

АВТОМАТИЗАЦІЯ ВИРОБНИЧИХ ПРОЦЕСІВ

УДК 681.3:004.8:622.867

В.В. Слесарев, д-р. техн. наук, А.В. Миргородский, В.С. Маргитич

(Украина, г. Днепропетровск, Государственное высшее учебное заведение “Национальный горный университет”)

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ МОНИТОРИНГА АВАРИЙНОГО СОСТОЯНИЯ ШАХТЫ

Анотація. У статті на підставі аналізу інформації про параметри значень концентрації окису вуглецю, вуглекислого газу, температури і метану на аварійній ділянці, а також визначаючи поточну ситуацію усередині системи, виявлено кількісні характеристики відносин між екземплярами і класами пожеж. Відповідно до запропонованої методикою формування онтологій прецедентів, проведено кластерний аналіз даних моніторингу характеристик пожеж. На підставі цього виділено групи проблемних ситуацій на шахтах і були сформульовані правила розпізнавання класів аварій.

Ключові слова: Кластерний аналіз, кластеризація, дескриптори, ієрархічний агломеративного алгоритм, аналіз методом К-середніх.

Аннотация. В статье на основании анализа информации о параметрах значений концентрации окиси углерода, углекислого газа, температуры и метана на аварийном участке, а также определяя текущую ситуацию внутри системы, выявлены количественные характеристики отношений между экземплярами и классами пожаров. В соответствии с предложенной методикой формирования онтологий прецедентов, проведён кластерный анализ данных мониторинга характеристик пожаров. На основании этого выделены группы проблемных ситуаций на шахтах и были сформулированы правила распознавания классов аварий.

Ключевые слова: Кластерный анализ, кластеризация, дескрипторы, иерархический агломерационный алгоритм, анализ методом k-средних.

Abstract. The article based on the analysis of information about the parameters of the concentrations of carbon monoxide, carbon dioxide, methane, and the temperature at the emergency site, as well as defining the current situation within the system revealed the quantitative characteristics of the relationship between instances and classes of fires. In accordance with the proposed method of forming the ontology of precedents, conducted a cluster analysis of the data monitoring of the fires. Based on this isolated group of problem situations in the mines and were formulated rules of recognition classes of accidents.

Cluster analysis, clustering, descriptors, hierarchical agglomerative algorithm analysis of k-means.

Введение

Горное производство является сложной и опасной отраслью. В ходе ведения горных работ могут возникать различные опасности природного и технологического характера. Поэтому проблемы безопасности на объектах горного производства имеют особое значение. Анализ и оценка опасности возможных аварий являются одной из ключевых проблем промышленной безопасности. Для комплексного охвата состояния горного производства применяют системы контроля и управления, такие как УТАС [1-3]. Анализ производится на основании показаний датчиков, выдаваемых применяемой системой контроля, или устных по средствам оповещения от работников.

На данном этапе часто возникают затруднения. Автоматически измеряется лишь малое количество показателей, а доступная информация имеет высокую долю неоднозначности. В результате чего наблюдается отсутствие однозначных правил, критериев и методов оценки аварийных ситуаций. Особенно это проявляется в процессе уже ликвидации аварий, когда ограничен доступ в аварийную зону и в результате действия поражающих факторов разрушены датчики и системы контроля аварийных ситуаций. Все это приводит к ошибочным решениям, которые могут привести к тяжёлым последствиям. Чисто механического фиксирования показаний датчиков о состоянии среды зачастую недостаточно. Это создает сложности для точной и объективной классификации вида аварий и процессов их ликвидации. Таким образом, было установлено, что задача создания интеллектуальной системы для определения вида аварийной ситуации и способов ее устранения является актуальной для шахт и рудников.

Формулировка цели и задачи исследования

Конечная цель данной работы – реализовать метод классификации аварийных ситуаций в процессе их определения и последующей ликвидации.

Для этого необходимо разработать методы и систему классификации данных ситуаций с использованием элементов нечёткой логики и искусственного интеллекта.

