

# Экспертная классификация: ресурсный подход

П. И. Падерно<sup>1</sup>, Е. А. Бурков<sup>2</sup>, Е. А. Толкачева<sup>3</sup>

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В. И. Ульянова (Ленина)

<sup>1</sup>pipaderno@list.ru, <sup>2</sup>eaburkov@gmail.com,  
<sup>3</sup>eatolкачева@etu.ru

Е. А. Лавров

Сумский государственный университет  
prof\_lavrov@mail.ru

О. Е. Сирьк

Киевский национальный университет  
им. Т. Г. Шевченко  
lavrova\_olia@ukr.net

**Аннотация.** Рассмотрено проведение классификационной экспертизы с позиций эффективного использования ресурсов для обеспечения работы экспертной группы. Предложен комплекс показателей, позволяющих оценить как величину дохода от проведения экспертизы, связанную с получением правильного решения задачи классификации, так и возможные потери, вызванные допущенными экспертами ошибками. Сформулирована задача оптимального подбора экспертов в зависимости от априорного распределения объектов по классам.

**Ключевые слова:** эксперты; классификация; ресурсы; вероятности; оценки; расходы; доход; штрафы

Задача классификации сложных объектов той или иной природы может быть отнесена к категории трудно формализуемых задач, для решения которых достаточно часто используются методы экспертного оценивания [1, 2]. Когда речь идет о решении задачи классификации, экспертная группа должна на основе анализа всего множества предоставленных для экспертизы объектов отнести каждый из этих объектов к одному из предварительно определенных классов. В том случае, если суждения о классе принадлежности каждого объекта принимаются экспертами индивидуально, то для получения итогового результата классификационной экспертизы (обобщенного мнения экспертов), необходимо проведение анализа и агрегирования мнений отдельных экспертов [3–7]. Таким образом, проведение экспертизы представляет собой комплексный многоступенчатый (в некоторых случаях – итеративный) процесс, и его реализация нередко требует значительных ресурсных затрат.

**Замечание:** будем далее полагать, что программа и методика проведения классификационной экспертизы, регламентирующие порядок и содержательное наполнение всех стадий экспертизы (форма опроса экспертов, способ получения единого мнения, критерий проверки согласованности, процедура оценки достоверности полученных результатов и пр.) в достаточной степени отработаны, по-

этому затраты, связанные с организационно-подготовительной стадией экспертизы, рассматриваться не будут.

Пусть задана некоторая исходная система классов для всего множества возможных объектов, т. е. определены классы  $V_1, V_2, \dots, V_n$ . Будем полагать, что объекты с меньшим численным индексом класса превосходят по качеству объекты с большим численным индексом, а все привлеченные к решению задачи классификации объекты эксперты  $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_s$  обладают достаточно высоким уровнем квалификации. Последнее условие в данном случае означает, что при классификации конкретного объекта, принадлежащего к некоторому априорно неизвестному классу  $V_i$ , любой из экспертов может отнести его либо к классу  $V_i$ , к которому этот объект действительно принадлежит, либо к какому-то из соседних классов  $V_{i-1}$  или  $V_{i+1}$ , с вероятностями  $(q_i, q_{i-1}, q_{i+1})$  соответственно. Таким образом, для экспертной группы численностью  $s$  может быть сформирован набор квадратных вероятностных матриц  $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_s$ , отражающих способность (вероятности правильных и ошибочных суждений) каждого из экспертов осуществлять классификацию рассматриваемого множества объектов, используя систему классов  $V_1, V_2, \dots, V_n$ .

