:

$$S_i = \begin{pmatrix} S_{11} & S_{21} & \cdots & S_{1g} \\ S_{12} & 0 & \cdots & S_{2g} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{ng} \end{pmatrix} \quad (2)$$

$$\Theta = \{S_1, S_2, \dots, S_{N_W}\}, N_W = |W| \quad (3)$$

Представим множество матриц (3) в виде транзакционной базы D , в которой каждый элемент d_b , $b = 1, 2, \dots, N_D$, $N_D = |D|$ содержит информацию о расстояниях кластеризованной системы:

$$d_b = (\phi_b, S_b),$$

где ϕ_b — идентификатор транзакции, S_b — матрица расстояний по всем измерениям некоторого объекта $w_b \in W$.

Поскольку при генерации необходимо получить логические правила базы данных D в порядке уменьшения важности признаков, рационально использование метода генерации логических правил на основе теории приближенных множеств [4]. Пусть A_b — приближенное множество транзакции d_b . Такое множество характеризуется двумя аппроксимациями — нижней, которая определяет элементы, однозначно принадлежащие данному множеству, и верхней, определяющей элементы, которые могут принадлежать данному множеству. Найдем отношение неразличимости I через функцию принадлежности:

$$\mu_{A_b}^I(d_b) = \frac{|A_b \cap I(d_b)|}{|I(d_b)|}.$$

Видим, что $0 \leq \mu_{A_b}^I(d_b) \leq 1$.

Используем функцию принадлежности для определения верхней и нижней аппроксимации приближенного множества A_b , а также граничного региона (boundary region) BN_r .

Нижней аппроксимацией является:

$$I^*(A_b) = \{d_b \in D: \mu_{A_b}^I(d_b) = 1\}.$$

Верхней аппроксимацией:

$$I^*(A_b) = \{d_b \in D: \mu_{A_b}^I(d_b) = 0\}.$$

Алгебраический метод нахождения верхней и нижней аппроксимаций приближенного множества, основанный на описании объектов и их свойств с помощью системы унарных предикатов, был предложен в работах Д. Э. Ситникова [7]. Он позволяет определять аппроксимации на заданном множестве за один просмотр данных. В результате возможно получение логических правил набора данных D в порядке уменьшения важности признаков, фигурирующих в правых частях ассоциативных правил.

Поиск ассоциативных зависимостей осуществляется на основе нахождения покрытий (covers) как набора признаков, удовлетворяющих условиям поддержки (support). Способом последовательной проверки каждого признака покрытия выбирается наиболее важный, исключение которого приводит к предельному увеличению граничного региона, после чего такой признак исключается из рассмотрения. Далее определяется несократимый набор признаков, позволяющий определить набор неразличимых элементов, и помещается в левую часть ассоциативного правила, в то время как наиболее важный признак помещается в правую его часть. В результате формируется база ассоциативных правил в порядке уменьшения важности признаков.

Полученные правила главным образом используются для получения зависимостей между характеристиками объекта, выраженными в номинальной шкале, и свойствами, представленными в виде числовых значений.

Таким образом, становится возможным определение локальных нормативов на множестве характеристик. Полученные зависимости могут быть приняты

экспертом в качестве параметров ранжирования неоднородных объектов с заданным весом, совместно со значимыми характеристиками и заданными экспертом зависимостями.

Определим функцию g_i для нахождения численного значения характеристик i -го объекта для показателя, выраженного в номинальной шкале. Его нахождение сводится в выбору максимума на массиве упорядоченных согласно признаку важности ассоциативных правил при сегментации объектов.

Неоднородная система принимает вид:

$$W_{comp} = \begin{Bmatrix} p_{11} & f_{12} & \cdots & f_{1k} & g_{1l} & \cdots & g_{1n} \\ f_{21} & 0 & \cdots & f_{2k} & g_{2l} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mk} & g_{ml} & \cdots & g_{mn} \end{Bmatrix}. \quad (4)$$

Полученная система неоднородных объектов (4) позволяет сравнивать объекты и сегментировать их по ряду общих свойств. Определим функцию зависимости между сравниваемыми объектами, используя характеристики, влияющие на энергопотребление объекта. Полученный показатель возможно использовать в качестве коэффициента мощности объекта, используемого в предварительной подготовке данных для задачи прогнозирования эффективности.

