

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 622:519.25:004.42

А.В. Дерюшев, М.П. Лазеева, А.Г. Пимонов

СТАТИСТИЧЕСКИЙ АНАЛИЗ И РАСПОЗНАВАНИЕ ОБРАЗОВ В УПРАВЛЕНИИ ГОРНЫМ ПРЕДПРИЯТИЕМ

По прогнозу вероятного уровня спроса на уголь в ближайшие десятилетия [1] и с учетом обеспечения минимального «порогового» уровня энергетической безопасности государства [2] годовой объем добычи угля в России, а также, соответственно, и в Кузбассе будет увеличиваться (рис. 1).

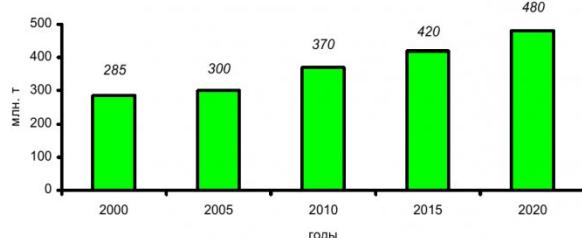


Рис. 1. Прогноз добычи угля в России

Балансовые запасы каменного угля в Кузбассе по категориям A+B+C1 оценивают в 58,8 млрд. т [3]. При сегодняшних темпах добычи 150 млн. т в год [4] его хватит почти на 400 лет. Но в связи с реструктуризацией угольной промышленности в России прекратили добычу 140 особо убыточных и опасных по условиям труда шахт [5]. Следовательно, увеличение объемов добычи угля в ближайшие 20 лет необходимо связывать с реконструкцией существующих и строительством новых высокопроизводительных горных предприятий с подземной и открытой добычей угля, с применением новейших технологий, основанных на автоматизации производственных и вспомогательных процессов. Проблема стабильного роста угольной промышленности тесно связана с проблемой стабильного и комплексного развития конкретного горнодобывающего предприятия, его "устойчивости" [6]. Концепция устойчивого развития предприятия и горнодобывающей отрасли в целом основывается на принятии своевременных, всесторонне взвешенных, практических и

научно обоснованных эффективных управленческих решений [7]. Поэтому в настоящее время наука об оптимальном управлении сложными динамическими системами получила значительное развитие и в горном деле. Наряду с разработкой и углублением сложившихся направлений возникли новые, возросло значение теории и техники управления. Практическим результатом научных исследований является **автоматизированная система управления** (АСУ) – управляющая система, часть функций которой, главным образом функцию принятия решений, выполняет человек. В зависимости от объектов управления различают, например: АСУП, когда объектом управления является предприятие; АСУТП, когда объектом управления является технологический процесс; ОАСУ, когда объектом управления является организационный объект или комплекс.

Управленческие решения принимают на основе анализа, в том числе и статистического, информации, которая может быть получена различными путями. Если предприятие рассматривать как сложную динамическую систему, то его состояние можно охарактеризовать различными социально-экономическими, технологическими, геомеханическими и другими стохастическими (случайными, вероятностными) факторами, которые могут служить источниками такого рода информации. Если удастся собрать необходимую информацию