Изложение основного материала исследований

Предложенная онтологическая модель процесса ликвидации аварий на шахтах позволила выявить основные категории базы знаний и связи между ними, которые описывают данную предметную область. Количественные характеристики эта модель выявить не может. Выявление и группирование значений характеристик позволит сформировать правила, определяющие состояние конкретной шахты в аварийных ситуациях и реализованных решений по ликвидации аварий. Под такими отношениями понимают данные, собранные при ликвидации всех известных успешно ликвидированных аварий на угольных шахтах (прецедентов). Исследование по взаимовлиянию друг на друга таких аварийных параметров, по данным автора, не проводилось. На самом же деле внутри этого процесса все аварийные показатели взаимодействуют друг с другом. Более того, каждая авария, и в частности пожар является уникальным событием. Поэтому выявление этих закономерностей позволит прогнозировать проблемные ситуации на основе анализа характеристик, поступающих с аварийного объекта в процессе ликвидации пожара. Кроме того, это позволит получить дополнительные данные об аварийных ситуациях, характерных для конкретного объекта управления.

Выявление классов проблемных ситуаций осуществляется одним из методов интеллектуального анализа данных – кластеризацией [4-6]. Цель кластерного анализа заключается в создании групп сходных (однородных) объектов. В то же время его действие состоит в привнесении структуры в анализируемые данные. Кластерный анализ является эвристической процедурой, не имеющей статистического обоснования, но даёт возможность произвести своеобразную разведку – изучить «структуру совокупности». Он применяется в случае, если нет априорных гипотез относительно классов на описательной стадии исследования. Разные кластерные методы могут порождать и порождают различные решения для одних и тех же данных.

В качестве дескрипторов (A'') для кластерного анализа характеристик аварийного процесса выбраны следующие параметры: окись углерода, углекислый газ, температура, метан (рис. 1):

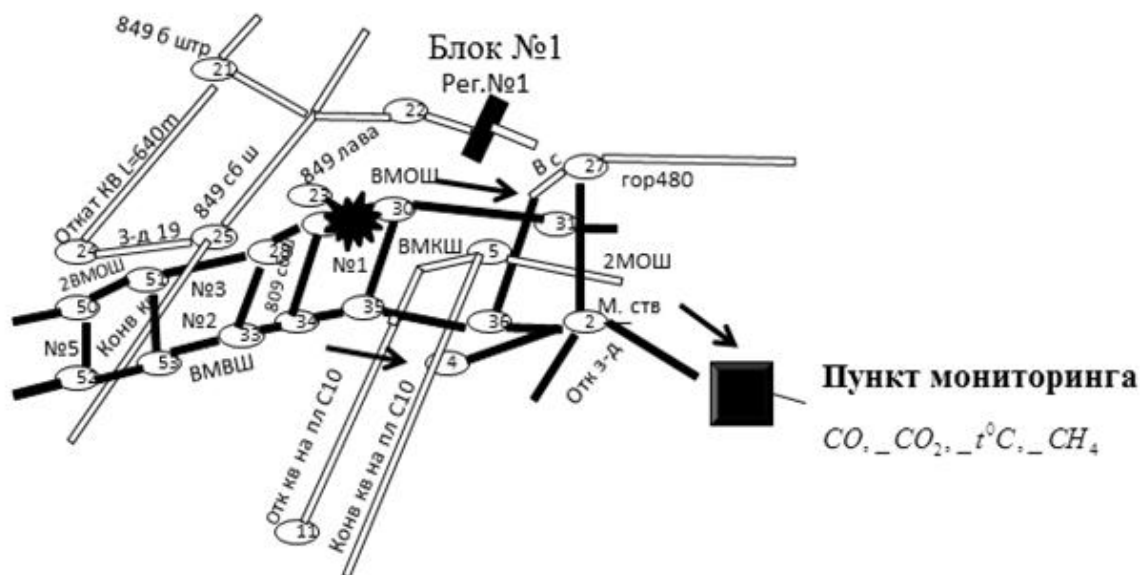


Рис. 1. Пункт мониторинга аварийных ситуаций на шахте

Для выявления непересекающихся подклассов и установления значений признаков необходимо решить задачу классификации прецедентов – пожаров, имевших место на угольных шахтах. Предположим, что можно разбить множество значений признаков, характеризующих проблемную ситуацию в реализованных решениях (управляющих воздействиях), обозначенное A'' , на непересекающиеся подмножества, соответствующие классам проблемной ситуации ($a_{i1}, \dots, a_{im}, Category_j$). То есть была подготовлена выборка данных из протоколов успешно ликвидированных аварий на шахтах в виде

$$\begin{aligned}
 A^n &= \{A_i, i = 1, \dots, n\}; \\
 A_i &= (a_{i1}, \dots, a_{ij}, \dots, a_{im}); \\
 \psi &: A_i \rightarrow \text{Category}_k; \\
 \forall A_i \in A^n \exists \text{category}_k \in \text{Category} : \psi(A_i) &= \text{category}_k,
 \end{aligned}
 \tag{1}$$

где A^n – множество данных протоколов ликвидации аварий; A_i – вектор значений данных; a_i – аварийный параметр; $\text{Category}_k = (\text{category}_i, \text{category}_z, \dots, \text{category}_j)$ – множество классов аварийных ситуаций.