Заключение

В работе предложен новый подход к предварительной обработке данных для задачи прогнозирования потребления электроэнергии. Основные особенности данного подхода заключаются в следующем:

- сегментация неоднородных объектов позволяет проводить прогноз энергопотребления для каждого сегмента отдельно, агрегирование данных выполняется для сравнимых объектов;
- возможно использование показателей сравнения объектов при нормировании и предварительной обработке характеристик энергопотребления;
- для построения прогнозов, в том числе по объектам, фактические показатели энергопотребления которых отсутствуют (новые и планируемые к постройке объекты), возможно использовать среднее значение по объектам внутри сегмента;
- представленные в статье результаты использованы в работах по исследованию зависимостей при построении прогнозных моделей энергопотребления в рамках реализации прототипа подсистемы анализа энергопотребления спортивных объектов.

СПИСОК ЛИТЕРАТУРЫ

1. Аюев Б. И. Рынки электроэнергии и их реализация в ЕЭС России / Б. И. Аюев. Екатеринбург: УРО РАН, 2007. 107 с.
2. Безруков А. С. Моделирование процесса оценки показателей однородных и неоднородных объектов / А. С. Безруков, М. С. Воробьева // Вестник Тюменского Государственного Университета. 2014. № 7. С. 218-225.
3. Безруков А. С. Применение искусственных нейронных сетей для прогнозирования динамики показателей объекта управления на примере спортивных соору-

- жений / А. С. Безруков // Математическое и информационное моделирование: сборник научных трудов. Тюмень: Издательство Тюменского государственного университета. 2013. Вып. 13. С. 49-54.
4. Болдырев М. В. Решение задач с применением нечеткой логики / М. В. Болдырев // Энергосбережение, автоматизация в промышленности, интеллектуальные здания и АСУТП. 2010. Вып. 5. С. 5-7.
 5. Демьяненко Т. С. Прогнозирование потребления электрической энергии на оптовом рынке электроэнергии и мощности / Т. С. Демьяненко, В. Г. Мохов // Вестник Южно-Уральского государственного университета. Серия: Экономика и менеджмент. 2014. Т. 8. № 2.
 6. Кирпичникова И. М. Прогнозирование объемов потребления электроэнергии / И. М. Кирпичникова, Л. А. Саплин, К. Л. Соломахо // Вестник ЮУрГУ. Серия: Энергетика. 2014. № 2. С. 16-25.
 7. Ситников Д. Э. Метод поиска логических правил в экспертных системах на основе ассоциативных зависимостей между дискретными признаками / Д. Э. Ситников // Збірник наукових праць Харківського університету Повітряних сил. 2014. № 2. С. 147-150.
 8. Талалаев А. А. Анализ эффективности применения искусственных нейронных сетей для решения задач распознавания, сжатия и прогнозирования / А. А. Талалаев // Искусственный интеллект и принятие решений. 2008. Т. 2. С. 24-33.
 9. Чучуева, И. А. Модель экстраполяции временных рядов по выборке максимального подобия / И. А. Чучуева // Информационные технологии. 2010. № 12. С. 43-47.
 10. Щербаков М. В. Специфика применения интеллектуальных моделей анализа данных для повышения энергетической эффективности / М. В. Щербаков // Известия Волгоградского государственного технического университета. 2010. Вып. 9. № 11 (71). С. 72-76.

