

УДК 004.827

Л.В. Галицкая, А.А. Голубева, Ю.Б. Гриценко, О.И. Жуковский, И.Ю. Пудуль

## Идентификация параметров объектов электронного генерального плана металлургического предприятия

Приведено описание метода решения задачи идентификации параметров объектов электронного генерального плана металлургического предприятия на основе теории нечеткого моделирования с использованием методов наименьших квадратов, роящихся частиц и гибридного метода.

**Ключевые слова:** электронный генеральный план, геоинформационная система, нечеткое моделирование, идентификация.

**Постановка задачи.** Современные промышленные металлургические предприятия имеют в своем составе большое количество разнообразных производств: доменное, мартеновское, коксохимическое, литейное и др., которые в свою очередь занимают большое территориальное пространство, плотно насыщенное разнообразными производственными объектами. В настоящий момент времени на предприятиях используют такой подход к автоматизации, при котором основными задачами, требующими внедрения компьютерных технологий, являются те, эффект от успешного решения которых оценивается наиболее явным образом. При использовании данного подхода управление производственной инфраструктурой предприятия является одним из основных автоматизируемых процессов.

Решение задачи управления производственной инфраструктурой металлургического предприятия во многом зависит от эффективности разработки и использования генерального плана. В этом случае применение технологии ГИС – закономерный этап на пути перехода к безбумажной технологии обработки информации. Географическая информационная система (ГИС) – это современная компьютерная технология для картирования и анализа объектов реального мира и событий, происходящих с этими объектами [1].

В состав этой технологии входят традиционные операции работы с базами данных, такие как запрос и статистический анализ, с преимуществами полноценной визуализации и пространственного анализа, которые предоставляет карта [2].

На сегодняшний день применение ГИС-технологий, лежащих в основе комплексных информационных систем, все чаще ограничивается решением задач информационного сопровождения, а их роль в решении задач анализа недооценивается. Причиной подобного обстоятельства является либо наличие в современных ГИС слаборазвитого аппарата анализа данных, либо полное его отсутствие. Такое положение дел требует совершенствования подходов к созданию информационных систем на базе ГИС с учетом реализации функций анализа данных. Одной из ключевых задач анализа данных на металлургическом предприятии является задача идентификации (восстановления) параметров объектов электронного генерального плана (ЭГП). Решение задачи идентификации позволит узнать причину, свойства взаимосвязей объектов, а также место нахождения связанных объектов. Получение этой информации может помочь более глубоко понять ситуацию в данном месте, сделать правильный выбор или лучше подготовиться к предстоящим событиям и условиям.

**Описание метода решения.** Для решения задачи идентификации параметров объектов ЭГП авторы статьи предлагают использовать теорию нечеткого моделирования. Использование данной теории обусловлено тенденцией увеличения сложности математических и формальных моделей реальных систем и процессов управления, связанной с желанием повысить их адекватность и учесть все большее число различных факторов, оказывающих влияние на процессы принятия решений [3]. Авторами статьи разработан обобщенный алгоритм решения задачи идентификации (восстановления) пространственно-временных данных ЭГП металлургического предприятия, представленный на рис. 1. Данный алгоритм был реализован в программном обеспечении интегрированным в веб-ориентированную геоинформационную технологию формирования и мониторинга электронного генерального плана инженерной инфраструктуры [4] в качестве модуля анализа данных.

Важным этапом диагностики параметров объектов ЭГП является сбор априорной информации об их характеристиках. Важно помнить, что объекты ЭГП весьма разнообразны. Например, в определенном производстве металлургического предприятия могут пролегать самые разнообразные типы инженерных коммуникаций (водопроводные, канализационные сети, паропроводы, продуктопроводы и др.), и их параметры (толщина труб, материал, глубина залегания и др.) могут существенно отличаться. Проводя обучение системы (реализованной в модуле анализа данных) на существующих данных, нужно помнить о том, что применять задачу идентификации можно только к такому же типу данных. Другими словами, информация должна принадлежать одной генеральной совокупности.

После того как данные были собраны, можно приступить к процессу обучения системы. На первом этапе необходимо определить входные и выходные параметры системы, которыми являются характеристики объектов, и проинициализировать их, задав соответствующее количество термов и выбрав подходящую функцию принадлежности.

На данном этапе развития исследований в области нечеткого моделирования практически невозможно заранее точно сказать, как повлияют на результат вычислений выбор функций принадлежности и количество термов. В силу этого необходимо проводить эксперименты по инициализации переменных и наблюдать за результатами вычислений. Выбрав удовлетворяющую точность вычислений, определить, какая функция принадлежности наиболее пригодна и показывает адекватные результаты.

