

Особые темы *First Break*

Рудная геология и геофизика

Mining Geoscience

Из материалов рубрики «Особая тема» этого номера, посвященной рудной геологии и геофизике, становится очевидным, что специалисты рудной и нефтегазовой отраслей применяют в целом сходные подходы и методики. Одно из существенных различий состоит в том, что в рудной отрасли назрела более очевидная потребность в улучшении методов исследования, о чем говорят авторы всех трех статей рубрики. В обзорной статье Барретта и Уильямса говорится, что основной проблемой, с ростом объемов данных, могут оказаться их визуализация и доступ к ним; с этой же проблемой сталкивается нефтегазовая отрасль. В статьях, представленных компаниями Geosoft и Quantec речь идет о более «лобовых», но все равно полезных решениях.

Образование и повышение квалификации

Education and Training

Вторая «Особая тема» – образование и повышение квалификации – посвящена вопросам «смены команды» разведочной отрасли в целом, привлечению и закреплению выпускников, которые заменят когорты, уходящие на заслуженный отдых. Рассматриваются также вопросы создания постоянных программ повышения квалификации, которые позволят поддерживать работников на современном уровне быстро развивающихся технологий. Приводятся точки зрения двух компаний, занимающихся практическим решением этих задач.



Тематика рубрики «Особая тема»

Январь	<i>Информационные технологии и управление данными</i>
Февраль	<i>Наземная сейсморазведка</i>
Март	<i>Визуализация и интерпретация данных</i>
Апрель	<i>Нефтяная геология</i>
Май	<i>Скважинная сейсморазведка</i>
Июнь	<i>Междисциплинарный выпуск</i>
Июль	<i>Образование и повышение квалификации / Рудная геология и геофизика</i>
Август	<i>Экологическая и инженерная геология и геофизика</i>
Сентябрь	<i>Обработка данных</i>
Октябрь	<i>Геология и геофизика на эксплуатируемых месторождениях</i>
Ноябрь	<i>Несейсмические методы и аэрогеофизика</i>
Декабрь	<i>Морская сейсморазведка</i>
<i>В течение года тематика рубрики может быть расширена.</i>	

Разведка месторождений минерального сырья и современные методы работы с данными Mineral exploration using modern data mining techniques

Колин Т. Барнетт (Colin T. Barnett)¹ и Питер М. Уильямс (Peter M. Williams)² обсуждают, как новые идеи работы с данными, например визуализация и вероятностный подход к решению обратной задачи, могут способствовать совершенствованию разведочных работ в рудной отрасли. Эта статья является сокращенным вариантом статьи из тематического сборника* Общества геологов-экономистов (Society of Economic Geologists), которое в этом году отмечает свое столетие.

В этой работе анализируется состояние дел в разведке месторождений золота, в которой, как показано в (Schodde, 2004) предложение с трудом покрывает спрос. За последние 20 лет затраты на открытие нового месторождения выросли вчетверо, а средний размер запасов на месторождении сократился на 30%. Рентабельность отрасли составляет 5-7 %, что сопоставимо со ставкой кредита. Почему так происходит, и что с этим делать? В работе (Paterson, 2003) отмечается, что открытия следовали волнами, после внедрения новых методов или новых представлений о генезисе руд. С скачок в количестве открытых месторождений произошел, например в 1950-1975 гг., после внедрения новых методов и аппаратуры в разведочной геофизике и геохимии. В последнюю четверть XX в. произошел прорыв в микроэлектронике и компьютерных технологиях, что привело к значительному росту качества и количества данных. Очевидно, однако, что сами по себе эти новшества не смогли остановить снижение прироста открытых месторождений за тот же период.

Где же искать новые методы, которые вызовут следующую волну открытий? Похоже, что мы получаем данные быстрее, чем можем их потребить. Но та же проблема стоит в биоинформатике с расшифровкой геномов, или в осмыслиении огромных объемов информации в Интернете. В этой работе проводится тезис о том, что необходимо найти новые методы извлечения полезной информации из данных. Особое внимание следует обратить на недавние разработки в области визуализации и работы с данными.

Многие новые методы работы с данными имитируют человеческий разум и дают новый подход к использованию вычислительной техники. Обычно он применяется в задачах, для которых человек может найти алгоритм решения. Но живые существа программируются по мере накопления опыта, а не явным перечислением всех шагов процесса.

Идея нового подхода состоит, таким образом, в том, чтобы машины, как люди, обучались, получая данные.

Обучение машин разделяют в целом на контролируемое и неконтролируемое. Контролируемое обучение, или обучение на эталонных объектах (образцах), подразумевает наличие и доступность большого количества объектов с известными свойствами и откликами. Эти объекты составляют множество так называемых известных парных соответствий, называемое обычно обучающим множеством, и задача состоит в том, чтобы научить правильно определять эти соответствия в примерах. В разведочной отрасли обучающее множество составляется обычно из известных месторождений и разведанных областей. При неконтролируемом обучении известны только отклики, а свойства неизвестны. В этом случае цель состоит в обнаружении «интересных» особенностей данных – кластеров, границ, скрытых структур – о происхождении которых можно сделать предположения. В этой статье рассматривается только обучение на образцах; оба подхода обсуждаются подробнее в работе (Williams, 2002).

В начале этой статьи дается обзор последних достижений в визуализации и методах обучения на образцах, таких как нейронные сети. Далее приводится пример использования этих идей при работах на золото на участке Уокер Лейн (Walker Lane) на западе США. В конце работы показано, как эти результаты применяются для количественной оценки достоверности результатов и как повышение точности определения объектов позволяет снизить стоимость разведки и повысить вероятность успеха.

Визуализация

Визуализация является важным методом работы с данными. Человеческий мозг имеет огромные возможности расшифровки и осмыслиения сложных данных, если они представлены в визуальной форме. За последние 20 лет сделаны важные шаги в визуальном представлении информации, что позволяет опытному интерпретатору легче их воспринимать.

* Barnett, C.T. and Williams, P.M. [2006] Mineral exploration using modern data mining techniques. In Doggett, M.E. and Parry, J.R. (Eds.) *Wealth Creation in the Minerals Industry: Integrating Science, Business and Education*. Society of Economic Geologists Special Publication 12, 295-310.

¹BWG Mining, 424 Mapleton Avenue, Boulder, Colorado 80304, USA. ²BWG Mining, 28 Eaton Place, Brighton BN2 1EG, UK.

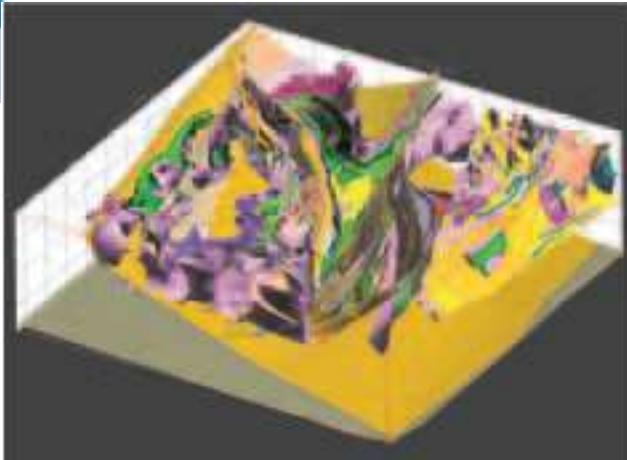


Рис. 1. Объемная геологическая модель рудного района Калгури (Kalgoorlie). Вид с юго-востока. Размеры модели 112x115x36 км. Красным отмечены шахты Кундана (Kundana) и Кановна Бель (Kanowna Belle). Серый слой в основании показывает границу Мохо, коричневый – плоскость разлома Ида (Ida). Гранитные интрузии показаны пурпурным цветом. С разрешения Грэнни Смит (Granny Smith), компания Placer.

2D представление данных

Одним из таких шагов было внедрение цветного изображения. Это дало возможность строить цветные рельефные карты или, для космоснимков, представлять их в условных цветах. Цветные изображения гораздо более наглядны, чем ранее употреблявшиеся черно-белые карты изолиний и схемы расположения точек; их проще интерпретировать. Кроме того, меняя угол освещения, можно выделить и проследить тонкие особенности в отдельном наборе данных.