Распознавание аварийных ситуаций состоит в том, чтобы на основании мониторинга объекта и анализа вектора признаков $a[i, j]$ данные в определённый момент времени отнести к определённому классу возможных аварийных ситуаций в случае, если производится оценка реализуемых решений (категорий) $cat_j \in \text{Cat}$, либо же уже наступившей аварийной ситуации (в случае если анализируются характеристики реализуемых решений). Далее полученный класс соотносится с характерным для него управляющим воздействием $b[i, j]$ и выдаётся рекомендации руководителю аварийных работ.

Исследования проводились в соответствии с методологией кластерного анализа, включающей пять основных этапов[5]:

- 1) отбор выборки для кластеризации;
- 2) определение множества признаков, по которым будут оцениваться объекты в выборке, и способов их стандартизации;
- 3) определение той или иной меры сходства между объектами;
- 4) применение метода кластерного анализа для создания групп сходных объектов;
- 5) проверка достоверности результатов кластерного анализа.

Для реализации кластерного анализа данных использовался модифицированный пакет SPSS 11.5. На начальном этапе проведения кластерного анализа неизвестно количество классов в онтологии прецедентов оперативных планов ликвидации аварий (ОПЛА). Для определения количества кластеров (классов ОПЛА) применяется иерархический агломерационный алгоритм. Его смысл заключается в следующем. Перед началом кластеризации все объекты считаются отдельными кластерами, которые в ходе кластеризации объединяются (агломерация). Вначале выбирается пара ближайших кластеров, которые в ходе алгоритма объединяются в один кластер. В результате количество кластеров становится равным $(N-1)$. Процедура повторяется, пока все кластеры не объединятся. На любом этапе объединение можно прервать, получив нужное число кластеров. Таким образом, в результате работы алгоритма агрегирования определяются способы вычисления расстояния между объектами, а также близости между кластерами. Выбор метода образования кластеров, метода расчёта дистанционной меры и меры подобия определяется с экспертом и учётом того, что рассматриваемые признаки представляют собой интервальные значения. В работе - это метод называется «связь между группами» (*Between groups linkage*). Дистанция между кластерами равняется среднему значению дистанций между всеми возможными парами наблюдений. Информация, необходимая для расчёта расстояния, находится на основании всех возможных пар наблюдений. В качестве дистанционной меры выбран квадрат евклидова расстояния (squared Euclidean distance): $dist = \sum_{i,j=1}^n (x_i - y_j)^2$. Значения признаков стандартизуются модифицированными средствами

SPSS 11.5. Для определения того, какое количество кластеров следовало бы считать оптимальным, была проанализирована агломерационная матрица, построенная на основании дистанционной меры – квадрат евклидова расстояния, и стандартизованных значений признаков. Нам необходимо проверить, является ли верным предположение о разбиении исходного множества данных мониторинга шахтной атмосферы в зоне пожара после реализации очередного оперативного плана на 3 класса проблемных ситуаций. Для этого после определения кластеров был проведён анализ данных методом k – «средних», на основании предшествующих исследований. Первые наблюдения, содержащиеся в файле с данными, используются как первый кластер. Вычисляется расстояние между центрами кластеров и каждым объектом, и объект приписывается к тому кластеру, к которому он ближе всего. Завершив приписывание, согласно алгоритму вычисляются средние значения для каждого кластера. Этих средних будет столько, сколько используется переменных для проведения анализа: в данном случае три. Набор k – средних представляет собой координаты нового положения кластера. По алгоритму вновь вычисляем расстояние от каждого объекта до центра кластеров и приписываем объекты к ближайшему кластеру. Вновь вычисляются центры тяжести кластеров, и этот процесс повторяется до тех пор, пока центры тяжести не перестанут «мигрировать» в пространстве.

На рис. 2 приведены результаты кластерного анализа, а в таблице представлен фрагмент, состоящий из 15 наблюдений – данных мониторинга шахтной атмосферы в исходящей струе воздуха зоны пожара,

дается информация о том, к какому из трёх классов было отнесено конкретное наблюдение, а также расстояние до центра одного из трёх кластеров анализируемых наблюдений.