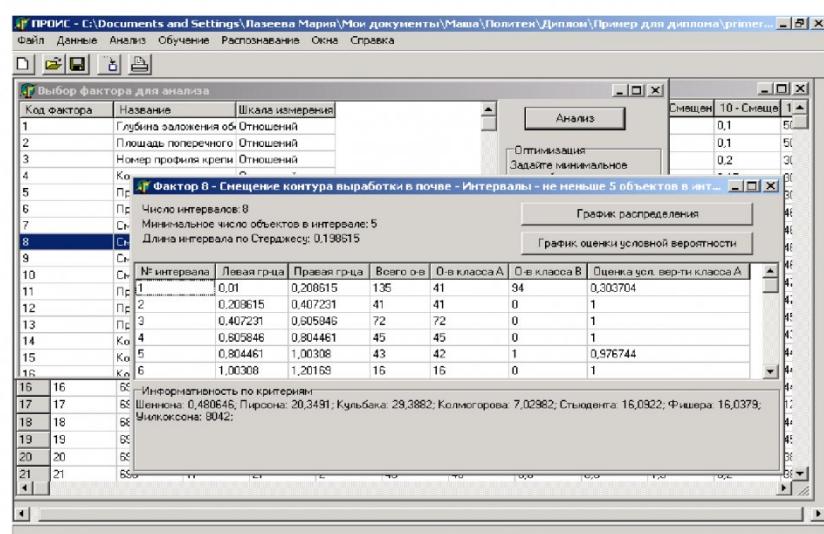


Рис. 2. Интерфейс ППП "ПРОИС" 2.0.0

о системе, можно осуществить текущий, краткосрочный или долгосрочный прогноз изменения интересующих нас показателей, характеризующих состояние системы, в том числе её устойчивость или, пользуясь терминами теории надежности [8], – безотказность работы. Результаты такого прогноза [9] могут быть использованы в АСУ для принятия проектных, плановых и оперативных управлеченческих решений. Таким образом, роль прогноза устойчивой (безотказной) работы предприятия как сложной системы и отдельных его подсистем в настоящее время является крайне важной, а разработка методов прогноза – актуальной проблемой горного производства.

При управлении сложными техническими системами и процессами горного производства часто необходимо принимать наиболее эффективные в заданных условиях решения. Во многих случаях такого рода задачи специалисты решают на интуитивном уровне, основываясь на своем профессиональном опыте и знаниях. Зачастую для принятия таких решений необходимо установить принадлежность исследуемого объекта к одному из классов [10] (образов) – множеству объектов, объединенных общими свойствами. Для решения указанной проблемы можно использовать методы теории **распознавания образов**, как научно-технического направления, связанного с разработкой алгоритмов и построением на базе ЭВМ **распознающих систем** для установления принадлежности некоторого объекта к одному из заранее выделенных классов. Процесс распознавания основан на сопоставлении признаков (характеристик) исследуемого объекта с признаками других известных объектов, в результате чего делается вывод о наиболее правдоподобном их сопоставлении. В качестве распознающей системы следует рассматривать систему «человек-машина», где человеком является специалист в определенной области знаний, принимающий управлеченческие решения, а машиной – **распознающий аппарат** (ЭВМ). Распознающая система может являться подсистемой автоматизированной системы управления горным предприятием. К числу типичных подсистем относятся подсистемы: измерения, сбора, обработки, распределения и передачи информации о состоянии объекта управления, входных и внутренних воздействиях; принятия решения; контроля; диагностирования; противоаварийного управления; восстановительного управления; передачи и осуществления воздействий; сбора и представления информации человеку-оператору; документирования. Подсистемы могут содержать в себе управляющие устройства, регуляторы, контроллеры и другие элементы.

Чтобы осуществить обучение распознающего аппарата, необходимо произвести большой объем вычислений. Поэтому интерес к использованию методов распознавания возрос с появлением средств вычислительной техники с высоким быст-

родействием. Первоначально широкое распространение получили различные алгоритмы распознавания, основанные на методах математической статистики. Однако, с развитием нейроинформатики [11] эти методы были, на наш взгляд, незаслуженно отодвинуты на второй план. Распознавание на основе нейронных сетей позволяет эффективно решать ряд задач классификации. Но в большинстве случаев для решения конкретных задач необходимо создавать новую сеть, приспособленную для использования в данной области знания. Статистические же методы более универсальны, многие из них могут быть использованы для решения задач классификации и прогноза в различных отраслях науки.

С целью создания распознающего аппарата нами был разработан **пакет прикладных программ (ППП) "ПРОИС"** [12] версии 2.0.0 (рис. 2), в котором реализован алгоритм вероятностно-статистического непараметрического метода распознавания образов [13]. Он предназначен для решения задач классификации и прогноза и позволяет разделить незнакомые объекты – предметы или явления, исследуемые в конкретной задаче и характеризующиеся некоторым конечным числом признаков, на два класса.

Пакет позволяет автоматизировать расчеты при обучении распознающего аппарата и непосредственно при распознавании, предоставляет возможности для статистического анализа исходных данных путем вычисления основных точечных статистик выборки значений факторов: минимального и максимального значений, математического ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации, а также определяет число объектов в выборке, число объектов первого и второго классов, оценку условной вероятности каждого из двух классов.

Для каждого фактора производится построение эмпирической функции распределения с возможностью графического анализа распределения объектов по интервалам.

В реализованном в ППП "ПРОИС" 2.0.0 алгоритме можно выделить четыре основных этапа:

- 1) формирование обучающей выборки с эффективной для распознавания системой факторов;
- 2) построение решающего правила распознавания;
- 3) оценка надежности решающего правила при распознавании контрольной выборки;
- 4) собственно распознавание неизвестных объектов.