Следующим шагом стало совмещение данных из разных источников в так называемых геоинформационных системах (ГИС), позволяющих совмещать наборы данных по площадям, оконтуривая их, прослеживая границы, или накладывая друг на друга в едином масштабе. В современных системах ГИС (ArcView, MapInfo) имеются также средства для управления базами данных и формирования запросов к ним, для сбора информации, которую пользователь хочет визуализировать, в точках, по профилям и на площадях.

3D интерпретация

Значительный прогресс достигнут также в обработке данных перед визуализацией. Стало возможно определение свойств пород путем решения трехмерной обратной задачи (Li and Oldenburg, 1996, 1998). Поэтому теперь, вместо того, чтобы просто разглядывать карты поля силы тяжести или магнитного поля в миллигалах или нанотеслах, можно использовать результаты автоматической интерпретации для распределения плотности и магнитной восприимчивости. Поскольку закономерности потенциальных полей хорошо

известны, решение обратной задачи для любого набора полевых данных можно проводить, задавая ограничения на основе любой геологической и физической априорной информации и получать удовлетворительный результат.

За последние несколько лет эти и другие технологии вошли в состав пакетов 3D визуализации, например, Gocad. Появился термин «общая модель земли» (common earth model, CEM [Cр. «физико-геологическая модель», ФГМ, в Никитин, Хмелевской, «Комплексирование геофизических методов», 2004 и др. учебники на русском; далее используется этот термин – Прим. перев.]), используемый для моделей геологического строения, построенных по данным различных дисциплин (Garrett et al., 1997; McGaughey and Vallee, 1998). ФГМ может включать 1D данные по бурению в точках (или по профилям), 2D данные – карты и разрезы, и 3D блочные модели распределения геологических, геохимических и геофизических параметров. Модель при этом должна быть топологически корректной, с замкнутыми границами, без разрывов между формациями.

На рис. 1 представлена ФГМ рудного района Калгури (Kalgoorlie). От начала (карты рельефа поверхности) до конца (геологическое строение разреза) эта 3D модель построена в пакете Gocad с учетом всех доступных данных бурения и геофизики. Положение поверхности Мохо и плоскости крупного разлома Ида (Ida). Другие формации, например гранитные интрузии (пурпурный цвет) и сопутствующие железорудные формации (показаны черным), выделяются по обнажениям, профилям скважин, и, в не меньшей степени, по результатам решения обратной задачи для данных гравиразведки и магниторазведки. (Р. Kowalczyk, личная беседа, 2004). Созданной ФГМ могут пользоваться геологи и другие специалисты, совместно работающие над дальнейшим исследованием этого важного рудного района..

Совмещение наборов данных

Разнообразие данных, используемых при разведке рудных месторождений, требует понятной 3D визуализации. По этому пути уже прошли другие отрасли, например поиски и разведка на нефть и газ или прогноз погоды; сложилась практика применения интегрированных компьютерных систем, создающих рабочую среду в виртуальной реальности (Stark et al., 2000). В таких системах размером с комнату («пещерных»), интерпретатор, по сути, лично входит внутрь среды и использует специальные средства управления и шлем виртуальной реальности для доступа к окружающим его данным. Из-за высокой стоимости и сложности эти системы станут считать основным инструментом разведки руд лишь через несколько лет.

Пока же, разрабатываются средства совмещения наборов данных на имеющихся более дешевых системах (Treinish, 2001). В этом подходе для одновременного показа значительного количества разнородных данных используются цвета и прозрачные

слои в разумном сочетании с изолиниями и векторами. В качестве примера 3D визуализации на рис. 2 приведено изображение прогноза погоды на Гавайских островах. Для ориентировки на топосовне черным показаны контуры островов. Поле ожидаемой интенсивности осадков показано цветными изолиниями. Изоповерхность относительной влажности 90%, показанная полупрозрачным белым цветом, обозначает границы облачного слоя. Вертикальная скорость ветра показана в разрезе по острову Гавайи. В двух местах на карту наложены результаты моделирования ветрового поля. Длина и цвет стрелок и линий потока представляют горизонтальную скорость ветра. Оси профилей используются также для передачи (цветами) относительной влажности воздуха.

Подобные технологии совмещения данных могут быть почти без дополнительных затрат перенесены в рудную отрасль для отображения разнородных данных в трех измерениях и извлечения полезной информации из больших объемов данных, получаемых сегодня.

Вероятностный подход к решению обратной задачи

Зрение человека способно различать сложные элементы, например, выделять линии в сложных изображениях, и этот процесс зачастую сложно смоделировать на компьютере. Тем не менее, в разнородных данных, встречающихся в разведочной деятельности, обычно содержатся тонкие особенности, которые не удается выделить визуально. Поэтому визуализацию необходимо дополнить статистическими

средствами работы с данными и распознавания образов. В разведочной отрасли самым полезным средством статистического анализа является расчет условной вероятности наличия месторождения внутри некоторого района, при наличии некоторого набора данных по району. Этую величину обозначают $P(D|x)$, где D – гипотеза о наличии месторождения, а x – вектор значений, соответствующих данным изысканий на месторождении или в его окрестностях.

Веса признаков

Прямыми способом оценки $P(D|x)$ является расчет с весами признаков. Это подход уже давно применяется в разведке минерального сырья для оценки вероятности наличия целевого объекта (Bonham-Carter, 1994; Raines, 1999; Mihalasky, 1999; Kemp et al., 2001). Обычно используется процедура расчета через логарифмические остатки, откуда, собственно, и появилось название метода. Сила метода в его простоте. При некоторых допущениях вероятность $P(D|x)$ может быть оценена непосредственно по набору данных путем подсчета и простых арифметических действий. Эти допущения таковы: (1) данные условно независимы, и (2) каждая компонента вектора параметров может принимать лишь конечное (небольшое) число значений. Реальные данные, заметим, лишь весьма приблизительно удовлетворяют этим предположениям. Несмотря на это ограничение, тщательный анализ с использованием весов признаков может дать полезные результаты, расширяющие понимание вопроса. Примером является блестящая диссертация Михальского (Mihalasky, 1999).