**Случай 1.** Пусть в классификационной экспертизе одновременно принимают участие эксперты  $\mathcal{E}_i, \mathcal{E}_j$  и  $\mathcal{E}_m$ . В рамках другого доклада, подготовленного для данной конференции и посвященного построению и применению вероятностной модели для формального описания классификационных экспертиз, демонстрируется, что вероятности правильных и ошибочных результатов такой экспертизы также могут быть представлены в виде ленточной матрицы следующего общего вида:

$$Q_{ijm} = \begin{pmatrix} q_{11} & q_{12} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q_{21} & q_{22} & q_{23} & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & q_{ll-1} & q_{ll} & q_{ll+1} & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & q_{m-1} & q_{mm} \end{pmatrix} \quad (1)$$

Для оценки эффективности классификационной экспертизы с ресурсной точки зрения рассмотрим расходы, доходы и штрафы, связанные с проведением подобной экспертизы.

**Расходы.** Расход ресурсов обусловлен привлечением к решению задачи классификации специалистов, работа которых в роли экспертов должна быть определенным образом обеспечена, и зависит от квалификации этих специалистов. Также имеется зависимость от количества и природы объектов, подлежащих классификации, поскольку эти факторы определяют степень нагрузки на экспертов.

Следовательно, еще на этапе формирования экспертной группы для проведения классификационной экспертизы известного множества объектов может быть определен набор векторов  $\bar{r}_1, \bar{r}_2, \dots, \bar{r}_k, \dots, \bar{r}_s$ , элементы  $r_{kl}$  которых представляют собой стоимости анализа объекта  $l$ -го класса ( $l = 1, 2, \dots, n$ )  $k$ -м экспертом ( $k = 1, 2, \dots, s$ ). Если множество объектов, подлежащих классификации, имеет мощность  $M$ , т. е. экспертам предстоит проанализировать  $M$  объектов, и известно некоторое априорное распределение объектов по классам, задаваемое вероятностным вектором  $\bar{h} = (h_1, h_2, \dots, h_n)$ , то расходы на проведение классификационной экспертизы могут быть вычислены по следующей формуле:

$$R_{ijm} = M \sum_{l=1}^n (r_{il} + r_{jl} + r_{ml}) h_l. \quad (2)$$

Тогда можно сформулировать задачу оптимизации расходов на проведение классификационной экспертизы в виде задачи минимизации выражения (2), которая может быть решена полным или усеченным перебором. При этом необходимо учитывать, что минимальное значение выражения (2) существенно зависит от распределения объектов по классам в заданном множестве объектов, и поэтому для классификации различающихся по этому показателю исходных множеств объектов может быть целесообразно привлечение различных экспертов.

*Замечание:* при отсутствии какой-либо достоверной информации об априорном распределении объектов по классам можно сделать предположение, что оно является равномерным, т. е. допустить, что вероятностный вектор для заданного множества объектов имеет вид  $\bar{h} = (1/n, 1/n, \dots, 1/n)$ .

**Доход.** Введем в рассмотрение вектор  $\bar{u} = (u_1, u_2, \dots, u_n)$ , где  $u_l$  – доход, который образуется в случае правильной классификации объекта, т. е. когда объект из класса  $V_l$  был отнесен экспертами к  $l$ -му же классу. Тогда общий доход

от результатов проведения классификационной экспертизы может быть вычислен по формуле

$$U_{ijm} = M \cdot \sum_{l=1}^n u_l h_l q_{ll}, \quad (3)$$

где  $q_{ll}$  – диагональные элементы матрицы (1).

**Штрафы.** Введем в рассмотрение два вектора штрафов:

- $\bar{v} = (v_2, v_3, \dots, v_n)$  – вектор штрафов за оптимизм: элементы этого вектора представляют собой штрафы, накладываемые за ошибочное отнесение объекта из класса  $V_l$  к классу  $V_{l-1}$  (завышение класса объекта);
- $\bar{w} = (w_1, w_2, \dots, w_{n-1})$  – вектор штрафов за пессимизм: элементы этого вектора представляют собой штрафы за ошибочное отнесение объекта из класса  $V_l$  к классу  $V_{l+1}$  (занижение класса объекта).