Для обучения распознающего аппарата необходимо сформировать **исходную обучающую выборку** из N объектов с указанием принадлежности каждого из них к одному или другому классу. Каждый объект характеризуется наибольшим количеством M факторов (признаков), количественно или качественно описывающих те или иные свойства исследуемого объекта и влияющих на

принадлежность объекта к тому или иному классу, с точки зрения пользователя ППП "ПРОИС". При вводе данных исходной обучающей выборки в электронные таблицы ППП производит автоматическую проверку вводимых значений факторов по установленным пользователем формату и граничным условиям.

Каждый i -й объект исходной обучающей выборки можно представить в виде точки T_i в гиперпространстве (рис. 3). Координатами таких точек являются значения факторов Y_j ($j=1, 2, 3, \dots, M$). Тогда объекты класса A с одинаковыми свойствами сформируют в этом гиперпространстве одну область (образ), объекты класса B с противоположными свойствами – другую.

Можно определить некоторые характеристики положения образов в гиперпространстве, например: координаты центров тяжести образов класса A - O_A и класса B - O_B , расстояние между центрами тяжести образов - ρ_{AB} , расстояние от начала координат до центров тяжести образов - ρ_{OA} и ρ_{OB} , угол между радиус-векторами центров тяжести образов - φ и т. д.

Значения каждого фактора содержат в большей или меньшей степени информацию о принадлежности объекта к тому или иному классу. ППП "ПРОИС" 2.0.0 производит оценку информативности факторов по семи критериям: J - критерий Шеннона; χ^2 - критерий Пирсона; λ - критерий Колмогорова; I - критерий Кульбака; t - критерий Стьюдента; F - критерий Фишера; U - критерий Уилкоксона.

Все оценки информативности сведены в таблицу, а факторы отсортированы и соответственно пронумерованы по мере уменьшения их информативности методом обобщенной ранжировки с учетом мнений всех экспертов.

На основе полученных оценок информативности факторов пользователь осуществляет профессиональный анализ исходной обучающей выборки и производит **минимизацию признакового пространства** путем удаления из выборки неинформативных и малоинформационных факторов. При минимизации признакового пространства большую роль играет профессиональный анализ факторов пользователем – экспертом в конкретной области знаний. Оставшаяся в результате минимизации признакового пространства совокупность наиболее информативных факторов является **окончательной обучающей выборкой** с эффективной для распознавания системой факторов, которая будет использована в процессе обучения распознавающего аппарата.

Поскольку факторы могут быть измерены в различных шкалах, иметь разные пределы и единицы измерения, то объекты рассеяны во всевозможных октантах гиперпространства и компактность образов мала. Для устранения этих проблем и повышения эффективности обучения необходимо ввести безразмерные координаты объектов в

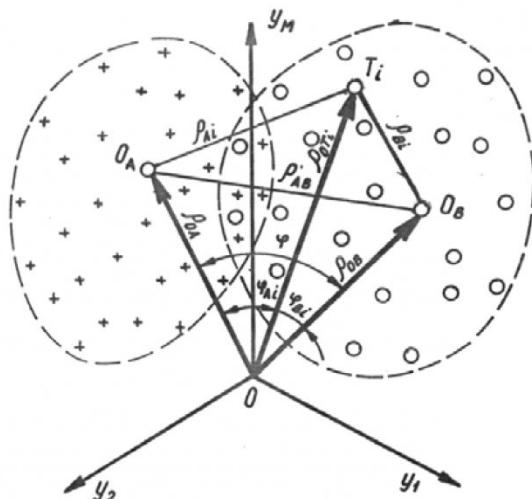


Рис. 3. Геометрическая интерпретация положения образов в гиперпространстве

гиперпространстве. Для перевода значений факторов в безразмерные координаты в качестве новой безразмерной координаты может быть использована оценка условной вероятности принадлежности объекта первому из классов, которая рассчитывается в таблицах перевода, содержащих границы интервалов распределения значений факторов и оценки условной вероятности первого из классов, соответствующие каждому интервалу. На основе полученных по каждому фактору безразмерных координат объектов строится **решающее правило распознавания** – алгоритм, по которому на основе анализа значений вероятностных характеристик факторов делается вывод о том, к какому классу принадлежит объект.