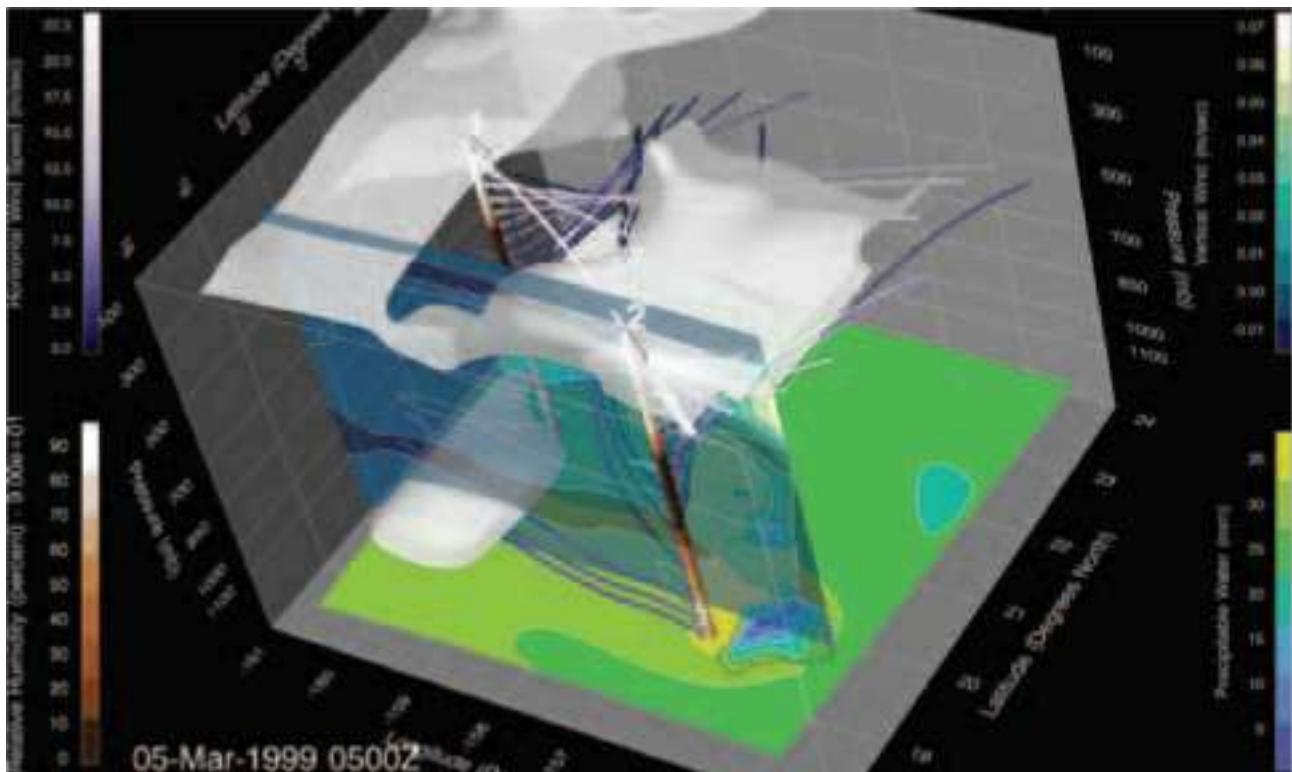


Рис. 2. Совмещение разнородных данных. Вид с высоты на погодную схему Гавайских островов. На одном рисунке показаны давление, относительная влажность, объем осадков, горизонтальная и вертикальная скорость ветра. С разрешения Ллойда Трейниши (Lloyd Treinish), Исследовательский центр IBM им. Томаса Дж. Ватсона.

Рудная геология и геофизика

Существуют и другие методы, не подверженные этим ограничениям, и, поэтому, способные извлекать из данных больше информации.

Аппроксимация функциями

Одним из представлений $P(D|x)$ является наилучшая в смысле наименьших квадратов аппроксимация D как функции вектора параметров x . Это значит, что за $P(D|x)$ принимается такая функция $f(x)$, которая минимизирует значение нормы $\|D - f(x)\|^2$, где D рассматривается как случайная величина, равная 1 при наличии месторождения и 0 – при его отсутствии. Такая функция, разумеется, существует, но в общем случае она нелинейна относительно x . Ясно, что о виде этой функции заранее сказать ничего нельзя.

Для решения таких задач есть ряд современных математических методов, например, опережающие нейронные сети (feed-forward neural networks), сети на основе радиальных базисных функций (radial basis function networks), или более общие представления через ядро. Нейронные сети используются для обнаружения объектов и оценки перспективности (см. Poulton et al., 1992; Brown et al., 2000, 2003) и Bougrain et al., 2003). Применение сетей на основе радиальных базисных функций описано в работах (Singer and Kouda, 1997, 1999), (Harris and Pan, 1999), (Harris et al., 2003) и (Porwal et al., 2003).

Подбор параметров

Независимо от вида выбранная для использования функция, скажем, $f(x, w)$, будет зависеть от параметров w , которые должны соответствовать данным. Задача состоит в выборе таких w , которые сделают $f(x, w)$ хорошей аппроксимацией $P(D|x)$ при близких к реальным значениях параметров модели x . Этого можно достичь, выбирая w так, чтобы минимизировать значение нормы $\|D - f(x, w)\|^2$, хотя возможны и другие функции ошибки (Barnett and Williams, 2006).

Обычно оценку нормы минимизируют, используя обучающее множество с известными x и D . Допустим, мы располагаем набором из N известных пар $(x_1, d_1), \dots, (x_N, d_N)$, где каждое x_i – наблюденное значение на участке $d_i = 1$ или 0, в зависимости от наличия или отсутствия месторождения. Тогда модель можно подобрать, выбирая w так, чтобы минимизировать выражение $\frac{1}{N} \sum_{i=1}^N |d_i - y_i|^2$, где $y_i = f(x_i, w)$.

Оценка мощности

Минимизация невязки не гарантирует, что вне обучающего множества ошибки также будут малы, если только не ограничить как-либо сложность модели. По сути, центральной идеей теории обучения является наложение ограничений на величину невязки в терминах мощности класса аппроксимирующих функций (Vapnik, 1998; Poggio and Smale, 2003).

first break том 24, июль 2006

Слишком хороший подбор на обучающем множестве может уменьшить пригодность в рабочем случае. Это особенно верно для разведочной отрасли, где известных месторождений немного, а количество данных велико (van der Baan and Jutten, 2000). При анализе этот факт всегда следует тщательно учитывать. В рассмотренном ниже полевом примере используется регуляризация по Байесу, описанная в работе (Williams, 1995).

Другие методы

Существуют и нестатистические методы компьютерного картирования месторождений, в том числе использующие методы математической и нечеткой логики (Bonham-Carter, 1994). Полезный обзор и анализ некоторых общих вопросов и особенностей разных методов содержится в работе (Groves et al., 2000), где также приведен полевой пример использования специально разработанного подхода на основе векторной нечеткой логики (Knox-Robinson, 2000), дающего непрерывную меру перспективности и оценку достоверности результатов.

В остальной части этой статьи показаны преимущества полностью статистического подхода, использующие нейронные сети.

Полевой пример: Уокер Лейн

Разломная зона Уокер Лейн, находящаяся на границе Невады и Калифорнии на западе США, изучается и разрабатывается уже давно, с открытия в конце 1850-х гг. знаменитой жилы Комсток (Comstock) (рис. 3,4). Этот освоенный район является хорошим примером и образцом для проверки процедур работы с данными, в том числе – по процедуре, описанной в этой работе.



Рис. 3. Карта западных штатов США. Белым показаны границы участка работ по Уокер Лейну

Не так известная, как соседний район Карлин (Carlin), и, безусловно, не столь богатая, зона Уокер Лейн примечательна многочисленными эптермальными месторождениями золота и серебра, связанными с вулканами. Помимо жилы Комсток, давшей в свое время 200 млн. унций серебра и 9 млн. унций золота, и месторождения Круглая гора (Round Mountain) с разведанными запасами в 14 млн. унций золота, здесь имеется еще как минимум 10 месторождений с запасами более 1 млн. унций золота каждое (рис. 4, табл. 1). С учетом всех меньших (с запасами менее 1 млн. унций) общий объем разведанных запасов золота в рассматриваемой зоне составляет около 50 млн. унций.

Большинство этих месторождений находится в разломной зоне Уокер Лейн – полосой шириной 100 км, простирающейся с северо-запада на юго-восток от Рено до Лас-Вегаса. Эта система сбросо-сдвигов содержит серию глубоких правосторонних сдвигов и оперяющих секущих разломов, по которым, предположительно, с глубины поступают магма и термальные воды. Считается, что разломная зона Уокер Лейн связана с мегаразломом Мохав-Сонора (Mojave-Sonora) и существует с мезозойского времени (Stewart, 1988).

На юго-западе зона Уокер Лейн ограничена поднятием гор Сьерра-Невада, а на северо-востоке обширным массивом Хребтов и Долин (Basin and Range terrain). В обнажениях этого района встречаются в основном третичные вулканиты, а также мезозойские интрузии в основном монzonитового и гранодиоритового состава. В основании разреза лежат сильно смятые в складки и разбитые разломами палеозойские осадки. Эти породы представлены в обнажениях на юго-востоке зоны в районе Лас-Вегаса, где в них содержится значительная доля карбонатов.



