

вития. Научно-обоснованные рекомендации с оценкой эффективности и последствий предлагаемых мероприятий по своевременной перестройке необходимо обосновывать с помощью методов системного анализа и комплексных системных моделей, адекватных реальным процессам, происходящим в экономике страны и в отрасли.

Результаты такого комплексного и системного анализа и прогнозирования позволят выявить «узкие места» в отрасли, что будет способствовать повышению отраслевой инвестиционной привлекательности, а это будет положительным фактором, влияющим на социально-экономическое положение страны в целом. В то же время применение предложенной методики в процессе формирования стратегии инвестирования позволит повысить ее эффективность, так как предотвратит вкладывание средств в отрасль, находящуюся на сегодняшний момент «на плаву», но тенденция развития которой регрессивная.

ЛИТЕРАТУРА

1. *Изыгзон Н. Б.* Методология формирования и реализации инвестиционной политики угольной промышленности России. — М.: Статистика, 2000.
2. *Бойко В. И.* Применение многокритериального подхода при оценке перспектив освоения минеральных ресурсов Арктической зоны // Сб. тр. ВНИИ системных исследований. — 1991. — № 5.
3. *Гранин И. В.* Мировая стратегия формирования угольного производства // Научные сообщения НИЦ горного производства. — 1999. — № 311.

e-mail: docent-ostu@yandex.ru

Статья представлена к публикации членом редколлегии
Р. М. Нижнегородцевым. □

УДК 681.325...2

ИДЕНТИФИКАЦИЯ ЭКОЛОГИЧЕСКИХ СИТУАЦИЙ В ПРОЦЕССЕ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА

Г. С. Вересников

Институт проблем управления им. В. А. Трапезникова РАН, г. Москва

Предложен подход к построению модуля идентификации экологической ситуации для поддержки принятия оперативных решений. Разработана методика формирования репрезентативной выборки, необходимой для его создания.

ВВЕДЕНИЕ

В процессе экологического мониторинга потенциально опасных объектов важно своевременно идентифицировать сложившуюся экологическую ситуацию и применять адекватные меры для ее разрешения.

Закономерности, используемые для сопоставления оценки экологической ситуации текущим измерениям экологических параметров обычно трудно формализуемые. Для поиска и построения этих закономерностей предлагается применять методы, являющиеся частью технологий Data Mining [1]. Они позволяют сформировать логические правила, основываясь на эмпирической выборке вида <экологические параметры; оценка эколо-

гической ситуации>. Основная проблема заключается в том, что необходимо составить репрезентативную выборку, которая будет охватывать все предусмотренные экологические ситуации. Системы экологического мониторинга характеризуются недостаточным количеством накопленной информации о нештатных и аварийных ситуациях, что объясняется их малой вероятностью. В связи с этим необходимы методы получения дополнительной информации о тех ситуациях, которые могут возникнуть в процессе экологического мониторинга. Для решения данной проблемы предлагается моделировать нештатные ситуации, измерения по которым отсутствуют.

Другой аспект целесообразности применения технологий Data Mining для обработки данных экологичес-



кого мониторинга состоит в необходимости принятия решений в условиях нечеткой исходной информации. Нечеткость информации может быть связана как с недостаточным числом производимых измерений (например, при отказе измерительной аппаратуры, каналов связи), так и с задержкой получения необходимой информации.

1. ЗАДАЧА ИДЕНТИФИКАЦИИ ЭКОЛОГИЧЕСКОЙ СИТУАЦИИ

Сформулируем задачу идентификации экологической ситуации. В каждый момент времени t для каждой точки контроля существует вектор экологических и метеорологических параметров $X_t = (x_{t,1}, x_{t,2}, \dots, x_{t,n})$, который формируется из значений, поступающих от измерительных приборов, установленных в точках контроля.

Обозначим через $Y_t = (y_{t,1}, y_{t,2}, \dots, y_{t,m})$ вектор, с помощью которого дается оценка экологической ситуации в момент времени t . Элементы вектора Y_t представляют собой качественную оценку экологической ситуации по заранее определенной шкале, в которой кодируются:

- тип нештатной или аварийной ситуации;
- место возникновения нештатной или аварийной ситуации;
- уровень опасности;
- тип опасности и др.

