Использование формулы Байеса при групповом экспертном рейтинговом оценивании

Г. И. Кожомбердиева[1], Д. П. Бураков[2], Г. А. Хамчичев[3]
Петербургский государственный университет путей сообщения Императора Александра I
[1]kgi-liizht@yandex.ru, [2]burakovdmitry8@gmail.com, [3]gyqn@yandex.ru

***Аннотация*. В докладе обсуждается байесовский подход к решению задачи рейтингового оценивания объектов. Результатом оценивания группой экспертов каждого объекта является апостериорное распределение вероятностей на множестве гипотез о ранге этого объекта. Используемые при вычислении апостериорного распределения значения условных вероятностей понимаются как степени соответствия полученных свидетельств (экспертных оценок) предположениям об истинности гипотез. При этом эксперт имеет право в случае неуверенности присваивать объекту смежные ранги, что является преимуществом подхода. Итоговый рейтинг объектов формируется в результате обработки всех полученных апостериорных распределений.**

Ключевые слова: рейтинг; экспертное оценивание; индивидуальные и групповые предпочтения; принятие решений малой группой; байесовский подход; формула Байеса.

I. Введение

Байесовские технологии широко используются при решении задач, требующих анализа данных для выбора из альтернатив, выяснения причин событий или явлений, прогнозирования событий, выполнения измерений в условиях неопределённости, неточности или неполноты информации [1 − 4].

Байесовский подход как направление в науке об управлении основан на принципе использования имеющейся априорной информации, ее непрерывного пересмотра и переоценки с учетом получаемых выборочных данных (свидетельств) об исследуемом явлении или процессе. Такой пересмотр трактуется как обучение, и сам процесс управления при байесовском подходе понимается как процесс обучения (адаптации). При этом используется интерпретация понятия вероятность, отличная от частотного определения, − *байесовская вероятность*. Такая вероятность понимается как степень уверенности в истинности суждения, а для определения такой степени уверенности при получении новой информации используется теорема Байеса [1, 2, 5].

Активному внедрению байесовских вероятностных моделей способствует новое актуальное и мощное средство – *системы вероятностного программирования* (СВП). Под СВП понимается система вероятностных рассуждений, для которой языком представления моделей является язык программирования [6, 7]. В монографии [6], например, показано, как язык вероятностного программирования (*Probabilistic Programming Language − PPL*) Figaro используется для представления модели байесовской сети при создании приложений фильтра спама и диагностики вычислительной системы.

Использование формулы Байеса не ограничивается традиционными задачами. В работах авторов [8 − 12] предложен способ применения байесовского подхода к решению задач оценивания качества на основе неточных результатов измерений и/или субъективных экспертных оценок. В принципе, оценивание качества некоторого объекта по ряду показателей можно рассматривать как частный случай рейтингового оценивания множества объектов группой экспертов [13].

Целью работы, результаты которой представлены в настоящем докладе, является адаптация байесовского подхода к оцениванию качества, изложенного в [8 − 12], к задаче группового экспертного рейтингового оценивания.

II. Рейтинговое оценивание множества объектов группой экспертов

Под *рангом* $\rho(x)$ некоторого объекта x понимается порядковый показатель, отражающий его значимость и характеризующий его место среди других объектов: $\rho(x) \in \boldsymbol{N}$, $x \in X$. *Рейтингом* называется определяемый по специальной методике или путем опроса компетентных экспертов список объектов, расположенных в порядке возрастания рангов. В случае, если объекты предварительно каким-либо образом пронумерованы, т.е. каждому $x_i \in X$ в соответствие поставлен номер $i \in \{1, \ldots, N\}$, где $N = |X|$ – число оцениваемых объектов, то рейтинг можно представить вектором $\mathbf{r} = (\rho_1, \ldots, \rho_i, \ldots, \rho_N)$, где $\rho_i = \rho(x_i)$.

Экспертиза является частным случаем коллективного принятия решений − принятием решения *малой группой*, причем приемлемой с точки зрения объективности группового решения и удобства организации экспертизы является численность группы 7 ± 2. Важным фактором группового выбора является поиск согласованных оценок. В соответствии с этим под групповым выбором понимается выработка согласованного решения о порядке предпочтения рассматриваемых объектов на основе индивидуальных мнений членов группы [13, 14].

