УДК 510.22

ИЗВЛЕЧЕНИЕ ЗНАНИЙ ДЛЯ ОЦЕНКИ КРЕДИТОСПОСОБНОСТИ: ПОДХОД ТЕОРИИ МУЛЬТИМНОЖЕСТВ *

А.Б.Петровский[1]

Мультимножество или множество с повторяющимися элементами служит удобной математической моделью для представления объектов, которые характеризуются многими разнородными (количественными и качественными), в том числе противоречивыми признаками. В работе предложен новый метод классификации таких многопризнаковых объектов как точек метрического пространства мультимножеств и использован для оценки кредитоспособности владельцев кредитных карт на основе анализа их финансового поведения.

Введение

Оплата товаров и услуг с помощью кредитных карт стало одной из обыденных и удобных возможностей современного мира. Вместе с тем банки и кредитные компании, выпускающие кредитные карты, ежегодно несут многомиллионные убытки из-за перерасхода денежных средств, допущенного владельцами кредитных карт. Поэтому оценка возможной кредитоспособности лица, желающего получить кредитную карту, является одной из актуальных и типичных задач в банковском деле. На практике для принятия решения о выдаче или отказе в выдаче кредитной карты желательно иметь одно или несколько достаточно простых решающих пра-

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (проекты 02-01-01077, 04-01-00290), программ фундаментальных исследований РАН «Математическое моделирование и интеллектуальные системы», «Фундаментальные основы информационных технологий и систем», гранта НШ1964.2003.1 Президента Российской Федерации по поддержке ведущих научных школ.

[1] *117312, Москва, проспект 60-летия Октября, 9, ИСА РАН, pab@isa.ru*

вил, которые позволяли бы отнести заявителя к соответствующей категории в зависимости от представленных им персональных данных.

К настоящему времени в базах данных кредитных организаций накоплены большие объемы информации о реальном финансовом поведении десятков и сотен тысяч владельцев кредитных карт. Каждый владелец карты описывается в базе данных набором многих разнородных признаков, включающих личные сведения (пол, возраст, образование, семейное положение, род занятий, доходы, место жительства и прочее) и финансовые показатели (баланс счета, получение и возврат кредитов, оплата покупок и услуг, выдача наличных и тому подобное). Анализ этой информации дает возможность выявить закономерности финансового поведения владельцев кредитных карт и построить их классификацию по платежеспособности, в простейшем случае разбить всех владельцев на две условные категории: «хорошие», не допускающие перерасхода средств или своевременно погашающие возникающие задолженности, и «плохие», совсем не возвращающие кредиты или возвращающие их с большой задержкой.

Большая размерность анализируемых массивов данных (десятки и сотни тысяч владельцев, несколько десятков признаков) и значительная разнородность признаков, включающих количественные и качественные (порядковые, номинальные) переменные, создает определенные трудности построения решающих правил для оценки кредитоспособности предполагаемых владельцев кредитных карт. Для решения задачи классификации таких многопризнаковых объектов разработан ряд методов, использующих линейную и логистическую регрессию, деревья решений, нейронные сети, линейное программирование, кластерный анализ [Shi *et al*, 2001]. Однако в основе этих подходов лежит тот или иной вид «обучения» или «настройки» алгоритмов классификации на некотором предварительно отобранном массиве данных. Вместе с тем, очевидно, что среди многих тысяч владельцев карт окажутся такие, которые имеют одинаковые наборы личных сведений, но разные финансовые показатели, то есть, относятся к разным категориям. Другими словами, эти объекты описываются противоречивыми признаками, что будет приводить к ошибкам при их классификации. Кроме того, такие объекты могут присутствовать и в выборочном массиве данных, что существенно снижает качество обучения классифицирующих алгоритмов.

Мультимножество или множество с повторяющимися элементами служит удобной математической моделью для описания объектов, которые характеризуются многими разнородными (количественными и качественными) признаками и могут существовать в нескольких экземплярах с отличающимися, в частности, противоречивыми значениями признаков, свертка которых или невозможна, или математически некорректна. Крат-

ность элементов – существенная особенность мультимножества, позволяющая отличать его от множества и рассматривать мультимножество как качественно новое математическое понятие. Множественность и повторяемость факторов, описывающих объекты, усложняет и затрудняет решение задачи их классификации.