Задача идентификации экологической ситуации заключается в том, чтобы на основе анализа информации, полученной в процессе экологического мониторинга, определить вектор Y_t , характеризующий сложившуюся экологическую ситуацию.

Предлагается следующий подход к идентификации экологической ситуации. В каждой точке контроля на основе текущего вектора X_t , определяется вектор Y_t . Затем анализируются результаты идентификации, полученные во всех точках контроля, и делается вывод о текущей экологической ситуации (делается окончательный вывод о векторе Y_t).

Согласно этому подходу, для каждой точки контроля строится функция идентификации экологической ситуации, которая является отображением $\varphi: X \rightarrow Y$, множества входных векторов $X \in R^n$ в множество выходных векторов $Y \in R^m$.

Для построения отображения φ предлагается использовать алгоритмы, для работы которых требуется набор эмпирических данных, в достаточной мере отражающий зависимости между входными и выходными переменными. Такие алгоритмы реализуются, например, с помощью нейронных сетей, деревьев решений, метода группового учета аргументов (МГУА). Для этого необходимо сформировать выборку, состоящую из пар вида $\langle X; Y \rangle$. Компоненты вектора X используются в качестве входа создаваемых формальных конструкций, а компоненты вектора Y в качестве выхода.

Однако неэффективно строить отображение φ в виде одной функции, поскольку число экологических ситуаций может быть большим, а указанные алгоритмы дают

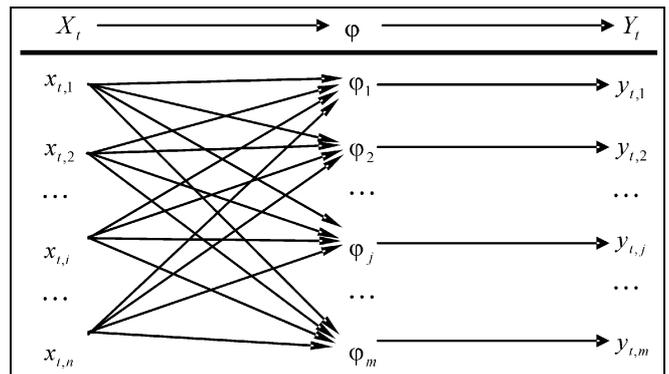


Рис. 1. Схема реализации отображения φ с помощью отображений $\varphi_1, \varphi_2, \dots, \varphi_m$

наилучший результат, когда число выходных состояний мало. Поэтому для каждой компоненты вектора Y целесообразно строить отдельное отображение из множества входных векторов X , т. е. отображение φ , реализовать с помощью набора отображений $\varphi_1, \varphi_2, \dots, \varphi_m$ (рис. 1).

Основная проблема при построении отображения φ состоит в необходимости получения обучающей выборки пар $\langle X; Y \rangle$, охватывающей различные экологические ситуации.

2. ПОДГОТОВКА ОБУЧАЮЩЕЙ ВЫБОРКИ. ПРОБЛЕМА РЕПРЕЗЕНТАТИВНОСТИ

Построение отображения φ предполагает, что для каждой экологической ситуации, задаваемой вектором Y , существует набор векторов X , который будет ей сопоставлен.

В базе данных, предназначенной для хранения экологической информации, накапливаются большие объемы данных в виде пар $\langle X_t; Y_t \rangle$. Например, в базу данных системы экологического мониторинга объекта хранения и уничтожения химического оружия по каждой точке контроля вносится более 20 000 записей в год. Однако имеющиеся пары $\langle X_t; Y_t \rangle$ в основном характеризуют штатный режим работы объекта. Большинство нештатных ситуаций либо вообще не возникает в силу их малой вероятности, либо очень редки, поэтому измеренной информации недостаточно для построения репрезентативной выборки.

Методики получения отсутствующей информации о возможных экологических ситуациях благодаря опыту и интуиции высококвалифицированных специалистов-экспертов в данном случае не дают нужного результата. Задача сопоставления наборов экологических и метеорологических параметров конкретным экологическим ситуациям для каждой точки контроля будет не только трудоемкой, но может оказаться невыполнимой в силу невозможности учета экспертами большого числа факторов.