Таблица

Фрагмент анализируемых наблюдений и результатов кластерного анализа

Наблюдения	Окись углерода CO, %	Температура, t ^o C	Углекислый газ CO ₂ , %	Метан, % CH ₄	№ класса	Расстояние до центра кластера
1	2	3	4	5	6	7
1	0,0012	38,00	3,60	0,50	1	0,0300767
2	0,0011	35,00	3,25	1,10	1	0,0570733
3	0,0077	50,00	2,55	0,95	2	0,0356737
4	0,0030	20,00	2,25	0,10	2	0,0415996
5	0,0082	25,00	2,50	1,10	2	0,0420257
6	0,0076	15,00	4,25	0,95	2	0,0707088
7	0,0104	18,00	3,30	0,01	2	0,0505453
8	0,0117	20,00	3,85	0,09	2	0,0402976
9	0,0073	18,00	3,60	1,40	3	0,0334702
10	0,0080	20,00	4,05	0,95	3	0,0201597
11	0,0067	22,00	3,05	0,20	3	0,0451907
12	0,0032	14,00	4,25	0,50	3	0,0510968
13	0,0026	41,00	4,50	1,10	1	0,0593515
14	0,0033	18,00	2,55	1,50	2	0,0512553
15	0,0028	36,00	2,25	0,90	1	0,0026668

HIERARCHICAL CLUSTER ANALYSIS
Dendrogramma using Average Linkage (Between Groups)
Rescaled Distance Cluster Combine

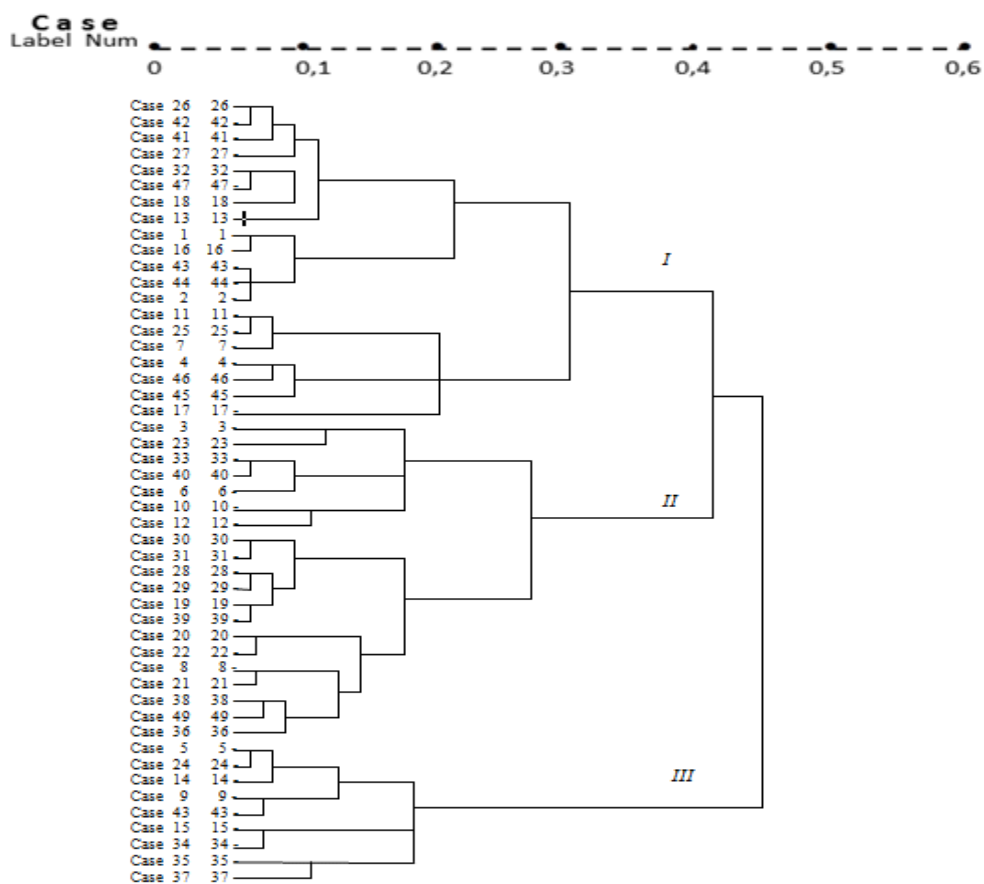


Рис. 2. Совокупность результатов кластерного анализа

На рис. 3 изображено распределение данных мониторинга трёх величин в исходящей струе аварийных зон: окиси углерода, углекислого газа, температуры, что подтверждает возможность классификации аварийных ситуаций на шахтах методом кластеризации.