Рис. 4. Золотые месторождения зоны Уокер Лейн. Площадь шестиугольника – 100000 кв. км. Кроме обозначенных месторождений, в зоне имеется еще около 140 рудопроявлений золота с запасами более 50000 унций. Серым цветом показан испытательный полигон и прилегающие территории, выведенные из хозяйственного оборота

Результаты полевых работ

Для систематического статистического исследования рудоперспективного района нужны два условия: во-первых, достаточное количество известных месторождений для создания обучающего множества, и, во-вторых, достаточное количество современных данных. Зона Уокер Лейн отвечает обоим требованиям

Табл. 1. Золотые месторождения зоны Уокер Лейн с запасами св. 1 млн. унций.
По Р Уорленду (R. Worland), 2003, личная беседа.

Месторождение	Запасы Au (млн. унций)	Содержание Au (г/т)	Добыча Au (млн. унций)	Добыча Au (млн. унций)
Round Mountain	14.1	0.69	6.4	5.1
Comstock Lode	8.6	41.02	8.4	193.0
Northumberland	5.0	3.73	n/a	n/a
Goldfield	4.6	26.63	4.2	1.5
Bullfrog	2.5	8.26	1.9	2.1
Borealis	2.1	0.66	0.6	0.1
Tonopah	2.0	5.61	1.9	174.1
Aurora	1.8	2.33	1.8	20.6
Paradise Peak	1.6	2.07	1.6	24.0
Rawhide	1.5	1.35	1.1	8.9
Bodie	1.5	31.10	1.5	7.3
Daisy	1.5	1.18	n/a	n/a
Totals	46.7	1.56	29.4	436.7

По : R. Worland, личная беседа, 2003.

© 2006 EAGE

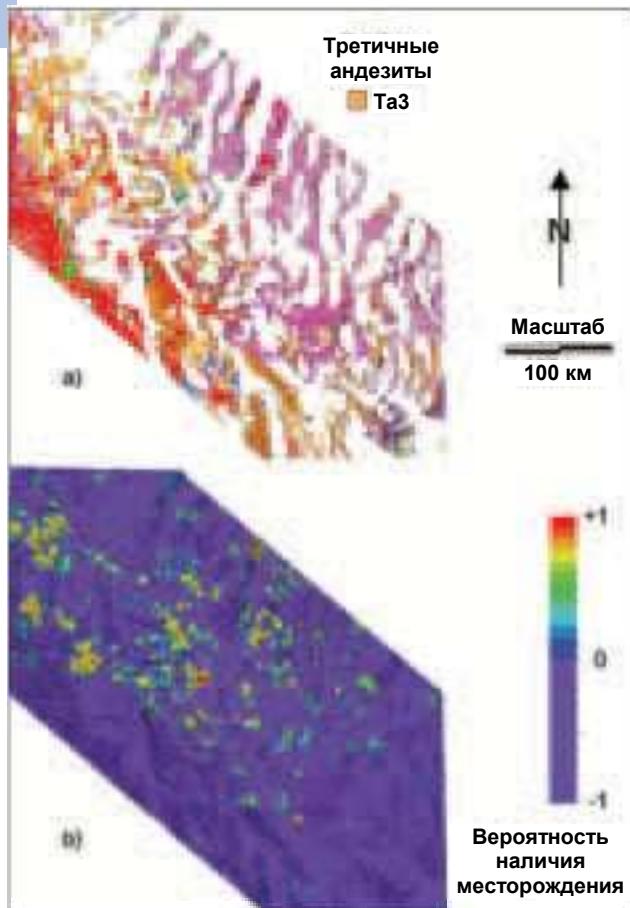


Рис. 5. a) Геологическая карта зоны Уокер Лейн. Области развития современных покровных отложений не закрашены. b) Карта вероятности наличия месторождения по результатам анализа геологической карты. Объекты наложены на рельефную карту местности, на которой белым обозначены известные месторождения. Отметим корреляцию формации третичных андезитов Та3 и объектов с высокой вероятностью (красных).

Известно много золотых месторождений. На протяжении XX в. Геологическая служба США (US Geological Survey), Бюро горного дела и геологии штата Невада (Nevada Bureau of Mines and Geology), Геологическая служба штата Калифорния (California Geological Survey), и другие организации провели там многочисленные работы. Таким образом, накоплен большой объем данных, которые просто ждут осмыслиения.

Аналитический подход к интеграции данных, очевидно, весьма зависит от качества и полноты данных. При отборе данных и включении их в процесс нужно быть очень внимательным. Распространенной ошибкой является неосмотрительное совмещение данных в разных системах координат. Например, на западе Невады расхождение систем координат WGS84 и NAD27 составляет около 250 м, что может серьезно повлиять на результаты. Для хранения и визуализации данных весьма полезной может оказаться ГИС, но для статистического представления необходим также доступ к самим данным, например, для расчета производных.

При исследовании зоны Уокер Лейн использованы следующие наборы данных: региональная геология, 48

космоснимки (Landsat thematic mapper, TM), геохимия русловых осадков, изостатические аномалии силы тяжести, данные аэромагниторазведки и аэрорадиометрии, цифровая модель местности, региональные структуры. Большая часть этих данных находится в открытом доступе, например в Национальном центре геофизических данных (National Geophysical Data Center). Исключением являются данные о региональных структурах, которые, как сказано ниже, получены авторами путем рассуждений.

Рассматриваемая зона имеет размер 500 x 200 км, т.е. ее площадь составляет 100000 кв. км. Сначала нужно было выбрать размер ячейки сети, по которой будет проводиться работа с данными. В нашем случае, с учетом качества и масштаба имеющихся данных, размер ячейки был принят 250 x 250 м. Это размер является компромиссным между космоснимками с высоким (28.5 м) разрешением и меньшей детальностью других данных. В других случаях можно использовать более крупные или более мелкие ячейки.

Самым важным слоем [в терминах ГИС – Прим. перев.] является слой известных месторождений, поскольку по его данным проводится обучение всей сети работе с другими данными. В этот слой включены все известные месторождения с запасами более 50000 унций, установленные путем дешифрирования аэрофотоснимков, анализа публикаций и полевых выходов с привязкой по GPS. Было выявлено около 150 отдельных месторождений. На рис. 4. традиционным знаком «скрещенные молотки» показаны положения некоторых крупных месторождений зоны Уокер Лейн. В ходе работы по возможности использовались заверенные месторождения. В качестве региональной геоподосновы использованы листы геологической карты штата Невада (M1:500000, Stewart and Carlson, 1978) и геологической карты штата Калифорния (M1:750,000, Jennings, 1985). Эти карты представляются в цифровом виде соответствующими службами штатов. Поскольку карта Невады покрывает большую часть района, обе карты приведены к системе обозначений карты Невады.

При отборе космоснимков использовались спутниковые данные с начала 1980-х гг., что позволило учесть все искажения, связанные с горными работами за последние 20 лет. Данные по девяти перерывающимся участкам были тщательно объединены в единообразную мозаику, покрывающую всю зону Уокер Лейн. В статистическом анализе использовались все шесть диапазонов высокого разрешения. Геохимические данные предоставлены Национальной программой оценки запасов урана (National Uranium Resource Evaluation, NURE), осуществленной Министерством энергетики США (US Department of Energy) в 1973-1984 гг. (Smith, 2001). Данные получены несколькими партиями, использовавшими разные методы сбора и обработки данных в четырехугольниках $1^\circ \times 2^\circ$ на территории западных штатов США. Поэтому, прежде всего, нужно было нормализовать данные по десяти четырехугольникам, покрывающим зону Уокер Лейн. Это удалось сделать с удовлетворительным качеством и получить полезные карты на весь район.

В исследовании использовались не только такие общепринятые индикаторы золота как мышьяк и ртуть, поскольку съемка по этим элементам была неполной. Дополнительно привлекались данные еще по 11 элементам: алюминию, барнию, железу, калию, магнию, марганцу, натрию, скандию, торию, титану и ванадию. Хотя некоторые из них малоподвижны, допускалось, что внедрение золотоносной системы должно было влиять на весь массив вмещающих пород целиком.

По работам NURE получены также данные аэродиагностики. Они проводились на больших площадях с большой высоты при поисках месторождений урана, поэтому качество данных не соответствовало нуждам работ на золото. Современные данные с высокой детальностью были бы гораздо более полезны для исследования литологии и изменчивости молодых вулканитов зоны Уокер Лейн. Однако на всю площадь исследования имелись только данные NURE.