В качестве альтернативного способа получения недостающей информации для формирования обучающей

выборки предлагается моделирование предусмотренных нештатных ситуаций на объекте экологического мониторинга.

В зависимости от типа источника выброса используются различные математические модели и методики расчета распространения вредных веществ в атмосфере. Официально утверждена «Методика расчета концентраций в атмосферном воздухе вредных веществ, содержащихся в выбросах предприятий ОНД-86», позволяющая рассчитывать распределения концентраций в приземном слое атмосферы в зависимости от условий окружающей среды [2]. Данная методика предназначена для расчета приземных концентраций в двухметровом слое над поверхностью земли, а также вертикального распределения концентраций от точечных, линейно и пространственно распределенных источников. Она предназначена для расчета значений концентраций при постоянном источнике выброса без учета распространения вредных веществ во времени, и в этом ее недостаток. Для расчета динамики распространения вредных веществ в атмосфере используются трехмерные модели переноса и диффузии примеси и их упрощенные варианты [3–5], основанные на решении начально-краевой задачи для трехмерного уравнения параболического типа.

Целесообразно использовать модели расчета распространения вредных веществ и информации об источнике загрязнения, чтобы сформировать обучающую выборку (для каждой точки контроля), необходимую при построении функции для идентификации экологической ситуации.

Опишем процесс построения обучающей выборки для точки контроля с координатами (x_k, y_k) . Он состоит из четырех этапов.

Этап 1. Формирование перечня нештатных ситуаций с указанием диапазонов выбрасываемых вредных веществ.

Нештатная ситуация описывается:

- расположением источника выброса;
- условиями выброса, задаваемыми технологическими особенностями источника;
- набором вредных веществ, выбрасываемых в атмосферу;
- диапазонами возможных изменений уровня выброса каждого вредного вещества.

Когда объект мониторинга представляет серьезную экологическую опасность, создается документ, описывающий потенциальные нештатные и аварийные ситуации. В нем для каждой такой ситуации указывается перечень выбрасываемых вредных веществ, а для каждого вещества возможные минимальный и максимальный уровни выбросов в атмосферу.

Этап 2. Формирование возможных вариантов экологических ситуаций на основе варьирования уровней выброса и параметров окружающей среды в заданных диапазонах.

Конкретный вариант нештатной ситуации задается комбинацией параметров: $\langle(x_b, y_b); S; U\rangle$, где x_b и y_b — координаты источника выброса; $S = (s_1, s_2, \dots, s_g)$ — вектор, элементами которого являются уровни выброса вредных веществ; U — вектор значений, влияющих на

распространение вредных веществ и определяемых технологическими особенностями источника.

На распространение вредных веществ в атмосфере также влияют условия окружающей среды. Обозначим через E вектор параметров окружающей среды, необходимых для расчета концентраций вредных веществ в точке контроля с координатами $(x_k; y_k)$. Часть из них изменяется во времени и входит в состав параметров, измеряемых в процессе экологического мониторинга. Обозначим ее через $E_{изм}$. Другая часть параметров постоянная и характеризует местность, где распространяются вредные вещества. Обозначим ее через $E_{пост}$.

Варьируя элементы (s_1, s_2, \dots, s_g) в заданных диапазонах и параметры вектора $E_{изм}$ в допустимых пределах, получаем различные варианты нештатной ситуации $\langle(x_b, y_b); S; U\rangle$, возникающие в различных условиях E , т. е. формируем варианты экологических ситуаций, учитывая условие $E: \langle(x_b, y_b); S; U; E\rangle$.

Этап 3. Расчет концентраций в точках контроля для каждого варианта экологической ситуации.

Для каждого сформированного варианта экологической ситуации по значениям (x_b, y_b) , (x_k, y_k) , S , U , E рассчитываются концентрации в точке контроля по всем вредным веществам. Затем рассчитываются концентрации вредных веществ в других точках контроля, и оценивается тип и уровень опасности каждого варианта экологической ситуации.