В работе предложен новый метод классификации совокупности многопризнаковых объектов, основанный на их представлении в виде точек метрического пространства мультимножеств. Метод позволяет строить обобщенное решающее правило для сортировки объектов, выраженное через значения их признаков. Метод классификации может работать с реальными большими массивами объектов, имеющих любые наборы признаков, в том числе противоречивых, и не требует предварительной настройки на выборочном массиве.

1. Мультимножества и пространства мультимножеств

Мультимножеством $\boldsymbol{A}$, порожденным обычным множеством $U=\{x_1,x_2,\ldots\}$, все элементы которого различны, называется совокупность групп элементов вида $\boldsymbol{A}=\{k_A(x)\bullet x|x\in U,\ k_A(x)\in\mathbb{Z}_+\}$. Здесь $k_A:U\to\mathbb{Z}_+$ называется функцией числа экземпляров мультимножества, определяющей кратность вхождения элемента $x_i\in U$ в мультимножество $\boldsymbol{A}$, что обозначено символом •. Если $k_A(x)=\chi_A(x)$, где $\chi_A(x)=1$ при $x\in\boldsymbol{A}$ и $\chi_A(x)=0$ при $x\notin\boldsymbol{A}$, то мультимножество $\boldsymbol{A}$ становится обычным множеством. Если все мультимножества семейства $\mathsf{A}=\{\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots\}$ образуются из элементов множества G, то G называется доменом для семейства A, а множество $\mathrm{Supp}\boldsymbol{A}=\{x|x\in G,\ \chi_{\mathrm{Supp}A}(x)=\chi_A(x)\}$ – опорным множеством или носителем мультимножества $\boldsymbol{A}$. Мощность мультимножества $|\boldsymbol{A}|=\Sigma_x k_A(x)$ определяется как общее число экземпляров всех его элементов; размерность мультимножества $/\boldsymbol{A}/=\Sigma_x\chi_A(x)=|\mathrm{Supp}\boldsymbol{A}|$ – как общее число различных элементов. Мультимножество называется пустым $\varnothing$, если $k_\varnothing(x)=0$, и максимальным $\boldsymbol{Z}$, если $k_Z(x)=\max\limits_{A\in\mathsf{A}} k_A(x)$, $\forall x\in U$ [Петровский, 1995, 2003].

Введены следующие основные операции над мультимножествами:

объединение

$$\boldsymbol{A}\bigcup\boldsymbol{B}=\{k_{A\cup B}(x)\bullet x\mid k_{A\cup B}(x)=\max(k_A(x),\,k_B(x))\};$$

пересечение

$$\boldsymbol{A}\bigcap\boldsymbol{B}=\{k_{A\cap B}(x)\bullet x\mid k_{A\cap B}(x)=\min(k_A(x),\,k_B(x))\};$$

арифметическое сложение

$$\boldsymbol{A}+\boldsymbol{B}=\{k_{A+B}(x)\bullet x\mid k_{A+B}(x)=k_A(x)+k_B(x)\};$$

арифметическое вычитание

$$\boldsymbol{A}-\boldsymbol{B}=\{k_{A-B}(x)\bullet x\mid k_{A-B}(x)=k_A(x)-k_{A\cap B}(x)\};$$

симметрическая разность

$$\boldsymbol{A}\Delta\boldsymbol{B} = \{k_{A\Delta B}(x)\bullet x \mid k_{A\Delta B}(x)=|k_A(x)-k_B(x)|\};$$

дополнение

$$\overline{\boldsymbol{A}} = \boldsymbol{Z}-\boldsymbol{A} = \{ k_{\overline{A}}(x) \bullet x \mid k_{\overline{A}}(x) = k_Z(x)-k_A(x)\};$$

умножение на число

$$h\bullet\boldsymbol{A} = \{k_{h\bullet A}(x)\bullet x \mid k_{h\bullet A}(x)=h\cdot k_A(x),\ h\in\mathbb{Z}_+\};$$

арифметическое умножение

$$\boldsymbol{A}\bullet\boldsymbol{B} = \{k_{A\bullet B}(x)\bullet x \mid k_{A\bullet B}(x) = k_A(x)\cdot k_B(x)\};$$