Данные региональной грави- и магниторазведки довольно высокого качества были получены в Геологической службе США. Данные магниторазведки были получены за многие годы в ходе работ с различным расстоянием между маршрутами и на разных высотах. Они были приведены к одной высоте и сведены в единую карту магнитного поля зоны Уокер Лейн. Данные гравиразведки, также полученные в ходе различных работ за долгое время, были сведены в единую карту изостатических аномалий поля силы тяжести в редукции Буге. Поправка Буге компенсирует влияние промежуточного слоя, а изостатическая поправка – влияние компенсационных выступов под максимумами рельефа. Цифровая модель местности хорошего качества изначально создавалась для расчета поправки за рельеф при обработке данных гравиразведки. В этой модели отражаются также локальные структуры – направления хребтов и водотоков, а также кольцевые структуры. Поэтому эти данные самостоятельно участвуют в процедуре анализа в нейронной сети.

Завершает набор данных структурная карта района. Это единственный в своем роде набор данных (принадлежащий лично г-ну Ньюмонту (Newmont), составленный путем анализа геологических данных, космоснимков и данных гравимагниторазведки. Анализ включал не только прослеживание линейных структур, но и геодинамическую интерпретацию с учетом региональной тектоники, например надвиг массива Хребтов и Долин, подпирающий эффект батолита Сьерра-Невада, и сбросо-сдвиговые движения в разломе Сар-Андреас (B. Davies, 2003, разговор).

Все виды этих данных имели свои достоинства и недостатки. Геологи увидят, что геология в масштабах штата слишком генерализована и содержит ошибки картирования; геофизики заметят потерю данных гравиразведки из-за переменного шага по профилю; геохимиков смутят разнородные процедуры сбора и анализа образцов русловых отложений. Эти сомнения обоснованы. Разумеется, чем лучше данные, тем лучше результат анализа. Тем не менее, это типичные данные региональных исследований, с которыми мы умеем работать вручную. Более того, статистические связи

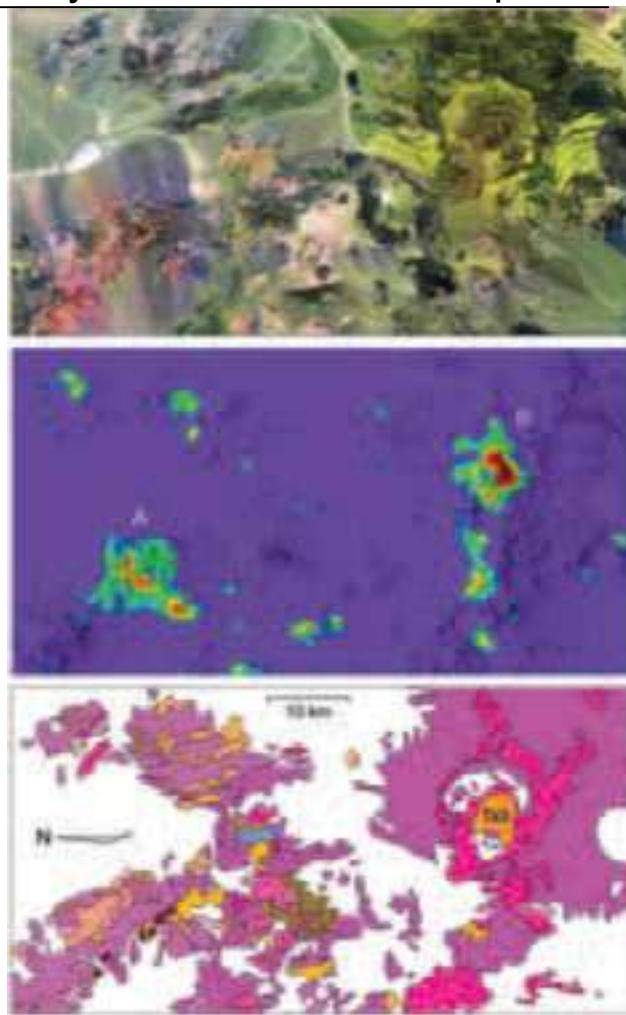


Рис. 6. Окончательные результаты по части площади испытательного полигона – вымороочной территории в зоне Уокер Лейн. Карта вероятности наличия месторождения (в середине) в сопоставлении с космоснимками в диапазоне 571 (вверху) и геологической картой (внизу). Объекты наложены на рельефную карту местности. Вблизи объекта В отметим кольцевую структуру, выделяемую как на космоснимке, так и на геологической карте, а также выход третичных андезитов Та3.

между разнородными данными повышают достоверность и позволяют компенсировать отдельные недостатки одних данных достоинствами других.

Представление данных

Описанные выше данные организованы в 25 первичных слоев, что нормально для современного набора данных. Из них 6 слоев содержат космоснимки, 11 – данные геохимии, 3 – радиометрии. Однако в отдельных отсчетах геофизических или, скажем, Топографических данных большого смысла нет. Нет причин полагать, например, что все золото находится на одной и той же высотной отметке, скажем, 3000 м. То же можно сказать о значениях силы тяжести. Существенным является место точки наблюдения в общей картине поля. Статистически значимыми могут быть, например, борта максимумов силы тяжести. Но нет



Рис. 7. Справа – распределение рангов вероятности для известных месторождений золота, слева – для остальной части района.

необходимости предупреждать об этом нейронную сеть, достаточно обеспечить ее необходимым для принятия решений количеством информации. Такие особенности можно выразить через производные и интегралы исходных данных, гравиметрических, например, и включить их в анализируемый набор данных наравне с первичными. Эта процедура сродни разложению в ряд Тейлора с удержанием нескольких первых членов. Для полноты картины следует использовать первые и вторые производные, горизонтальные, вертикальные и смешанные. Таким образом, для каждого первичного слоя образуется около десяти вторичных.

Данные геологии также требуют особого подхода. После объединения геологических карт Невады и Калифорнии в зоне Уокер Лейн выделяется 69 формаций, то есть на вход аналитической процедуры в нейронную сеть нужно подать 69 наборов данных. Следует также учитывать, что соседние формации и геологические границы также являются важными деталями геологического строения. Поэтому геология в районе каждой точки описывалась в скользящем окне, в котором подсчитывалась доля каждой формации. После ряда проб размер окна был принят 5 x 5 км.

Структурная карта также является особым случаем. Процедура введения буферной зоны, применяемая в ГИС-технологиях, оказалась неадекватной. Поэтому воспользовались приемом из геофизики – разломным зонам были присвоены произвольные плотности, после чего на всей площади было рассчитано суммарное поле силы тяжести. Поскольку информации об углах падения не было, все выделенные структуры считались вертикальными. В результате получилось плавно меняющееся поле, в котором хорошо видны пересечения, а эффект отдельных структур убывает с расстоянием по простому ($1/r$) закону. Более того, это непрерывное поле можно дифференцировать и интегрировать, как результаты измерений потенциального поля.

В связи с такой предварительной обработкой количество входных наборов данных быстро растет и превышает количество исходных наборов данных. В исследовании на

Уокер Лейн общее количество входных наборов данных составило 139. Поэтому потребовалось применять специальные процедуры регуляризации, описанные выше в разделе про вероятностное решение обратной задачи, и процедуры оптимизации, подобные описанным в (Williams, 1991).

Результаты анализа в нейронной сети

Перед полномасштабным анализом всего набора данных полезно рассмотреть отдельные наборы данных и убедиться в том, что статистические процедуры работают правильно и дают осмысленные результаты. Это позволит также установить, какие наборы данных дают наибольший вклад в итоговую карту. На рис. 5 представлена карта вероятности наличия месторождения по данным только геологического слоя. Как уже говорилось, по геологической карте зоны Уокер Лейн, приведенной в верхней части рисунка, выделено 69 формаций, использованных в процедуре анализа. На этом рисунке не приведены молодые породы, перекрывающие золотоносные породы и развитые в большинстве долин. Но в процедуре анализа эти нераскрашенные области участвуют по умолчанию.