Этап 4. Формирование пар вида $\langle X; Y\rangle$.

Формируются векторы X , состоящие из рассчитанных концентраций вредных веществ и вектора $E_{изм}$. Векторы X образуются, исходя из конкретной экологической ситуации и места ее возникновения, поэтому эти качественные оценки экологической ситуации включаются в вектор Y .

В результате каждому сформированному вектору X сопоставляется вектор Y , т. е. образуется репрезентативная выборка, состоящая из пар вида $\langle X, Y\rangle$.

Состав параметров, входящих в векторы U и E , зависит от конкретной методики расчета распространения вредных веществ. Рассмотрим случай применения методики [2].

В состав вектора U входят:

H , м — высота источника выброса над уровнем земли (для наземных источников при расчетах принимается $H = 2$ м);

F — безразмерный коэффициент, учитывающий скорость оседания вредных веществ в атмосферном воздухе (зависит от коэффициента очистки выбросов);

T_r , °С — температура выбрасываемой газовой смеси;

D , м — диаметр устья источника выброса;

w_0 , м/с — средняя скорость выхода газовой смеси из устья источника выброса.

В состав вектора $E_{пост}$ входят:

A — безразмерный коэффициент, зависящий от температурной стратификации атмосферы;

η — безразмерный коэффициент, учитывающий влияние рельефа местности (в случае ровной или слабопе-

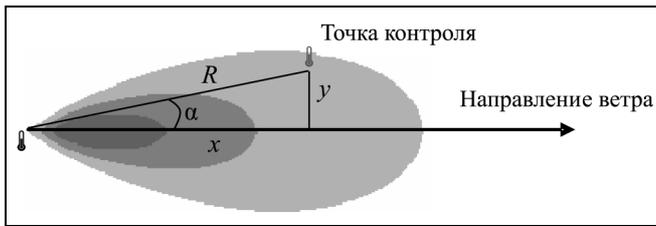


Рис. 2. Факел распространения вредных веществ

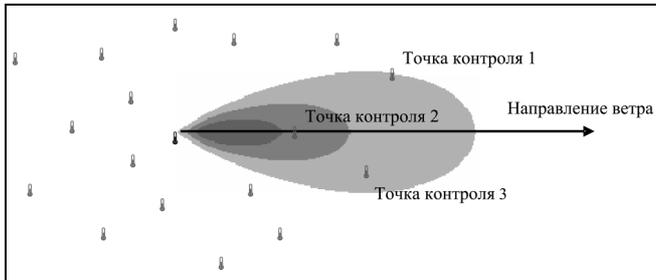


Рис. 3. Факел распространения вредных веществ

ресеченной местности с перепадом высот, не превышающем 50 м/км, $\eta = 1$).

В состав вектора $E_{\text{изм}}$ входят:

T_r , °C — температура окружающего атмосферного воздуха;

ϑ , м/с — скорость ветра;

x , м — расстояние от точки выброса до точки контроля по оси факела;

y , м — расстояние от точки выброса до точки контроля по перпендикуляру от оси факела.

При расчете концентраций вредных веществ вместо значений координат x и y используются значения $R\sin\alpha$ и $R\cos\alpha$, где R — расстояние от точки выброса до точки контроля, α — угол между вектором направления ветра и вектором, соединяющим точку выброса и точку контроля (рис. 2). Значения R и α вычисляются по известным значениям (x_k, y_k) , (x_b, y_b) и направления ветра.

Полученная обучающая выборка используется для построения отображений $\varphi_1, \varphi_2, \dots, \varphi_m$ путем обучения нейронных сетей, построения деревьев решений и уравнений МГУА.

Функции φ строятся для всех точек контроля и в дальнейшем используются, чтобы в каждой из них получить вектор Y_r , характеризующий сложившуюся экологическую ситуацию. Окончательно вектор Y_t выбирается на основе анализа результатов идентификации, полученных во всех точках контроля. Способ проведения анализа зависит от типа источника выбросов. Рассмотрим пример для точечного непрерывного источника выброса, когда диаграмма распространения вредных веществ имеет форму факела (рис. 3).