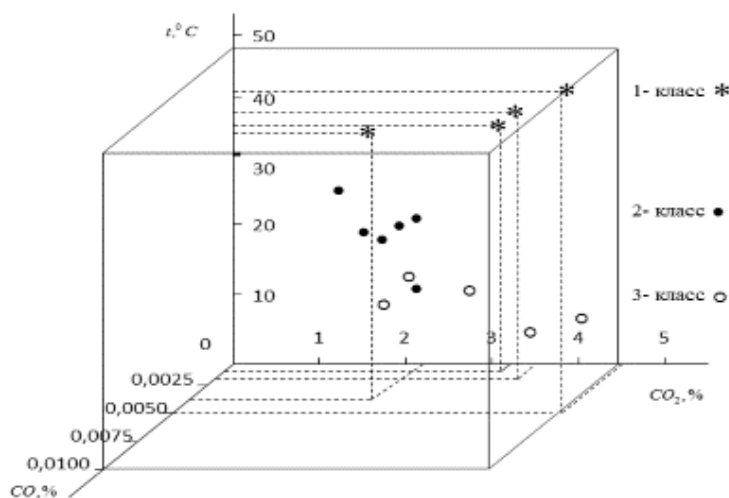


Рис. 3. Распределение данных мониторинга шахтной атмосферы в пространстве

Расстояние между центрами кластеров получилось достаточно большим по сравнению с расстоянием между результатами мониторинга и центрами соответствующих кластеров, что позволяет сделать вывод об успешном проведении кластерного анализа методом k – средних.

Окончательные значения кластеров были получены уже на третьей итерации. С помощью метода k – средних было подтверждено, что исходное множество данных мониторинга шахтной атмосферы в зоне пожара было верно разбито на три предопределённых класса, которые получились достаточно однородными и равноценными. Проанализировав результаты кластеризации: выявлены закономерности и были разработаны правила, позволяющие по значениям признаков отнести аварийную ситуацию к одному из трёх классов. Правила позволяют предсказать (на основании данных о поступающей аварийной ситуации) возможность возникновения аварийной ситуации, а также точнее определить границы возникновения аварийной ситуации для конкретного объекта управления.

Выводы

Анализируя поступающую информацию о значениях окиси углерода, углекислого газа, температуры и метана на аварийном участке, а также определяя текущую ситуацию внутри системы, выявлены количественные характеристики и отношения между составляющими и классами аварий. В соответствии с предложенной методикой формирования онтологий прецедентов, проведён кластерный анализ данных мониторинга процессов ликвидации пожаров на шахтах. На основании выделенных групп с экспертами были сформулированы правила распознавания классов проблемной ситуации. С помощью этих правил представляется возможность руководителю аварийных работ установить класс возникшей аварийной ситуации на шахте, а также управлять процессом её ликвидации.

Список литературы

1. Лекции по системе УТАС: структура, эксплуатация и обслуживание / Нормативный документ Министерства топлива и энергетики Украины. – Донецк: 2004. – 100 с;
2. Научные основы автоматизации в угольной промышленности: опыт и перспективы развития: монография / В.Г. Курносов, В.И. Силаев; Международный институт независимых педагогических исследований МИНПИ – ЮНЕСКО, ОАО «АВТОМАТГОРМАШ им. В.А. Антипова». – Донецк: изд. «Вебер» (Донецкое отделение), - 2009. – 422
3. Курносов В.Г. Основные направления развития работ по созданию автоматизированных систем оперативно-диспетчерского управления горным предприятием // В.Г. Курносов, д.т.н., В.И. Силаев, д.т.н., В.В. Синенко, к.т.н., А.А. Винарик, д.т.н., // Донецк: «Вебер» (Донецкое отделение), - 2009. – 227 с.
4. Maximov V.N., Bulgakov N.G., Levich A.P. Quantitative methods of ecological control: Diagnostics, Standardization, and Prediction // Environmental indices: System Analysis Approach. London: EQLSS Publishers. – 1999. – pp.363-381.
5. Web Ontology Language (OWL)/ Edited by Miller E., 2004 // The World Wide Web Consortium (WWWC) // <http://www.w3.org/2004/OWL>.
6. Hrocks I., Patel-Schneider P.F., Bechhofer S. and Tsarkov D. OWL rules: A proposal and prototype implementation // Web Semantics: Seintics, Services and Agents on the World Wide Web, Vol. 3(1), 2005. – pp. 23-40.

Рекомендовано до друку: д-ром техн. наук, проф. Самусею В.І.