REFERENCES

1. Ajuev B. Rynki jelektroenergii i ih realizacija v EJeS Rossii [Electricity Markets and Their Implementation in the UES of Russia]. Ekaterinburg: URO RAN, 2007. 107 p. (In Russian)
2. Bezrukov A. S., Vorobyeva M. S. Modelirovanie processa ocenki pokazatelej odnorodnyh i neodnorodnyh obyektov [Simulation of the Evaluation Process of Indicators of Homogeneous and Heterogeneous Objects] // Vestnik Tjumenskogo Gosudarstvennogo Universiteta [Tyumen State University Herald]. 2014. No 7. Pp. 218-225. (In Russian)
3. Bezrukov A. S. Primenenie iskusstvennyh nejronnyh setej dlja prognozirovanija dinamiki pokazatelej objekta upravlenija na primere sportivnyh sooruzhenij [Application of Artificial Neural Networks to Predict the Dynamics of the Control Object on the Example of Sports Facilities] // Matematicheskoe i informacionnoe modelirovanie: sbornik nauchnyh trudov [Mathematical and Information Modeling: Collection of Scientific Papers]. Tyumen: Izdatel'stvo Tjumenskogo gosudarstvennogo universiteta [Tyumen State University Publishing House]. 2013. Issue 13. Pp. 49-54. (In Russian)
4. Boldyrev M. V. Reshenie zadach s primeneniem nechetkoj logiki [Solving problems with the use of fuzzy logic] // Jenergoberezhenie, avtomatizacija v promyshlennosti, intellektual'nye zdanija i ASUTP [Energy-saving, Industrial Automation, Process Control Systems and Intelligent Building]. 2010. Vol. 5. Pp. 5-7. (In Russian)

5. Dem'janenko T. S., Mohov V. G. Prognozirovanie potreblenija jelektricheskoy jenerгии na optovom rynke jelektrojenerгии i moshhnosti [Predicting Consumption of Electric Power in the Wholesale Electricity Market] // Vestnik Juzhno-Ural'skogo gosudarstvennogo universiteta. Serija: Jekonomika i menedzhment [South Ural State University Herald. Economics and Management]. 2014. Vol. 8. No 2. (In Russian)
6. Kirpichnikova I. M., Saplin L. A., Solomaho K. L. Prognozirovanie obyomov potreblenija jelektrojenerгии [Energy Consumption Forecasting] // Vestnik JuUrGU. Serija: Jenergetika [South Ural State University Herald. Power Engineering]. 2014. No 2. Pp. 16-25. (In Russian)
7. Sitnikov D. E. Metod poiska logicheskikh pravil v jekspertnyh sistemah na osnove asociativnyh zavisimostej mezhdu diskretnymi priznakami [Search Method of Logical Rules in Expert Systems Based on Associative Relationships between Discrete Signs] // Zbirnik naukovih prac' Harkivs'kogo universitetu Povitrjanih sil [Proceedings of Kharkiv Air Force University]. 2014. No 2. Pp. 147-150. (In Russian)
8. Talalaev A. A. Analiz jeffektivnosti primenenija iskusstvennyh nejronnyh setej dlja reshenija zadach raspoznavanija, szhatija i prognozirovanija [An analysis of the Effectiveness of Artificial Neural Networks to Solve the Problems of Recognition, Compression and Prediction] // Iskusstvennyj intellekt i prinjatje reshenij [Artificial Intelligence and Decision Making]. 2008. Vol. 2. Pp. 24-33. (In Russian)
9. Chuchueva I. A. Model' jekstrapoljaccii vremennyh rjadov po vyborke maksimal'nogo podobija [Time Series Extrapolation Model Based on Maximum Likeness Set] // Informacionnye tehnologii [Information Technologies]. 2010. No 12. Pp. 43-47. (In Russian)
10. Shherbakov M. V. Specifika primenenija intellektual'nyh modelej analiza dannyh dlja povyshenija jenergeticheskoy jeffektivnosti [Specific of Specificity of Application of Intelligent Data Analysis Models for Energy Efficiency] // Izvestija Volgogradskogo gosudarstvennogo tehničeskogo universiteta. Aktual'nye problemy upravlenija, vychislitel'noj tehniki i informatiki v tehničeskikh sistemah [Proceedings of Volgograd State Technical University. Actual Problems of Management, Computer Science and Informatics in Technical Systems]. 2010. No 11 (71). Issue 9. Pp. 72-76. (In Russian)

Авторы публикации

Безруков Андрей Сергеевич — старший преподаватель кафедры программного обеспечения Института математики и компьютерных наук Тюменского государственного университета

Воробьева Марина Сергеевна — кандидат технических наук, доцент кафедры программного обеспечения Института математики и компьютерных наук Тюменского государственного университета

Authors of the publication

Andrey S. Bezrukov — Senior Lecturer at the Department of Software, Institute of Mathematics and Computer Sciences, Tyumen State University

Marina S. Vorobyeva — Cand. Sci. (Engin.), Associate Professor at the Department of Software, Institute of Mathematics and Computer Sciences, Tyumen State University