С учетом геометрической интерпретации положения образов в гиперпространстве (рис. 3) производится разработка восьми решающих правил распознавания:

- 1) по разности расстояний от объекта до центра тяжести образов - $\rho_{Ab} \rho_{Bb}$;
- 2) по разности углов между радиус-векторами объектов и центров тяжести образов - $\varphi_{Ab} \varphi_{Bb}$;
- 3) по разности скалярных произведений радиус-векторов объектов и центров тяжести образов;
- 4) по разности между коэффициентами корреляции объектов с центрами тяжести образов;
- 5) по расстоянию от объектов до гиперплоскости, проходящей через середину отрезка, соединяющего центры тяжести образов и нормально к нему;
- 6) по расстоянию от начала координат до каждого объекта - ρ_{oTi} ;
- 7) с помощью разделяющего гиперэллипсоида;
- 8) по обобщенной функции желательности.

По каждому из решающих правил производится построение прогнозной таблицы, которая

содержит границы интервалов распределения обобщенного признака, а также оценки условной вероятности принадлежности объекта к первому из классов, соответствующие каждому интервалу.

Для **оценки надежности решающего правила** используется контрольная выборка – множество объектов, описываемых окончательной системой факторов, принадлежность которых к тому или другому классу известна пользователю и которые не были включены в обучающую выборку. Ввод контрольной выборки объектов производится аналогично вводу исходной обучающей выборки. ППП производит распознавание по каждому из 8 решающих правил и вычисляет оценки ошибок прогноза: ошибка первого рода – когда объект первого из классов идентифицируется как объект второго класса; ошибка второго рода – когда объект второго из классов идентифицируется как объект первого класса; а также суммарную ошибку. При этом распознавание объектов контрольной выборки можно провести с использованием различных порогов принятия решения – минимальных значений оценки условной вероятности, при котором объект будет отнесен к первому из классов. В результате анализа ошибок распознавания из 8 решающих правил пользователь выделяет одно **оптимальное решающее правило распознавания** и оптимальный порог принятия реше-

ния, при которых ошибки распознавания минимальны. Выбранное оптимальное решающее правило и оптимальный порог принятия решения в дальнейшем используются для распознавания неизвестных объектов. На этом обучение распознавающего аппарата можно считать завершенным.

Для **распознавания неизвестных объектов**, подлежащих классификации, составляют **экзаменационную выборку** аналогично обучающей и контрольной выборкам. Обученный распознавающий аппарат производит перевод значений факторов экзаменационной выборки в безразмерные координаты, по оптимальному решающему правилу распознавания производит оценку условной вероятности принадлежности каждого объекта к первому из классов, определяет класс объекта при установленном пороге принятия решения.

Разработанный ППП "ПРОИС" 2.0.0 является ядром программного комплекса, предназначенного для распознавания образов и статистической обработки данных. Пакет применим для решения задач классификации и прогноза в различных областях знаний. ППП "ПРОИС" предполагается использовать в Кузбасском государственном техническом университете для прогнозирования надежности горных выработок, геомеханических процессов и при разработке автоматизированной системы управления горным предприятием.

СПИСОК ЛИТЕРАТУРЫ

1. Кузнецов В. И. Концепция развития открытой добычи угля в Кузбассе// Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 18-21.
2. Кочетков В. Н. Об основных тенденциях в развитии угольной промышленности/ В. Н. Кочетков, С. Н. Лазаренко // Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 39-44.
3. Брагин В. Е. Состояние и перспективы развития угольной промышленности Кузбасса Кузбассе// Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 21-23.
4. Тулеев А. Г. Кемеровской области – 60 лет // Уголь.- 2003.- № 1.- С. 2.
5. Малышев Ю. Н. Приоритетные направления развития науки и производства в угольной промышленности// Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 9-14.
6. Кочетков В. Н. Системный подход к моделированию устойчивого развития угледобывающего региона/ В. Н. Кочетков, А. А. Новосельцев// Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 99-102.
7. Кузнецов А. С. К основам информатизации и автоматизации управления горным производством// Природные и интеллектуальные ресурсы Сибири: Материалы III Международ. науч.-практ. конф. "Сибресурс-99" (16-18 ноября 1999 г.) - Кемерово, 1999.- С. 80-83.
8. Барзилович Е. Ю. Вопросы математической теории надежности/ Е. Ю. Барзилович, Ю. К. Беляев, В. А. Каиштанов и др.; Под ред. Б. В. Гнеденко.- М.: Радио и связь, 1983.- 376 с.
9. Пимонов А.Г. Автоматизированный прогноз разрушения горных пород// Вестн. КузГТУ.- 1997.- №1. - С. 4-8.
10. Гитис Л.Х. Статистическая классификация и кластерный анализ.- М.: Издательство Московского государственного горного университета, 2003.- 157 с.
11. Осовский С. Нейронные сети для обработки информации / Пер. с польского И. Д. Рудинского.- М.: Финансы и статистика, 2002.- 344 с.