Именно экспертиза может использоваться при решении задачи рейтингового оценивания множества объектов.

III. Байесовский подход к решению задачи рейтингового оценивания

В настоящем докладе обсуждается использование формулы Байеса при групповом экспертном рейтинговом оценивании объектов. Предлагаемый способ основан на подходе, изложенном в работах [8 – 12] применительно к решению задач оценивания качества.

Предположим, что имеется множество ранжируемых объектов X, $|X| = N$ и группа из $m > 1$ экспертов. В процессе оценивания каждому объекту $x_i \in X, i = \overline{1, N}$ каждым из экспертов $j = \overline{1, m}$ присваивается ранг $\rho_j(x_i) \in \{1, \ldots, k\}$. При этом любой эксперт, в случае неуверенности, имеет право присваивать объекту одновременно два смежных ранга: $\rho_j^1(x_i)$ и $\rho_j^2(x_i)$, таких, что $|\rho_j^1(x_i) - \rho_j^2(x_i)| = 1$. Следует отметить, что в общем случае число используемых экспертами рангов k может быть меньше числа оцениваемых объектов N (в задачах выделения «призовой группы»).

В контексте рейтингового оценивания базовым понятием является апостериорное распределение вероятностей на множестве гипотез о ранге, присваиваемом объекту. Значения условных вероятностей, используемых при вычислениях по формуле Байеса, понимаются как степени соответствия полученных свидетельств (рангов, присвоенных объекту отдельными экспертами группы) предположениям об истинности гипотез о ранге объекта. Обозначим через H_l гипотезу о том, что произвольному объекту $x \in X$ присвоен ранг $\rho(x) = l$. Через $P(H_l)$ обозначим априорную вероятность данной гипотезы. При условии отсутствия у оценивающей группы экспертов предварительных сведений о «качестве» оцениваемых объектов естественно предположить, что эти вероятности формируют равномерное распределение: $P(H_l) = 1/k$, $l = \overline{1, k}$. Отметим, что данное априорное распределение может быть как единым для всей группы объектов X, так и индивидуальным для каждого объекта. Индивидуальное распределение вероятностей для каждого объекта обозначим через $P_i(H_l)$, $i = \overline{1, N}$, $l = \overline{1, k}$.

Результатом оценивания группой экспертов каждого объекта является апостериорное распределение вероятностей на множестве гипотез о ранге этого объекта:

$$P_i\left(H_l \mid e_{(i)}^{(l)}\right) = \frac{P_i\left(e_{(i)}^{(l)} \mid H_l\right) \cdot P_i(H_l)}{\sum_{s=1}^{k} P_i\left(e_{(i)}^{(s)} \mid H_s\right) \cdot P_i(H_s)}, \quad l = \overline{1, k};\ i = \overline{1, N}. \tag{1}$$

Если априорное распределение $P_i(H_l)$ является равномерным, формулу (1) можно переписать в следующем виде:

$$P_i\left(H_l \mid e_{(i)}^{(l)}\right) = \frac{P_i\left(e_{(i)}^{(l)} \mid H_l\right)}{\sum_{s=1}^{k} P_i\left(e_{(i)}^{(s)} \mid H_s\right)}, \quad l = \overline{1, k};\ i = \overline{1, N}. \tag{2}$$

Здесь $P_i\left(H_l \mid e_{(i)}^{(l)}\right)$ – апостериорная вероятность истинности гипотезы о том, что $\rho(x_i) = l$, пересмотренная в связи с имеющимися свидетельствами $e_{(i)}^{(l)}$ о назначении объекту x_i ранга l некоторым подмножеством экспертов, а $P_i\left(e_{(i)}^{(s)} \mid H_s\right)$ – условная вероятность, интерпретируемая как степень соответствия предположению об истинности гипотезы о том, что $\rho(x_i) = s$, полученных свидетельств $e_{(i)}^{(s)}$ о присвоении объекту x_i ранга s некоторым (возможно, пустым) подмножеством экспертов. Эта вероятность вычисляется как отношение числа экспертов, присвоивших объекту x_i ранг s, к общей численности группы экспертов:

$$P_i\left(e_{(i)}^{(s)} \mid H_s\right) = \frac{\left|\left\{\rho_j(x_i) \mid \rho_j(x_i) = s,\ j = \overline{1, m}\right\}\right|}{m}. \tag{3}$$