арифметическая n-ая степень

$$\boldsymbol{A}^n = \{ k_{A^n}(x) \bullet x \mid k_{A^n}(x) = (k_A(x))^n\};$$

прямое произведение

$$\boldsymbol{A}\times\boldsymbol{B} = \{k_{A\times B}\bullet\langle x_i, x_j\rangle \mid k_{A\times B}=k_A(x_i)\cdot k_B(x_j),\ x_i\in\boldsymbol{A},\ x_j\in\boldsymbol{B}\};$$

прямая n-ая степень

$$(\times\boldsymbol{A})^n = \{ k_{(\times A)^n} \bullet\langle x_1,\ldots,x_n\rangle \mid k_{(\times A)^n} = \prod_{i=1}^{n} k_A(x_i)\,,\ x_i\in\boldsymbol{A}\}.$$

Носители операций над мультимножествами удовлетворяют следующим соотношениям:

$$\mathrm{Supp}(\boldsymbol{A}\cup\boldsymbol{B}) = \mathrm{Supp}(\boldsymbol{A}+\boldsymbol{B}) = (\mathrm{Supp}\boldsymbol{A})\cup(\mathrm{Supp}\boldsymbol{B});$$
$$\mathrm{Supp}(\boldsymbol{A}\cap\boldsymbol{B}) = \mathrm{Supp}(\boldsymbol{A}\bullet\boldsymbol{B}) = (\mathrm{Supp}\boldsymbol{A})\cap(\mathrm{Supp}\boldsymbol{B});$$
$$\mathrm{Supp}(\boldsymbol{A}\Delta\boldsymbol{B}) = (\mathrm{Supp}(\boldsymbol{A}-\boldsymbol{B}))\cup(\mathrm{Supp}(\boldsymbol{B}-\boldsymbol{A}));$$
$$\mathrm{Supp}(h\bullet\boldsymbol{A}) = \mathrm{Supp}(\boldsymbol{A}^n) = \mathrm{Supp}\boldsymbol{A};$$
$$\mathrm{Supp}(\boldsymbol{A}\times\boldsymbol{B}) = (\mathrm{Supp}\boldsymbol{A})\times(\mathrm{Supp}\boldsymbol{B}).$$

Семейство мультимножеств, замкнутое относительно операций объединения, пересечения, сложения и дополнения, называется алгеброй мультимножеств $L\ (\boldsymbol{Z})$, где максимальное мультимножество $\boldsymbol{Z}$ является единицей алгебры, а пустое мультимножество $\varnothing$ – нулем. Действительная неотрицательная функция $m(\boldsymbol{A})$, определенная на алгебре $L\ (\boldsymbol{Z})$ и удовлетворяющая условию сильной аддитивности: $m(\boldsymbol{A}+\boldsymbol{B})=m(\boldsymbol{A})+m(\boldsymbol{B})$, называется мерой мультимножества. Мера мультимножества $m(\boldsymbol{A})$ обладает следующим свойствами: $m(\varnothing)=0$; слабая аддитивность $m(\boldsymbol{A}\cup\boldsymbol{B})=m(\boldsymbol{A})+m(\boldsymbol{B})$, $\boldsymbol{A}\cap\boldsymbol{B}=\varnothing$; монотонность $m(\boldsymbol{A})\le m(\boldsymbol{B})\Leftrightarrow\boldsymbol{A}\subseteq\boldsymbol{B}$; симметричность $m(\boldsymbol{A})+m(\overline{\boldsymbol{A}})=m(\boldsymbol{Z})$; непрерывность $\lim\limits_{i\to\infty} m(\boldsymbol{A}_i)=m(\lim\limits_{i\to\infty}\boldsymbol{A}_i)$; эластичность $m(h\bullet\boldsymbol{A})=hm(\boldsymbol{A})$. Меру мультимножества можно определить различными способами, например, как линейную комбинацию функций кратности: $m(\boldsymbol{A})=\sum_j w_j k_A(x_j)$, $w_j>0$. Заметим, что мощность мультимножества $|\boldsymbol{A}|$ также будет мерой мультимножества

Метрические пространства мультимножеств ($\mathcal{A}$,d) были введены в [Петровский, 1995, 2003], где определены следующие виды расстояний между мультимножествами:

$d_1(\boldsymbol{A},\boldsymbol{B}) = m(\boldsymbol{A}\Delta\boldsymbol{B})$; $d_2(\boldsymbol{A},\boldsymbol{B}) = m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{Z})$; $d_3(\boldsymbol{A},\boldsymbol{B}) = m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{A}\cup\boldsymbol{B})$. (1)