Функция принадлежности может быть представлена одним из следующих типов:

- треугольный – рационально использовать при необходимости быстрой классификации и отсутствии требований высокого качества (необходим экспресс-анализ системы), т.е. пользователю важна в первую очередь скорость обработки данных;

- гауссовское распределение (функция имеет гладкий вид) – применение обусловлено в случае, если важна устойчивость распределения и есть необходимость уточнения данных экспресс-анализа. Функция Гаусса является оптимальной и дает практически приемлемые результаты по скорости обучения, по скорости классификации и по эффективности классификации;

- трапеция (функция имеет линейный и местами ступенчатый вид) – использование трапециевидной функции обусловлено наличием специфической выборки (специфичность выборки определяется тем, что анализируемые значения выходят за рамки обучающей выборки) и отсутствием жестких ограничений по времени;

- парабола – параболическая функция принадлежности является наилучшей по скорости обучения. Эффективность классификации также находится на высоком уровне [5].

После инициализации переменных следует перейти ко второму этапу обучения системы, настройке антецедентов и консеквентов правил. Для настройки базы правил применяется ряд методов: метод наименьших квадратов (МНК), метод роящихся частиц (РЧ), гибридный алгоритм (настройка методом роящихся частиц, последующая подстройка методом наименьших квадратов).

При выборе метода настройки, также как и при выборе функции принадлежности, на данном этапе развития исследований невозможно с уверенностью сказать, какой метод будет наиболее адекватен. Но следует учесть тот факт, что метод РЧ – наиболее часто успешно применяемая эвристика для обучения нечетких систем, так как данный метод защищен от «застывания» в локальных решениях.



Рис. 1. Обобщенный алгоритм решения задачи идентификации пространственно-временных данных ЭГП металлургического предприятия

МНК является классическим алгоритмом оптимизации и находит оптимальное решение за одну итерацию, поэтому в совокупности он менее затратный по ресурсам, чем аналоги при достижении той же точности. Рекомендацией применения МНК является необходимость получить хорошее и быстрое решение.

Рекомендацией применения гибридного алгоритма является желание улучшить полученное решение, но данный метод требует больших временных затрат.

Одним из заключительных этапов процесса настройки обучения нечеткой системы является выбор типа нечеткой системы, так называемого алгоритма вывода. Для выбора типа нечеткой системы следует произвести предварительный статистический анализ (рассчитать коэффициент регрессии) [6]:

1) если коэффициент регрессии более 0,7, необходимо использовать систему нечеткого вывода типа Такаги–Сугено;

2) если коэффициент регрессии не превышает 0,2, то используем систему нечеткого вывода типа Мамдани;

3) если коэффициент от 0,2 до 0,7, то используем систему нечеткого вывода типа Сингтон.

После того как подобраны все необходимые параметры и методы, произведена настройка системы, можно запустить процесс восстановления данных.

**Проведение эксперимента.** Для выполнения эксперимента по аппроксимации данных смоделируем ситуацию отсутствия данных по объектам слоя «продуктопровод», для чего из таблицы наблюдений, состоящей из 104 записей и 4 полей, удалим девять значений атрибута «толщина стенки» в записях, представленных в табл. 1.

Таблица 1  
Перечень записей в экспериментальной таблице, для которых необходимо удалить значение атрибута «Толщина стенки»

Object Id	Диаметр, мм	Толщина стенки, мм	Рабочая температура, °С
119465	127,5	3,5	14
119980	110	3,2	13
119495	125	3,5	13
118025	150	2,9	45
79113	192,5	2,8	67
119267	125	2,3	33
121518	162,5	3	52
121717	127,5	3,5	14
119045	155	2,9	43

(диаметр и рабочая температура). Для настройки параметров нечеткой модели (антецедентов правил) будем использовать метод роящихся частиц. Для осуществления настройки параметров нечеткой модели (консеквентов правил) применим метод наименьших квадратов (МНК).

Далее непосредственно запустим процедуру восстановления экспериментальных данных. Результат работы процедуры восстановления данных отображен в табл. 2. Восстановлены отсутствующие значения у объектов «продуктопроводы» в свойстве «толщина стенки». Среднеквадратичное отклонение составило 0,016.