Общие штрафы за оптимизм и пессимизм при проведении классификационной экспертизы могут быть найдены по следующим формулам:

$$V_{ijm} = M \cdot \sum_{l=2}^n v_l h_l q_{ll-1}, \quad (4)$$

$$W_{ijm} = M \cdot \sum_{l=1}^{n-1} w_l h_l q_{ll+1}, \quad (5)$$

где  $q_{ll-1}$  и  $q_{ll+1}$  – поддиагональные и наддиагональные элементы матрицы (1) соответственно.

По аналогии с (2) необходимо учитывать, что как общий доход (3), так и общие штрафы за оптимизм (4) и пессимизм (5) в значительной степени зависят не только от квалификации экспертов, принимавших участие в экспертизе, но и от распределения объектов по классам в множестве, представленном на экспертизу. Однако при этом ни расход, ни доход, ни штрафы не зависят от порядка следования экспертов, т. к. все три эксперта работают параллельно.

**Случай 2.** Пусть теперь в экспертизе принимают участие те же эксперты  $\mathcal{E}_i, \mathcal{E}_j$  и  $\mathcal{E}_m$ , но в этот раз в два этапа. На первом этапе опрашиваются только эксперты  $\mathcal{E}_i$  и  $\mathcal{E}_j$ . На втором этапе задействуется только эксперт  $\mathcal{E}_m$ , которому для анализа предоставляются лишь те объекты, по которым на первом этапе мнение экспертов  $\mathcal{E}_i$  и  $\mathcal{E}_j$  не было единодушным, и они отнесли каждый такой объект к разным классам. Использование подобной двухэтапной схемы работы экспертов позволит в определенной степени снизить расходы на проведение классификационной экспертизы.

**Расходы.** Расходы, возникающие на первом этапе опроса экспертов, можно вычислить по следующей формуле:

$$R_{ij} = M \sum_{l=1}^n (r_{il} + r_{jl}) h_l. \quad (6)$$

Зная вероятностные матрицы  $\mathbf{Q}_i$  и  $\mathbf{Q}_j$ , отражающие способности к классификации объектов  $i$ -м и  $j$ -м экспертами соответственно, оценим количество объектов, для которых будет необходимо провести второй этап опроса с привлечением  $m$ -го эксперта:

$$M^* = M \sum_{l=1}^n (1 - (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + q_{ill+1} q_{jll+1})) h_l. \quad (7)$$

*Замечание:* исходя из введенной системы обозначений, будем полагать, что значения вероятностей  $q_{k10}$  и  $q_{km+1}$  равны 0 для всех  $k = 1, 2, \dots, s$ .

Далее оценим расходы, связанные с проведением второго этапа опроса и привлечением  $m$ -го эксперта:

$$R(m | i, j) = M \sum_{l=1}^n r_{ml} (1 - (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + q_{ill+1} q_{jll+1})) h_l. \quad (8)$$

Тогда совокупные расходы на работу экспертной группы при проведении классификационной экспертизы с двухэтапной схемой опроса экспертов можно представить в виде суммы расходов, возникающих на первом и втором этапах:

$$\begin{aligned} R_{ijm} &= R_{ij} + R(m | i, j) = M \sum_{l=1}^n (r_{il} + r_{jl}) h_l + \\ &+ M \sum_{l=1}^n r_{ml} (1 - (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + q_{ill+1} q_{jll+1})) h_l = \\ &= M \sum_{l=1}^n (r_{il} + r_{jl} + r_{ml}) h_l - M \sum_{l=1}^n r_{ml} (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + \\ &+ q_{ill+1} q_{jll+1}) h_l = R_{ijm} - M \sum_{l=1}^n r_{ml} (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + \\ &+ q_{ill+1} q_{jll+1}) h_l. \end{aligned} \quad (9)$$

Таким образом, экономия от проведения классификационной экспертизы с применением двухэтапной схемы опроса экспертов составит:

$$M \sum_{l=1}^n r_{ml} (q_{ill} q_{jll} + q_{ill-1} q_{jll-1} + q_{ill+1} q_{jll+1}) h_l. \quad (10)$$