Функции $d_2(\boldsymbol{A},\boldsymbol{B})$ и $d_3(\boldsymbol{A},\boldsymbol{B})$ удовлетворяют условию нормировки $0\le d(\boldsymbol{A},\boldsymbol{B})\le 1$. По определению принимается $d_3(\varnothing,\varnothing)=0$. Основное расстояние $d_1(\boldsymbol{A},\boldsymbol{B})$ является метрикой типа Хемминга, традиционно используемым во многих приложениях. Полностью усредненное расстояние $d_2(\boldsymbol{A},\boldsymbol{B})$ характеризует различие между двумя мультимножествами $\boldsymbol{A}$ и $\boldsymbol{B}$, отнесенное к расстоянию, максимально возможному в исходном пространстве. Локально усредненное расстояние $d_3(\boldsymbol{A},\boldsymbol{B})$ задает различие, отнесенное к максимально возможной «общей части» только этих двух мультимножеств в исходном пространстве.

2. Решающее правило классификации

Пусть $\mathcal{A}=\{A_1,...,A_k\}$ – совокупность объектов (владельцев кредитных карт), которые описываются m дискретными признаками $Q_1,\ldots,Q_m$, имеющими количественные и качественные (номинальные, либо порядковые) значения $q_s^{e_s}$, $e_s=1,\ldots,h_s$, $s=1,\ldots,m$. Порядковые значения качественного признака предполагаются упорядоченными от лучшего значения к худшему $q_s^{\,1}\succ q_s^{\,2}\succ\ldots\succ q_s^{\,h_s}$. Объекты A_i, $i=1,\ldots,k$ предварительно рассортированы по нескольким классам X_t, $t=1,\ldots,f$ путем прямой классификации. Принадлежность объекта A_i к некоторому классу X_t выражается правилом сортировки R, которое может считаться еще одним качественным признаком со значениями r_t. Признаки $Q_1,\ldots,Q_m$ выражают персональные сведения владельца карты, а R характеризует его кредитоспособность. Для описания каждого объекта используется только одно какое-то значение признака из каждой группы $Q_1,\ldots,Q_m,R$. Других дополнительных предположений об особенностях классов, признаков объектов и их значений (важности, предпочтительности, характерности и прочее) не делается. Требуется построить одно или несколько решающих правил, составленных из небольшого числа значений признаков, которые относили бы объекты к заданным классам наилучшим (в смысле близости к предварительной сортировке) образом.

Каждый объект A_i, $i=1,\ldots,k$ из совокупности $\mathcal{A}$ обычно представляется m-мерным вектором $q_i=(q_{i1}^{e_1}, q_{i2}^{e_2},\ldots, q_{im}^{e_m})$ в пространстве $Q=Q_1\times Q_2\times\ldots\times Q_m$, являющемся прямым произведением шкал значений признаков Q_s, и ищутся некоторые решающие правила для отнесения объекта к заданному классу X_t. Сопоставим теперь объект A_i с мультимножеством вида

$\boldsymbol{A}_i=\{(k_{Ai}(\,q_s^{e_s}\,)\bullet\, q_s^{e_s}\,),\,(k_{Ai}(r_t)\bullet r_t)\}$ над доменом $G=\{Q_1,\ldots,Q_m,R\}$. Для сокращения размерности мультимножества ряд признаков, особенно имеющих большое число значений, удобно агрегировать в небольшие группы. Например, вместо возраста владельца кредитной карты можно указывать возрастную группу (например, до 20 лет, 21-30 лет, 31-40 лет, и так далее), вместо точной величины дохода – некоторый интервал, вместо точного адреса – район или город, и тому подобное.

Представление объекта A_i мультимножеством $\boldsymbol{A}_i$ может трактоваться как способ выражения индивидуального правила сортировки вида:

ЕСЛИ ⟨условия⟩, ТО ⟨решение⟩. (2)

Терм ⟨условия⟩ ассоциируется с различными комбинациями значений признаков $q_s^{e_s}$, описывающими свойства объекта A_i, а терм ⟨решение⟩ – с принадлежностью объекта A_i к классу X_t.