В нижней части рисунка приведена карта вероятности наличия месторождения, наложенная на рельефную карту поверхности, на которой видны долины и хребты. Для наглядности области с положительной вероятностью даны более теплыми цветами от синего до красного, а с отрицательной (там, где апостериорная вероятность ниже априорной) – лиловым.

Очевидно, что только по геологическим данным перспективной оказывается лишь малая часть района. Сеть ясно показала, что ни в пределах батолита Сьерра-Невада, ни в осадочных породах палеозойской платформы искать связанные с вулканами месторождения золота не следует. Это может быть, очевидно для геолога, который, несомненно, интуитивно исключит эти области из рассмотрения, но сеть должна была прийти к такому выводу, опираясь только на входные данные из обучающего множества. Быстрое сравнение с геологической картой позволяет утверждать, что наилучшие прогнозы связаны с формацией Та3. Эта формация, представленная третичными андезитами, хорошо известна геологам Невады, как вмещающая порода многих крупных месторождений золота, связанных с вулканами (Mihalasky, 1999 и др.). При более тщательном рассмотрении в крупном масштабе выявляются и другие формации и геологические границы, с которым связана большая часть золотых месторождений.

Все наборы данных, которых слишком много, чтобы привести их все, дали близкие, но не обязательно совпадающие результаты. Как и ожидалось, карта, построенная только по космоснимкам, имеет высокую корреляцию с аномалиями ТМ, связанными обычно с высоким содержанием сульфидов, как на месторождении Голдфилд (Goldfield), шт. Невада. По результатам геохимических съемок получается четкая карта, но объяснить ее интуитивно не получается, поскольку в ее основе лежат карты содержания 11 элементов. Тем не менее, при рассмотрении содержания элементов по отдельности вблизи месторождений видны

ранее известные закономерности, например понижение содержания натрия и повышение содержания калия.

Более интересные результаты получаются при комбинировании наборов данных. Поскольку за право разведки в зоне Уокер Лейн идет конкуренция, окончательную общую карту месторождений мы решили не публиковать. Можно, тем не менее, показать часть карты, относящуюся к испытательному полигону в Неваде. Неудивительно, что после 40 лет ядерных испытаний эта часть зоны Уокер Лейн не представляет интереса для золотодобытчиков. На рис. 6 представлены результаты по участку площадью 2500 кв. км, расположенному примерно в 50 км к юго-востоку от рудного района Голдфилд, в северо-западном углу полигона. Карта вероятности месторождений построена по всем наборам данных, кроме геохимии NURE, которая не проводила работы на полигоне. Цветовая шкала этой карты такая же, как на рис. 5. Для сравнения в том же масштабе приведены космоснимки и геологическая карта.

В результате работы выделено два интересных объекта, которые стоило бы обследовать при наличии доступа к территории. Объект А соответствует аномалии на спутниковых данных, связанной с изменением минерализации. Объект В соответствует кольцевой структуре, определяемой по космоснимкам и по геологии. Дополнительный интерес вызывает маленький оранжевый объект, расположенный, возможно, в центре кальдеры. Это уже упоминавшаяся формация третичных андезитов *Ta3*. Отметим также, что объекты выделены по всей совокупности данных, включая грави- и магниторазведку и структурную карту, а не только по космоснимкам и геологии.

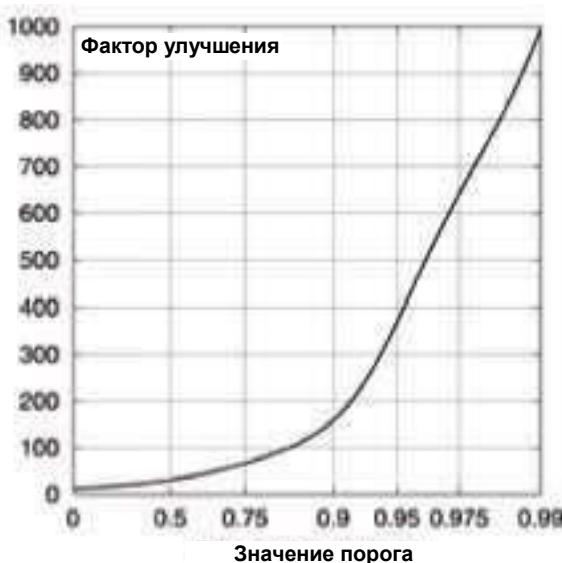


Рис. 8. Зависимость фактора улучшения от значения порогового ранга

Эти два участка с высокой вероятностью наличия месторождения достаточно малы, и, если бы планировалась разработка, их можно было бы обследовать за несколько дней с отбором образцов. На этом этапе можно было бы провести детальное геологическое картирование, геофизические работы, геохимическое опробование и рекомендовать места для закладки скважин. Стоит заметить, что площадь объектов по изолинии вероятности 0.5 не превышает 2.5% показанной площади. Такое соотношение площадей отмечается и для зоны Уокер Лейн в целом. Таким образом, площадь, подлежащая обследованию, значительно уменьшается, а вероятность обнаружения месторождения, как показано ниже, значительно повышается.

Экономический анализ

Результаты статистического анализа можно рассмотреть также с экономической точки зрения, обсуждая издержки и выгоды программы работ, и прибавочную стоимость точного выделения объектов (Mackenzie and Woodall, 1987; Green, 2004).

Вероятность успеха

Пусть возможные объекты обследования ранжированы в соответствии с ранговой функцией S , а программа изысканий состоит в обследовании N объектов, ранг которых превосходит заданный порог s . Во-первых, какова вероятность того, что определенный объект из числа этих N является промышленным месторождением? Ответ зависит от значения s и от того, насколько хорошо определяется ранг по S . В идеале S должна четко отделять месторождения от пустых областей, хотя на практике эти решения несколько перекрываются. На рис. 7 показано распределение рангов, полученное в ходе работ по зоне Уокер Лейн. Правое распределение построено только по известным месторождениям золота. Левое распределение включает все месторождения, в том числе – неоткрытые, сосредоточенные в правой части распределения. Левая часть правого распределения включает известные месторождения, которые нейронная сеть сочла нетипичными. Распределения с рис. 7 дают оценку условной вероятности $P(D|S>s)$ того, что объект с рангом выше s является месторождением. По правилу Байеса она выражается, как

$$P(D|S>s) = \frac{P(S>s|D)}{P(S>s)} P(D)$$

где $P(D)$ обозначает вероятность обнаружения месторождения при случайном поиске. Эта величина (априорная вероятность) домножается на фактор улучшения, который показывает, как возрастает вероятность обнаружения месторождения при обследовании только объектов с высоким рангом и дает установить апостериорную вероятность $P(D|S>s)$.

Для вычисления фактора улучшения из уравнения (1) нужно определить его числитель и знаменатель. Знаменатель $P(S > s)$ относится к распределению рангов по всей площади работ, то есть площадь правее s под левой кривой на рис. 7. Числитель $P(S > s|D)$ относится к объектам, которые нужно обследовать. К сожалению, он по определению неизвестен. Но, если допустить, что выборка известных месторождений репрезентативна для неизвестных месторождений, то для оценки распределения рангов неизвестных месторождений можно использовать распределение рангов известных месторождений, то есть правую кривую с рис. 7. Таким образом, фактор улучшения как функция порогового ранга s (рис. 8), оценивается как площадь правее s под обеими кривыми с рис. 7.

Целью разведки является обнаружение промышленных месторождений. В работах по зоне Уокер Лейн сеть была обучена находить месторождения с запасами более 50000 унций. Только часть из них будет промышленными. Чтобы вычислить вероятность того, точка с высоким рангом соответствует промышленному месторождению, нужно вероятность того, что это месторождение, на вероятность того, что месторождение с высоким рангом является промышленным. Если через f обозначить наличие промышленного месторождения, то из уравнения (1) следует, что $P(E|S > s)$ выражается так:

$$P(E|D \& S > s) \frac{P(S > s|D)}{P(S > s)} P(D) \quad (2)$$

Это значит, что вероятность наличия промышленного месторождения есть произведение трех множителей: (1) вероятности того, что месторождение с высоким рангом является промышленным, (2) фактора улучшения и (3) априорной вероятности. Например, если априорная вероятность равна 1/2500, а условная вероятность наличия промышленного месторождения равна 1/2, то их перемножение даст 1/5000. При случайном выборе объекта из числа тех, чей ранг превышает 0.95 (рис. 8), фактор улучшения равен примерно 375, и вероятность обнаружения промышленного месторождения будет примерно $375/5000 = 0.075$, то есть более 1/14.