12. Лазеева М.П. Программная реализация вероятностно-статистического непараметрического метода распознавания образов/ М.П. Лазеева, А.В. Дерюшев// Вестн. КузГТУ.- 2004.- №4.- С. 117-119.

13. Мякишева Л. Е. Использование ЭВМ для решения задач методами распознавания образов: Методические указания по применению библиотеки научных программ "ПРОИС" / Л. Е. Мякишева, А. В. Дерюшев; КузНИИшахтострой; Кузбас. политехн. ин-т.- Кемерово, 1986.- 40 с.

□ Авторы статьи:

Дерюшев
Александр Владимирович
- канд. техн. наук, доц. каф. строительства подземных сооружений и шахт

Лазеева
Мария Петровна
- асс. каф. вычислительной техники и информационных технологий

Пимонов
Александр Григорьевич
- докт. техн. наук, проф. каф. вычислительной техники и информационных технологий

УДК 519.21

А.С.Сорокин

ВЫБОР ОБОРУДОВАНИЯ И СХЕМ УГЛЕСОСНЫХ СТАНЦИЙ

На основе математических моделей надежности трех технологических схем углесосных станций производится выбор лучшей из них (рис. 1-3).

Функционирование различных систем, рассматриваемых в теории надежности, может быть описано полумарковским процессом с конечным числом состояний [1], [2], построенным соответствующим образом.

Из определения полумарковского процесса следует, что при переходе из i -го в j -ое состояние дальнейшее Поведение процесса зависит от времени пребывания его в j -ом состоянии. Это дает возможность составить систему стохастических уравнений, которая является аналогом формулы полной вероятности для случайных величин.

Эти системы уравнений представляют основную математическую модель [3], [4] (см. также [5], [6]).

На основе полумарковских процессов могут быть получены различные характеристики надежности. Предполагаем, что каждый работающий элемент углесосной станции создает простейший поток отказов с параметром $\lambda=1/T_p$ (I_{arp} , I_k , I_3 - соответственно для агрегата, обратного клапана, задвижки).

Время ремонта отказавшего элемента распределено по показательному закону с параметром $\mu=1/T_B$ (m_{arp} , m_k , m_3 - соответственно для агрегата обратного клапана, задвижки).

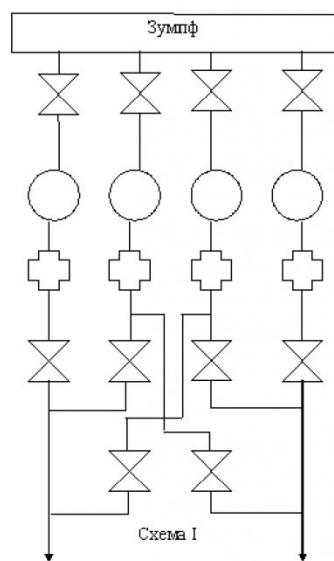


Рис. 1. Блок-схема углесосной станции

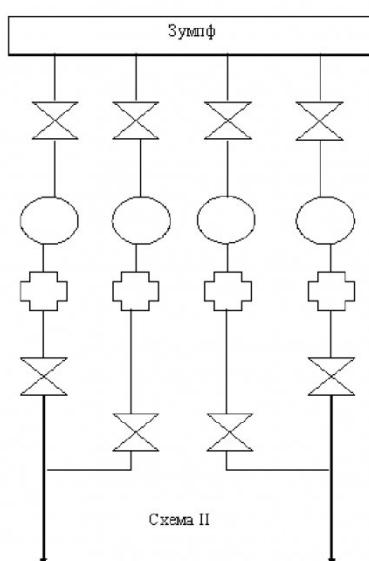


Рис. 2. Блок-схема углесосной станции

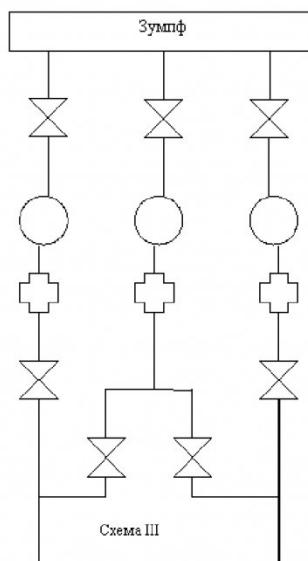


Рис. 3. Блок-схема углесосной станции



Углесосный агрегат



задвижка



Обратный клапан