Во всех точках контроля, представленных на рис. 1, определяется вектор Y_r . Пусть в точках контроля 1–3 получен результат идентификации, отличающийся от

результатов в других точках, и стали известны компоненты вектора Y_r — тип и место возникновения нештатной или аварийной ситуации. На основе этой информации строятся границы факела распространения вредных веществ. Так как внутри границ факела попадают точки контроля 1–3, то функции φ для этих точек дают одинаковый результат идентификации экологической ситуации (одинаковый вектор Y_r). Следовательно, делается вывод о том, что вектор Y_r , полученный в точках контроля 1–3, является результатом идентификации сложившейся экологической ситуации. На основе текущего вектора $Y_t = (y_{t,1}, y_{t,2}, \dots, y_{t,m})$ принимается решение о принятии мер по разрешению идентифицированной экологической ситуации.

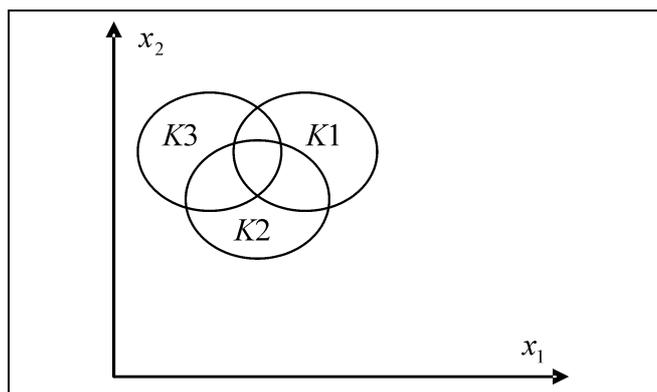
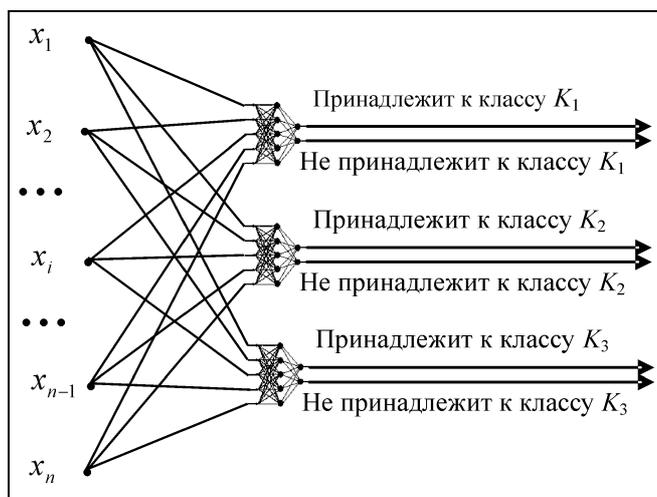
3. РЕШЕНИЕ ПРОБЛЕМЫ ПЕРЕСЕЧЕНИЯ КЛАССОВ В ЗАДАЧЕ ИДЕНТИФИКАЦИИ ЭКОЛОГИЧЕСКИХ СИТУАЦИЙ

Часто в процессе экологического мониторинга по имеющимся экологическим данным сложившуюся ситуацию нельзя однозначно отнести к определенному классу. Например, одна нештатная ситуация может привести к выбросу вещества в интервале от 1,5 до предельно допустимого выброса 2,5, другая к выбросу того же вещества в интервале от 2,3 до 5,4. Если рассчитать концентрации, которые будут в случае этих двух нештатных ситуаций (при прочих одинаковых условиях), то получим пересечение интервалов измеряемых концентраций; т. е. значения, попадающие в это пересечение, не могут быть однозначно отнесены к одному классу. На практике возможно пересечение более двух классов. Необходимо, чтобы подсистема анализа экологической информации указывала оператору на возможность нескольких вариантов развития экологической ситуации.

При построении выборки для обучения нейронной сети возникает проблема, которая заключается в том, что один и тот же вектор X_i можно отнести к нескольким выделенным классам или, другими словами, одному вектору X_i ставится в соответствие несколько векторов Y . Для двух экологических параметров x_1 и x_2 данная ситуация проиллюстрирована на рис. 4.