Для простоты будем считать, что результатом классификации должно быть разложение совокупности объектов A только на два класса X_a и X_b (в нашем случае – на категории «хороших» или «плохих» владельцев карт). Требование бинарной декомпозиции совокупности $\mathsf{A}=\{X_a, X_b\}$ не является принципиальным ограничением. Если необходимо рассортировать объекты на большее число классов, можно сначала разбить совокупность объектов на две группы, затем одну из них или обе группы – на подгруппы, и так далее.

Рассмотрим наиболее простой и типичный случай агрегирования многопризнаковых объектов, когда каждая группа объектов формируются как сумма соответствующих им мультимножеств. Тогда каждое из мультимножеств $\boldsymbol{X}_t=\{(k_{Xt}(\,q_s^{e_s}\,)\bullet\, q_s^{e_s}\,),\,(k_{Xt}(r_t)\bullet r_t)\}$, $t=a,b$, представляющее свой класс объектов X_t, можно записать в виде следующего разложения на мультимножества по группам признаков:

$$\boldsymbol{X}_t=\sum_{s=1}^{m}\boldsymbol{Q}_{st}+\boldsymbol{R}_t,$$

где каждое слагаемое есть в свою очередь разложение

$$\boldsymbol{Q}^{st}=\sum_{e_s=1}^{h_s}\boldsymbol{Q}_{st}^{e_s}\ ,\qquad \boldsymbol{Q}_{st}^{e_s}=\sum_{i\in I_{st}^{e_s}}\boldsymbol{A}_i\ ,\qquad \boldsymbol{R}_t=\sum_{i\in I_{rt}}\boldsymbol{A}_i$$

Здесь $I_{st}^{e_s}$ и I_{rt} – подмножества индексов i для объектов $\boldsymbol{A}_i$, имеющих соответственно $k_{Ai}(\,q_s^{e_s}\,)\neq 0$ и $k_{Ai}(r_t)\neq 0$, $\boldsymbol{A}_i\subset\boldsymbol{X}_t$.

Очевидно, что объекты $\boldsymbol{A}_i$, которые попали в разложение $\{\boldsymbol{R}_a, \boldsymbol{R}_b\}$, сделанное только по индивидуальным правилам сортировки, образуют наилучшую из всех возможных декомпозиций рассматриваемой совокупности объектов $\mathsf{A}=\{\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k\}$ на два класса для имеющегося набора правил сортировки. Рассмотрим теперь метрическое пространство мультимножеств (A,

d) с метрикой d, определяемой одним из выражений (1), и обозначим через $d^*=d(\boldsymbol{R}_a, \boldsymbol{R}_b)$ расстояние между мультимножествами $\boldsymbol{R}_a$ и $\boldsymbol{R}_b$. В каждом конкретном случае расстояние d^* является предельно возможным расстоянием между объектами, входящими в разные классы.

Задача поиска обобщенного решающего правила классификации многопризнаковых объектов сводится к m оптимизационным задачам для каждой группы признаков Q_s

$$d(\boldsymbol{Q}_{sa}, \boldsymbol{Q}_{sb}) \to \max d(\boldsymbol{Q}_{sa}, \boldsymbol{Q}_{sb}) = d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*), \qquad (3)$$

где мультимножества $\boldsymbol{Q}_{sa}^*$ и $\boldsymbol{Q}_{sb}^*$ расположены на максимально возможном расстоянии в метрическом пространстве мультимножеств (A,d) и принадлежат к разным классам. Решение каждой из задач (3) является наилучшей бинарной декомпозиций $\{\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*\}$ имеющейся совокупности многопризнаковых объектов $A=\{\boldsymbol{A}_1,...,\boldsymbol{A}_k\}$ по s-ой группе признаков. Когда число h_s значений $q_s^{e_s}$ каждого из признаков невелико, решение задачи (3) не вызывает существенных трудностей и может быть получено даже путем перебора.

Каждое мультимножество $\boldsymbol{Q}_{st}^*$ ($t=a,b$), относящееся к одному и тому же классу, представляется как сумма двух подмультимножеств $\boldsymbol{Q}_{st}^*=\boldsymbol{Q}_{st}^{*1}+\boldsymbol{Q}_{st}^{*2}$. Значение признака q_s^*, которое определяет границу между сгенерированными слагаемыми $\boldsymbol{Q}_{st}^{*1}$ и $\boldsymbol{Q}_{st}^{*2}$ внутри каждой пары, назовем граничным. Комбинация граничных значений признаков $\{q_s^*\}$ из разных групп признаков Q_s, задает условия отнесения объекта $\boldsymbol{A}_i$ к соответствующему классу $\boldsymbol{X}_t$ и образует искомое обобщенное правило классификации совокупности многопризнаковых объектов вида (2).