Таблица 2  
Перечень записей в экспериментальной таблице, для которых восстановили значение атрибута «Толщина стенки»

Object Id	Диаметр, мм	Толщина стенки, мм	Рабочая температура, °С
119465	127,5	3,500	14
119980	110	3,229	13
119495	125	3,509	13
118025	150	2,910	45
79113	192,5	2,791	67
119267	125	2,312	33
121518	162,5	2,990	52
121717	127,5	3,500	14
119045	155	2,893	43

Далее в модуле анализа данных осуществим задание входных и выходных параметров. Входными параметрами будут являться «диаметр» и «рабочая температура», а параметр «толщина стенки» обозначим как выходной. Выполним инициализацию и обучение на основе рассчитанного коэффициента корреляции, превышающего 0,7. В этом случае система рекомендует выбрать тип нечеткой системы Такаги–Сугено из реализованных в системе. Для осуществления инициализации переменных выберем функцию принадлежности «треугольная» и зададим по 5 термов для каждого входного параметра

(диаметр и рабочая температура). Для настройки параметров нечеткой модели (антецедентов правил) будем использовать метод роящихся частиц. Для осуществления настройки параметров нечеткой модели (консеквентов правил) применим метод наименьших квадратов (МНК).

Далее непосредственно запустим процедуру восстановления экспериментальных данных. Результат работы процедуры восстановления данных отображен в табл. 2. Восстановлены отсутствующие значения у объектов «продуктопроводы» в свойстве «толщина стенки». Среднеквадратичное отклонение составило 0,016.

**Заключение.** Разработанный авторами метод решения задачи идентификации параметров объектов ЭПП металлургического предприятия и реализованный в модуле анализа данных интегрированного в веб-ориентированную геоинформационную технологию формирования и мониторинга электронной генеральной плана инженерной инфраструктуры показал адекватный результат по восстановлению данных и может быть использован для решения задач управления производственной инфраструктурой предприятия.

Выполнение данной работы проводилось при финансовой поддержке Министерства образования и науки Российской Федерации в рамках мероприя-

тия 2.4 федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы», проект «Разработка Web-ориентированных геоинформационных технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры», государственный контракт № 07.524.11.4013 от 03 ноября 2011 г.

#### *Литература*

1. Ехлаков Ю.П. Архитектура веб-ГИС-сервера для ведения электронного генерального плана инженерной инфраструктуры / Ю.П. Ехлаков, О.И. Жуковский, Н.Б. Рыбалов // Известия Томского политехнического университета. – 2006. – Т. 309, № 7. – С. 146–152.
2. ДеМерс Майкл Н. Географические информационные системы. Основы. – М.: Дата+, 1999. – 490 с.
3. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – СПб.: БХВ Петербург, 2005. – 736 с.
4. Гриценко Ю.Б. Архитектура веб-ГИС-сервера для ведения электронного генерального плана инженерной инфраструктуры / Ю.Б. Гриценко, О.И. Жуковский, М.М. Милихин // Качество. Инновации. Образование. – 2012. – № 12. – С. 97–101.
5. Yuen Fuzzy linguistic variable matrix and parabola-based fuzzy normal distribution / K.K.F. Yuen, H.C.W. Lau // Intelligent Information Processing III. IFIP TC12 International Conference on Intelligent Information Processing (IIP 2006), September 20–23, Adelaide, Australia. – P. 205–215.
6. Голубева А.А. Интеллектуальный анализ пространственно-временных данных объектов инженерной инфраструктуры // Доклады ТУСУРа. – 2012. – № 2 (26), ч. 2. – С. 216–219.

---

#### **Галицкая Любовь Владимировна**

Мл. науч. сотрудник каф. АОИ ТУСУРа  
Тел.: (382-2) 90-01-80  
Эл. почта: lyb\_vlad@ngs.ru

#### **Голубева Александра Александровна**

Аспирант каф. АОИ ТУСУРа  
Тел.: 8-913-825-84-26  
Эл. почта: Sasha\_karateka@mail.ru

#### **Гриценко Юрий Борисович**

Канд. техн. наук, доцент каф. АОИ ТУСУРа  
Тел.: (382-2) 90-01-80  
Эл. почта: ubg@muma.tusur.ru

#### **Жуковский Олег Игоревич**

Канд. техн. наук, доцент каф. АОИ ТУСУРа  
Тел.: (382-2) 90-01-80  
Эл. почта: ol@muma.tusur.ru

#### **Пудуль Игорь Юрьевич**

Студент каф. АОИ ТУСУРа  
Тел.: (382-2) 90-01-80  
Эл. почта: svigani4ok@gmail.com

Galitskaya L.V., Golubeva A.A., Gritsehko Yu.B., Zhukovskiy O.I., Pudul I.Yu.

#### **Identification of the parameters of the objects in electronic master plan for metallurgical enterprise**

The paper describes the method for solving the parameter identification of objects in an electronic master plan for the metallurgical enterprise based on the theory of fuzzy modeling using the method of Ordinary Least Squares, swarming particles and hybrid method.

**Keywords:** electronic master plan, geographical information system, fuzzy modeling, identification.