Изыскания на нескольких объектах

Программа изысканий обычно включает в себя работы на более чем одном объекте. Вероятность того, что хотя бы один из них окажется промышленным месторождением, будет больше вероятности того, что таким месторождением окажется конкретный объект. Пусть выявлено N объектов с рангом больше s , и на них в каком-то порядке проводятся изыскания до тех пор, пока промышленное месторождение не будет найдено, или все N объектов не будут обследованы. Если вероятность того, что в данном месте имеется промышленное месторождение, есть $p = P(E|S > s)$ и все

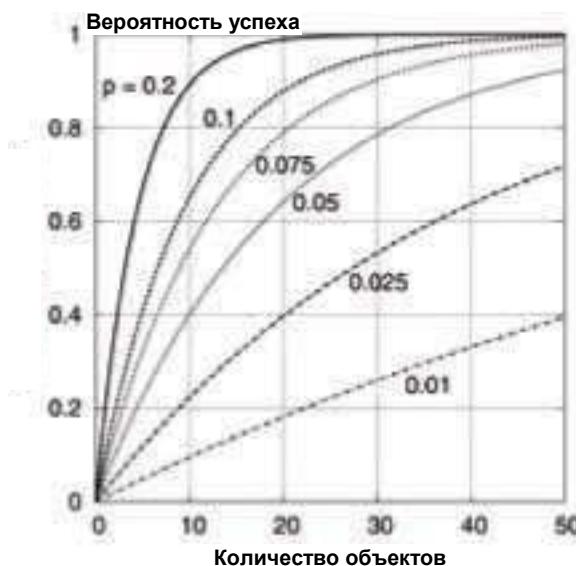


Рис. 9. Зависимости вероятности успеха программы изысканий от числа исследуемых объектов при различных значениях p – вероятности обнаружения месторождения на отдельном объекте. объекты независимы, то общая вероятность успеха будет

$$p_N = 1 - (1 - p)^N \quad (3)$$

а оценка количества объектов, которые нужно обследовать будет p_N/p . Зависимость вероятности успеха от числа обследуемых объектов при различных p приведена на рис. 9. В предположениях примера из конца предыдущего раздела $p = 0.075$ соответствует пороговому рангу $s = 0.95$.

Может показаться, что максимальная вероятность успеха достигается при максимально возможных значениях s , которое растет с ростом p . Но при слишком больших s никаких объектов вообще не будет обнаружено. При оценке количества объектов при данном пороговом значении следует принять некоторые допущения. В нашем случае объекты должны находиться достаточно далеко один от другого, чтобы выполнялось уравнение (3), для чего благоприятные и неблагоприятные исходы на каждом из объектов должны быть независимы. В табл. 2 приведено количество объектов в зоне Уокер Лейн в целом при различных значениях порога s и различных минимальных расстояниях между объектами. Например, в колонке «2 км» дано количество объектов, отстоящих один от другого и от известных месторождений более чем на 2 км. Это значит, что если обследовать район в целом, то при пороге менее 0.975 программа изысканий будет ограничена скорее бюджетом, чем количеством объектов поиска. На практике отбор целей производится с привлечением дополнительной информации, наряду со значениями рангов. Тем не менее, приведенная процедура анализа дает ожидаемый уровень вероятности.

Прибыли и издержки

Углубить анализ можно, располагая оценками затрат на разведку. Пусть затраты на выяснение, является ли объект месторождением, составляют C_D , а дополнительные затраты на выяснение, является ли выявленное месторождение промышленным – C_E . Пусть также ожидаемая стоимость запасов месторождения составляет V . Тогда стоимость изысканий на одном объекте оценивается как $C_D + qC_E$, где $q = P(D|S > s)$ – вероятность того, что объект с рангом выше порогового является месторождением. Для оценки стоимости всей программы изысканий эту сумму нужно умножить на ожидаемое количество объектов, которые придется обследовать p_N/p . Тогда, если считать, что обследуется не более N объектов, ожидаемая выгода равна $p_N V$ – ожидаемой стоимости запасов промышленного месторождения, умноженной на вероятность его обнаружения. Получается, что ожидаемая выгода превосходит ожидаемые издержки если

$$p > \frac{C_D + qC_E}{V} \quad (4)$$

и не зависит от N . Это значит, что вероятность того, что объект является промышленным месторождением, должна превышать отношение расчетных затрат на изыскания к расчетной стоимости запасов. В работах по Уокер Лейн вероятность p была больше 1/14 при пороговом значении 0.95. Это, по-видимому, существенно превышает отношение в правой части уравнения (4). Не составляет труда аналогично оценить риски такой программы изысканий, измеряя их в терминах изменений прибыли, но эта задача выходит за рамки данной статьи.

Выводы

Общий объем данных, получаемых при современных изысканиях, постоянно растет, что требует нового подхода к разведке минерального сырья. Этой задаче соответствуют новые методы работы с данными. В статье рассмотрены два различных подхода: визуализация и вероятностный подход к решению обратной задачи, которые, как считается, могут переломить тенденции как к снижению темпов прироста запасов, так и к росту затрат на открытие месторождения, отмечаемые в рудной отрасли в течение последних 20 лет.

Табл. 2. Зависимость количества объектов от минимального расстояния между ними при различных значениях порогового ранга.

Порог	Количество объектов		
	1 км	2 км	3 км
0.900	720	320	195
0.950	385	142	82
0.975	140	45	29
0.990	24	10	4

Эти новые статистические методы продемонстрированы на примере работ по Неваде. Располагая 25 слоями данных изысканий, процедура, реализованная в нейронной сети, дала хорошо разрешенную карту вероятность наличия месторождения. Наиболее убедительным доказательством эффективности этих методов, было бы, разумеется, открытие крупного месторождения. В настоящее время ведутся детальные изыскания на объектах, и время покажет, так ли это.

Рассмотрена также вероятность открытия этими методами промышленного месторождения. Теоретически показано, что системная интеграция всех результатов изысканий значительно повышает шансы на открытие. В золотоносном районе Уокер Лейн, например, априорная (без изысканий) вероятность открытия принята равной 1/2500. Улучшение прогноза после применения нейронной сети выражается множителем 375. Это значит, что апостериорная вероятность составляет 0.075 для отдельного объекта и 0.8 при обследовании 20 объектов.

Пример относится к большому (100000 кв. км) перспективному рудному району. В таких масштабах третья координата (глубина) не важна, поэтому исследование было в основном двумерным. Поскольку в процедуре использовались не данные по точкам, а просто наборы данных, связанные с эталонными объектами, естественным шагом развития будет переход к 3D ситуации. Метод должен работать в любом масштабе при наличии большого объема данных и возможности организовать хорошее обучающее множество. Другими словами его можно применять и при изучении строения известного месторождения, и при региональных поисках новых месторождений.

Будучи полезной в двумерных ситуациях, визуализация становится особенно важной в трехмерном случае, как для понимания данных так и для взаимодействия между специалистами. Авторы полагают, что сочетание вероятностного подхода к решению обратной задачи с наложением данных будет следующей вехой в развитии методов разведки, а их совместное применение снизит общие затраты на изыскания и в ближайшие годы приведет к росту числа открытий промышленных месторождений.

Благодарности

Авторы посвящают эту статью памяти Дэвида Марра (David Marr), своего близкого школьного друга и товарища по учебе в Кембридже. За время блестящей научной карьеры, безвременно оборвавшейся из-за лейкемии, Д.К.Марр заложил основы многих известных ныне методов визуализации и работы нейронных сетей.