Граничные признаки q_s^* можно упорядочить по величине расстояния $d(\boldsymbol{Q}_{sa}^*,\boldsymbol{Q}_{sb}^*)$. Для построения обобщенных правил классификации следует использовать признаки q_s^*, занимающие первые места в такой ранжировке. Чем ближе значение $d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*)$ к расстоянию $d^*=d(\boldsymbol{R}_a, \boldsymbol{R}_b)$, тем более точной будет аппроксимация первоначальной индивидуальной сортировки объектов. Оценить точность аппроксимации по s-ой группе признаков можно, например, выражением

$$\rho_s = d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*)/d(\boldsymbol{R}_a, \boldsymbol{R}_b).$$

В обобщенное решающее правило следует тогда включать значения признаков q_s^*, имеющие показатель точности ρ_s, превышающей некоторый желаемый пороговый уровень ρ_0. Заметим, что величина ρ_s показателя точности аппроксимации характеризует своего рода относительную важность s-ой группы признаков Q_s в обобщенном решающем правиле.

После построения обобщенных решающих правил, относящих объекты к каждому из заданных классов, необходимо провести содержательный анализ правил с целью выявления в них одинаковых комбинаций гранич-

ных значений признаков, которые (в зависмости от контекста решаемой задачи) нужно либо оставить, либо исключить из правил.

Заключение

Классификация объектов, которые описываются многими количественными и качественными признаками, причем каждый из объектов может существовать в нескольких различающихся «экземплярах», является достаточно трудной проблемой. Эти трудности имеют и содержательные основания (например, некорректность применения процедур «усреднения» качественных признаков), и формальные причины (например, большая размерность задачи). В работе предложен новый метод классификации многопризнаковых объектов, основанный на их представлении с помощью мультимножеств, который не содержит необоснованных преобразований исходной информации и не приводит к потере или искажению данных.

Аналогичный предложенному подход к построению обобщенного решающего правила для классификации объектов, описывающихся качественными признаками, был проверен на результатах конкурсного отбора проектов по государственной научно-технической программе по высокотемпературной сверхпроводимости [Ларичев и др., 1989]. Было сформулировано несколько решающих правил, одно из которых полностью совпало с примененным на практике [Петровский, 2003]. Обобщенное решающее правило классификации объектов позволило также выделить наиболее важные для отбора проектов критерии.

Список литературы

[Ларичев и др., 1989] Ларичев О.И., Прохоров А.С., Петровский А.Б., Стернин М.Ю., Шепелев Г.И. Опыт планирования фундаментальных исследований на конкурсной основе. // Вестник АН СССР, 1989, №7, 51-61.

[Петровский, 1995] Петровский А.Б. Метрические пространства мультимножеств. // Доклады Академии наук. 1995, Т.344, №2, 175-177.

[Петровский, 2003] Петровский А.Б. Упорядочение и классификация объектов с противоречивыми признаками. // Новости искусственного интеллекта. 2003, №4, 34-43.

[Петровский, 2003] Петровский А.Б. Пространства множеств и мультимножеств. – М.: Едиториал УРСС, 2003.

[Shi *et al*, 2001] Shi Y., Wise M., Luo M., Lin Y. Data mining in credit card portfolio management: a multiple criteria decision making approach. // M.Koksalan, S.Zionts (Eds.). Multiple Criteria Decision Making in New Millennium. – Springer-Verlag, Berlin, 2001, 427-436.

Петровский А. Б. Извлечение знаний для оценки кредитоспособности: подход теории мультимножеств // Труды Девятой национальной конференции по искусственному интеллекту с международным участием (КИИ-2004).— Т. 2. — М.: Физматлит, 2004.— С. 853–860.

```
@InProceedings{Petrovsky_2004c,
  author =       "Петровский, А. Б.",
  title =        "Извлечение знаний для оценки кредитоспособности: подход
                  теории мультимножеств",
  booktitle =    "Труды Девятой национальной конференции по искусственному
                  интеллекту с международным участием (КИИ-2004)",
  volume =       "2",
  address =      "М.",
  publisher =    "Физматлит",
  year =         "2004",
  pages =        "853--860",
  language =     "russian",
}
```