На основе формулы (1) или (2) формируется множество из N k-элементных апостериорных распределений $\left\{P_i\left(H_l \mid e_{(i)}^{(l)}\right)\right\}$, $l = \overline{1, k}$; $i = \overline{1, N}$.

Итоговые ранги объектов $\rho^*(x_i), i = \overline{1, N}$ определяются после обработки апостериорных распределений вероятностей согласно правилу:

$$\rho^*(x_{i^*}) = l;\ i^* = \arg\max_{i \in \overline{1, N}} \left\{\sum_{s=1}^{l} P_i\left(H_s \mid e_{(i)}^{(s)}\right)\right\};\ l = \overline{1, k}. \tag{4}$$

Согласно правилу (4), выявляется объект с номером i^*, имеющий максимальную сумму вероятностей для рангов $1, \ldots, l$, начиная с $l = 1$. При этом объекты $x \in X$, которым итоговый ранг $\rho^*(x)$ уже присвоен, исключаются из дальнейшего рассмотрения.

Альтернативным способом является назначение объекту *нормализованного* ранга, полученного на основе вычисленного математического ожидания [5] случайной величины $\hat{\rho}_i$, $i = \overline{1, N}$, множеством значений которой является множество возможных рангов (гипотез), а законом распределения – i-е апостериорное распределение вероятностей:

$$\mathrm{M}(\hat{\rho}_i) = \sum_{s=1}^{k} s \cdot P_i\left(H_s \mid e_{(i)}^{(s)}\right). \tag{5}$$

При нормализации ранга математические ожидания (5) заменяются натуральными числами из множества $\{1, \ldots, k\}$ с сохранением полученного порядка в рейтинге.

Отметим, что при любом из предлагаемых способов формирования итогового рейтинга могут возникать *коллизии*, то есть ситуации, когда невозможно назначить различные ранги двум произвольным объектам $x_i, x_j \in X$, основываясь на имеющихся оценках экспертов. В таком случае можно провести повторный раунд оценивания в

соответствии с описанной выше процедурой, ограничившись объектами, ранги которых не удалось определить.

Кроме того, при необходимости формулу (3), вычисляющую условную вероятность $P_i\left(e_{(i)}^{(s)} \mid H_s\right)$, несложно подвергнуть модификации, позволяющей учитывать при оценивании как различную компетентность экспертов, так и право эксперта в случае неуверенности не просто назначать объекту смежные ранги, но и присваивать ранги в виде долей целого голоса.

IV. Пример

a) *Исходные данные*

Пусть имеются 5 участников некоторого условного конкурса. К оцениванию привлекаются 7 экспертов, которые должны выявить трех призеров.

b) *Решение*

Каждый из экспертов присваивает ранг (место) каждому участнику. Ранг 4 проставляется участникам, не входящим в призовую тройку. Если эксперт не уверен в своих предпочтениях, он может назначить участникам смежные ранги (например, 2-е и 3-е места), как это показано в табл. I.

ТАБЛИЦА I Присвоение мест экспертом № 1

Эксперт 1		Места			
		1	2	3	4
Участники	*A*	*			
	B		*	*	
	C				*
	D		*	*	
	E				*

Затем для каждого участника x_i определяются отношения числа голосов, поданных экспертами за каждое из мест в рейтинге, к численности группы экспертов, т.е. оценки условных вероятностей $P_i\left(e_{(i)}^{(l)} \mid H_l\right)$. Они представляют собой степени соответствия экспертных заключений предположениям об истинности гипотез H_l, $i=\overline{1,5},\ l=\overline{1,4}$.