Авторы благодарят также компанию Newmont Mining Corporation за поддержку на ранних стадиях разработки описанных методов и за разрешение на публикацию некоторых материалов по Уокер Лейн.

Литература

Barnett, C.T. and Williams, P.M. [2006] Mineral exploration using modern data mining techniques. In Doggett, M.E. and Parry, J.R. (Eds.) *Wealth Creation in the Minerals Industry: Integrating Science, Business and Education*. Society of Economic Geologists Special Publication 12, 295-310.

Bonham-Carter, G.F. [1994] *Geographic information systems for geoscientists: Modelling with GIS*: Tarrytown, New York. Pergamon Press-Elsevier Science Publications, 398.

Bougrain, L., Gonzalez, M., Bouchot, V., Cassard, D., Lips, A.L.W., Alexandre, F., and Stein, G. [2003] Knowledge recovery for continental-scale mineral exploration by neural networks. *Natural Resources Research*, 12, 3, p. 173-181.

Brown, W.M., Gedeon, T.D., Groves, D.I., and Barnes, R.G. [2000] Artificial neural networks: A new method for mineral prospectivity mapping. *Australian Journal of Earth Sciences*, 47, 757-770.

Brown, W.M., Gedeon, T.D., and Groves, D.I. [2003] Use of noise to augment training data: A neural network method of mineral-potential mapping in regions of limited known deposit examples. *Natural Resources Research*, 12, 2, 141-151.

Garrett, S., Griesbach, S., Johnson, D., Jones, D., Lo, M., Orr, W., and Sword, C. [1997] Earth model synthesis. *First Break*, 15, 1, 13-22.

Green, A. [2004] Exploration, risk and the value of geophysics. *Australian Society of Exploration Geophysicists, 17th Geophysical Conference and Exhibition*, Sydney, NSW, Extended Abstracts, 5.

Groves, D.I., Goldfarb, R.J., Knox-Robinson, C.M., Ojala, J., Gardoll, S., Yun, G.Y., and Holyland, P. [2000] Late kinematic timing of orogenic gold deposits and significance for computer-based exploration techniques with emphasis on the Yilgarn block, Western Australia. *Ore Geology Reviews*, 17, 1-38.

Harris, D. and Pan, G. [1999] Mineral favorability mapping: A comparison of artificial neural networks, logistic regression, and discriminant analysis. *Natural Resources Research*, 8, 2, 93-109.

Harris, D., Zurcher, L., Stanley, M., Marlow, J., and Pan, G. [2003] A comparative analysis of favorability mappings by weights of evidence, probabilistic neural networks, discriminant analysis, and logistic regression. *Natural Resources Research*, 12, 4, 241-255.

Jennings, C.W. [1985] An explanatory text to accompany the 1:750,000 scale fault and geologic maps of California. *California Division of Mines and Geology Bulletin 201*, Appendix D, 125-197.

Kemp, L.D., Bonham-Carter, G.F., Raines, G.L., and Looney, C.G. [2001] Arc-SDM: Arcview extension for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neural network analysis. <http://ntserv.gis.nrcan.gc.ca/sdm/>.

Knox-Robinson, C.M. [2000] Vectorial fuzzy logic: A novel technique for enhanced mineral prospectivity mapping, with reference to the orogenic gold mineralisation potential of the Kalgoorlie terrane, Western Australia. *Australian Journal of Earth Sciences*, 47, 5, 929-941.

Li, Y. and Oldenburg, D.W. [1996] 3-D inversion of magnetic data. *Geophysics*, 61, 2, 394-408.

Li, Y. and Oldenburg, D.W. [1998] 3-D inversion of gravity data. *Geophysics*, 63, 109-119.

Mackenzie, B. and Woodall, R. [1987] Mineral exploration economics: The search for base metals in Australia and Canada: Kingston, Ontario, Queen's University, *Centre for Resource Studies Working Paper 40*, 110.

McGaughey, W.J. and Vallee, M.A. [1998] Integrating geology and borehole geophysics in a common earth model for improved three-dimensional delineation of mineral deposits. *Exploration and Mining Geology*, 7, 51-62.

Mihalasky, M.J. [1999] Mineral potential modelling of gold and silver mineralization in the Nevada Great Basin—a GIS-based analysis using weights of evidence. *Unpublished Ph.D. dissertation, University of Ottawa*, 360. Also available as USGS Open-File Report 01-291.

Paterson, N. R. [2003] Geophysical developments and mine discoveries in the 20th century. *The Leading Edge*, 22, 6, 558-561.

Poggio, T. and Smale, S. [2003] The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, 50, 5, 537-544.

Porwal, A., Carranza, E.J.M., and Hale, M. [2003] Artificial neural networks for mineral-potential mapping: A case study from Aravalli province, western India. *Natural Resources Research*, 12, 3, 155-171.

Poulton, M.M., Sternberg, B.K., and Glass, C.E. [1992] Location of subsurface targets in geophysical data using neural networks. *Geophysics*, 57, 12, 1534-1544.

Raines, G.L., 1999, Evaluation of weights of evidence to predict epithermal-gold deposits in the Great Basin of the western United States. *Natural Resources Research*, 8, 4, 257-276..

Schodde, R.C. [2004] Discovery performance of the western world gold industry over the period 1985-2003. *PACRIM 2004, Proceedings of the 2004 Adelaide Congress of the Australian Institute of Mining and Metallurgy*, Adelaide, 14.

Singer, D.A. and Kouda, R. [1997] Use of a neural network to integrate geoscience information in the classification of mineral deposits and occurrences. In Gubins, A.G. (Ed.) *Proceedings of Exploration 97: Prospectors and Developers Association of Canada, Fourth Decennial International Conference on Mineral Exploration*, Toronto, 127-134.

Singer, D.A. and Kouda, R. [1999] A comparison of the weights-of-evidence method and probabilistic neural networks. *Natural Resources Research*, 8, 4, 287-298.

- Smith, S.M. [2001] National Geochemical Database: Reformatted data from the National Uranium Resource Evaluation (NURE) Hydrogeochemical and Stream Sediment Reconnaissance (HSSR) Program, Version 1.30, U.S. Geological Survey Open-File Report 97-492, *www release only*, <http://greenwood.cr.usgs.gov/pub/open-file-reports/ofr-97-0492/index.html>.
- Stark, T.J., Dorn, G.A., and Cole, M.J. [2000] ARCO and immersive environments, Part 2: Oil industry experience with immersive environments. *The Leading Edge*, 19, 8, 884-890.
- Stewart, J.H. [1988] Tectonics of the Walker Lane belt, Western Great Basin: Mesozoic and Cenozoic deformation in a zone of shear, in Ernst, W.G. (Ed.), *Metamorphism and Crustal Evolution of the Western United States*. Prentice Hall, Englewood Cliffs, New Jersey, Rubey, 7, 683-713.
- Stewart, J.H. and Carlson, J.E. [1978] *Geologic map of Nevada*. U.S. Geological Survey and Nevada Bureau of Mines and Geology Map, scale 1:500,000.
- Treinish, L.A. [2001] How can we build more effective weather visualizations? *Proceedings of the Eighth European Centre for Medium-Range Weather Forecasting (ECMWF) Workshop on Meteorological Operational Systems*, Reading, England, 90-99.
- van der Baan, M. and Jutten, C. [2000] Neural networks in geophysical applications. *Geophysics*, 65, 4, 1032-1047.
- Vapnik, V. [1998] *Statistical learning theory*. John Wiley, New York, 768.
- Williams, P.M. [1991] A Marquardt algorithm for choosing the step-size in backpropagation learning with conjugate gradients. *University of Sussex, Cognitive Science Research Paper CSRP 229*, 8.
- Williams, P.M. [1995] Bayesian regularization and pruning using a Laplace prior. *Neural Computation*, 7, 1, 117-143.
- Williams, P.M. [2002] Probabilistic learning models, in Corfield, D. and Williamson, J. (Eds.) *Foundations of Bayesianism*, Dordrecht, Kluwer, 117-134.