Очевидно, что чем больше величина (10), тем целесообразнее проведение классификационной экспертизы по двухэтапной схеме, и тем меньше общие расходы на проведение экспертизы. При этом величина сэкономленных ресурсов (10) в значительной степени зависит не только от характеристик экспертов, задействованных на каждом из этапов опроса, но и в немалой степени от распределения объектов по классам в представленном на экспертизу множестве объектов. Отсюда следует, что при известном

априорном распределении объектов по классам (или хотя бы при наличии правдоподобной гипотезы об этом распределении), можно найти наилучший с точки зрения критерия минимизации расходов на проведение экспертизы вариант выбора трех экспертов из группы  $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_s$ , т. е. осуществить выбор такой комбинации трех экспертов, которая обеспечивает максимальное значение выражению (10). При этом следует принимать во внимание, что для другого априорного распределения вероятностей оптимальный состав экспертов, а также их распределение по двум этапам опроса могут быть совершенно другими.

*Замечание:* если в расчет не принимаются ожидаемые доходы и штрафы (риски), то максимизация выражения (10) обеспечивает минимальные расходы на проведение классификационной экспертизы.

В свою очередь чистый ожидаемый доход от проведения классификационной экспертизы можно оценить по формуле:

$$Z_{ijm} = U_{ijm} - R_{ijm} - V_{ijm} - W_{ijm}. \quad (11)$$

Как можно заключить на основании последнего сделанного замечания, набор экспертов, обеспечивающий максимальное значение выражению (10) в общем случае не обеспечивает максимального значения левой части выражения (11), ввиду наличия в правой части выражения (11) доходной и штрафных функций.

**Оценка объемов вычислений.** Очевидно, что для решения оптимизационной задачи в рассмотренном выше случае 1, когда все три эксперта работают параллельно и независимо, для нахождения наиболее выгодного состава экспертов методом полного перебора необходимо проверить  $C_s^3$  возможных комбинаций (троек) экспертов, т. е.  $s(s-1)(s-2)/6$  различных вариантов комбинаций из трех экспертов.

Для случая 2 при двухэтапном проведении опроса экспертов, когда на первом этапе два эксперта работают параллельно, а на втором этапе третий эксперт привлекается лишь в случае несогласованности первых двух экспертов, количество проверяемых вариантов будет в три раза больше, чем в случае 1, и составит  $s(s-1)(s-2)/2$ .

**Планирование экспертизы.** При планировании классификационной экспертизы рабочая группа, кроме желания обеспечить минимальный расход ресурсов, должна принимать во внимание возможные требования (ограничения) предъявляемые как к привлекаемым экспертам, так и к самой экспертизе, связанные с достоверностью получаемых результатов, основными из которых могут являться следующие:

- надежность каждого из экспертов должна быть не ниже некоторого  $\gamma_s$ , при этом под надежностью эксперта  $\mathcal{E}_i$  может пониматься как наименьшая возможная вероятность правильной классификации некоторого объекта этим экспертом  $\rho_1(\mathcal{E}_i) = \min_l q_{ill}$ , так и средняя вероятность пра-

вильной классификации им объекта

$$\rho_2(\Theta_i) = \sum_{l=1}^n q_{il} / n;$$

- надежность полученных результатов должна быть не ниже некоторого  $\gamma_p$ , при этом под надежностью результатов классификационной экспертизы может пониматься как наименьшая возможная вероятность правильной классификации некоторого объекта  $\rho_3(\Theta_i, \Theta_j, \Theta_m) = \min_l q_{il}$ , где  $q_{il}$  – диагональные элементы матрицы (1), так и средняя вероятность правильной классификации объекта по результатам экспертизы  $\rho_4(\Theta_i, \Theta_j, \Theta_m) = \sum_{l=1}^n q_{il} / n$ ;
- вероятность ошибки при классификации объекта должна быть не больше  $\delta = 1 - \gamma$ .