Обучить нейронную сеть на данных, в которых есть пересечение множеств выходных векторов, нельзя, так как нейронная сеть не может быть обучена на противоречивых данных. Процесс обучения никогда не прекратится, так как необходимый для прекращения обучения минимум ошибки не будет достигнут. Предлагаются два решения данной проблемы.

Первое решение заключается в образовании новых классов, являющихся результатом пересечения базовых классов. Увеличение числа классов означает возрастание числа состояний, которые кодируются выходными векторами Y . В случае, если используется нейронная сеть, организованная по принципу «первый забирает все», число выходов растет пропорционально числу образованных классов. Когда наблюдается большое число пересечений базовых классов, данный подход приводит к значительному усложнению нейронной сети. Это вызывает снижение эффективности решения задачи


 Рис. 4. Пересечение классов K_1 , K_2 и K_3

 Рис. 5. Ансамбль нейронных сетей для классификации на пересекающиеся классы K_1 , K_2 и K_3

классификации нейронной сетью и уменьшение логической прозрачности структуры нейронной сети.

Второе решение основывается на построении для каждого множества отдельной нейронной сети. Для случая, изображенного на рис. 4, строится три нейронных сети. Первая нейронная сеть производит классификацию на множество K_1 и его дополнение $K_2 \cup K_3/K_1$, вторая — на K_2 и $K_1 \cup K_3/K_2$, третья — на K_3 и $K_1 \cup K_2/K_3$. На рис. 5 изображен ансамбль нейронных сетей, предназначенный для классификации на классы K_1 , K_2 и K_3 , пересекающиеся между собой.

В результате получаем множество нейронных сетей, каждая из которых производит классификацию на два класса, целевой и его дополнение. При подаче на вход вектора X_i каждая нейронная сеть относит его либо к целевому классу, либо к его дополнению. В том случае,

когда две и более нейронных сети относят вектор к целевому классу, это означает, что вектор принадлежит пересечению классов, соответствующих этим нейронным сетям.

Второй подход более эффективный и универсальный, так как обеспечивает более понятные правила классификации и не требует образования большого числа новых классов.

ЗАКЛЮЧЕНИЕ

В работе предложен подход к построению модуля идентификации экологической ситуации. Преимущество предложенного подхода заключается в возможности формирования логических правил, позволяющих идентифицировать экологическую ситуацию, которая ранее не возникала в процессе экологического мониторинга. Однако его эффективность снижается из-за неполноты исходной информации, используемой для моделирования. В списке нестандартных и аварийных ситуаций не реально предусмотреть все возможные варианты ситуаций. Исследование особенностей функционирования объекта управления и анализ возникающих ситуаций позволит получить дополнительную информацию, необходимую для построения модуля идентификации экологической ситуации.

Построение модуля идентификации экологической ситуации на основе предложенного подхода позволяет повысить адекватность принятия оперативных решений по управлению сложившейся экологической ситуацией и способствует безопасному функционированию потенциально опасных технологических объектов.

ЛИТЕРАТУРА

1. Дюк В., Самойленко А. Data mining: учебный курс. — СПб.: Питер, 2001. — 368 с.
2. Методика расчета концентраций в атмосферном воздухе вредных веществ, содержащихся в выбросах предприятий. ОНД-86. — Л.: Госкомгидромет, 1987. — 140 с.
3. Yoshida A. Two-dimensional numerical simulation of thermal structure of urban polluted atmosphere (effects of aerosol characteristics) // Atmos. Env. — 1991. — N 1. — P. 17–23.
4. Bianconi R., Tamponi M. A mathematical model of diffusion from a steady source of short duration in a finite mixing layer // Atmos. Env. — 1993. — N 5. — P. 781–792.
5. Chrysikopoulos C. V., Hildmann L. M., Roberts P. V. A three-dimensional steady-state atmospheric dispersion-deposition model for emission from a ground-level area source // Atmos. Env. — 1992. — N 5. — P. 747–757.

☎ (495) 334-92-49; e-mail: veresnikov@ipu.rssi.ru

Статья представлена к публикации членом редколлегии В. В. Кульбой. □