*Замечание:* функции доходов и штрафов, а также всевозможные дополнительные требования и ограничения должны быть заданы до начала экспертизы, т. е. на стадии подготовки технического задания на ее проведение.

**Выводы.** На стадии подготовки классификационной экспертизы предложенный ресурсный подход может быть применен для прогнозирования ожидаемых результатов, а также предварительной оценки расходов, доходов и штрафов для различных по составу групп экспертов и варьирующихся условий проведения экспертизы: априорного распределения объектов по классам, функций доходов и штрафов, ограничений и т. п. Компьютерные модели, позволяющие проводить подобное прогнозирование и получать необходимые оценки, могут выступать частью интеллектуальной системы информационной поддержки подготовки и проведения классификационных экспертиз.

На основе результатов проведенного выше анализа можно заключить, что ресурсный подход к проведению классификационных экспертиз позволяет:

1. оценить по критерию чистого ожидаемого дохода целесообразность проведения классификационной экспертизы с заданным составом экспертной груп-

пы и известным множеством подлежащих анализу объектов;

2. спрогнозировать потери ресурса, обусловленные ошибками, допущенными экспертами при решении задачи классификации объектов;
3. формализовать процедуру определения наиболее эффективного по расходу ресурсов сочетания экспертов для проведения классификационной экспертизы на заданном множестве объектов, используя информацию об априорном распределении объектов по классам.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] E. A. Burkov, P. I. Paderno, Adaptation of Saaty's Analytic Network Process. Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM'2016. Saint-Petersburg, 2016. Vol. 1. pp. 38-41.
- [2] P. I. Paderno, E. A. Burkov, E. A. Lavrov, Issues of Organization of Expertise and Problems of Expert Assessments. Journal of Physics: Conference Series, 1703 012047, XXIII International Conference on Soft Computing and Measurement (SCM'2020) 27-29 May 2020, Russia. DOI: 10.1088/1742-6596/1703/1/012047
- [3] E. A. Burkov, P. L. Lyubkin, P. I. Paderno, Quantitive Estimation of Extent of Coincidence of Expertise's Objects Models. Proceedings of the 20th IEEE International Conference on Soft Computing and Measurements, SCM'2017. Saint-Petersburg, 2017. pp. 43-45. DOI: 10.1109/SCM.2017.7970490
- [4] E. A. Burkov, P. I. Paderno, About Correctness of Linear Transformations of Expert Estimates. Proceedings of the 21th International Conference on Soft Computing and Measurements, SCM'2018. Saint-Petersburg, 2018. Vol. 1. pp. 66-69.
- [5] E. A. Burkov, N. A. Nazarenko, S. S. S. Nasser, P. I. Paderno, Analysis of Correctness of Linear Transformations of Expert Estimates. Proceedings of the 22th IEEE International Conference on Soft Computing and Measurements, SCM'2019. Saint-Petersburg, 2019. pp. 29-32. DOI: 10.1109/SCM.2019.8903758
- [6] E. D. Dutova, N. A. Nazarenko, P. I. Paderno, Analysis of the Influence of Transformation and Integration Technology of the Expert Evaluations on the Result. Proceedings of the 19th IEEE International Conference on Soft Computing and Measurements, SCM'2016. Saint-Petersburg, 2016. pp. 21-24. DOI: 10.1109/SCM.2016.7519671
- [7] E. A. Burkov, P. I. Paderno, O. E. Siryk, E. A. Lavrov, N. B. Pasko, Analysis of Impact of Marginal Expert Assessments on Integrated Expert Assessment. Proceedings of the 23rd IEEE International Conference on Soft Computing and Measurements, SCM'2020. Saint-Petersburg, 2020. pp. 14-17. DOI: 10.1109/SCM50615.2020.9198772