Далее по формуле (2) вычисляются апостериорные условные вероятности $P_i\left(H_l \mid e_{(i)}^{(l)}\right)$, понимаемые как степень уверенности экспертов в справедливости гипотез о получении участником x_i l-го места в рейтинге, показывающие, насколько правдоподобными стали гипотезы H_l, $i=\overline{1,5},\ l=\overline{1,4}$, после оценивания экспертами. Пример данных и вычислений представлен в табл. II.

Затем, в целях большей наглядности, апостериорные распределения вероятностей на множестве гипотез о рангах (местах в рейтинге) для всех участников записываются в табл. III.

ТАБЛИЦА II Ранги и байесовские вероятности для участника *B*

Участник *B*		Места – гипотезы H_l			
		1	2	3	4
Эксперты	1		*	*	
	2	*			
	3	*	*		
	4			*	
	5			*	*
	6		*		
	7			*	*
Оценки $P(e \mid H_l)$		2/7	3/7	4/7	2/7
Ап. вер. $P(H_l \mid e)$		2/11	3/11	4/11	2/11

ТАБЛИЦА III Байесовские апостериорные вероятности

		Места			
		1	2	3	4
Участники	*A*	0.43	0.29	0.14	0.14
	B	0.18	0.27	0.37	0.18
	C	0.00	0.13	0.25	0.62
	D	0.20	0.30	0.20	0.20
	E	0.12	0.00	0.12	0.76

Обозначим для удобства и наглядности апостериорные вероятности, указанные в ячейках табл. III, идентификаторами соответствующих им участников и рангов (P_{A1}, P_{A2} и т.д.) и определим тройку призеров конкурса, используя критерий максимума сумм (4).

Победителем, очевидно, становится участник *A*, так как вероятность $P_{A1} = 0.43$ является максимальной из множества $\{P_{A1}, P_{B1}, \ldots, P_{E1}\}$.

Для определения участника, занявшего 2-е место, просуммируем вероятности P_{x1} и P_{x2}, где $x \in \{B, C, D, E\}$, так как уже получивший свой ранг участник *A* исключен из множества: $P_{B1}+P_{B2} = 0.45$; $P_{C1}+P_{C2} = 0.13$; $P_{D1}+P_{D2} = 0.5$; $P_{E1}+P_{E2} = 0.13$. Таким образом, $\rho^*(D) = 2$.

Аналогично, просуммировав по три вероятности на множестве $\{B, C, E\}$, определим, что 3-е место получает участник *B*: $P_{B1}+P_{B2}+P_{B3} = 0.82$; $P_{C1}+P_{C2}+P_{C3} = 0.38$; $P_{E1}+P_{E2}+P_{E3} = 0.24$.

Таким образом, участники *C* и *E* выбывают из тройки призеров и получают итоговый ранг 4: $\rho^*(C) = \rho^*(E) = 4$.

Тот же самый результат получится, если вычислить по формуле (5) математические ожидания рангов, а затем нормализовать их, как показано в табл. IV.

ТАБЛИЦА IV Математические ожидания рангов

Ранги		$M(\hat{\rho}_i)$	$\rho^*(x_i)$
Участники	*A*	1.99	1
	B	2.55	3
	C	3.49	4
	D	2.20	2
	E	3.52	4

V. Особенности и преимущества подхода

Как уже отмечалось выше, предлагаемый подход к групповому рейтинговому оцениванию, как, впрочем, и любые другие, в силу особенностей коллективного выбора,

отмеченных в известной теореме Эрроу [13, 14], может приводить к коллизиям.

В то же время, предлагаемый байесовский способ определения рейтинга располагает, на наш взгляд, следующими встроенными возможностями по устранению коллизий.

1. Повторный раунд оценивания с использованием при необходимости для каждого из оцениваемых объектов апостериорных вероятностей гипотез, полученных в ходе предыдущего раунда, в качестве априорных вероятностей.

2. Учет различной компетентности экспертов.

3. Предоставление экспертам права в случае неуверенности не просто назначать оцениваемому объекту смежные ранги, но и присваивать ранги в виде долей целого голоса.

VI. Заключение

В работе показана применимость байесовского подхода [8−12], ориентированного на решение задачи оценивания качества некоторого объекта по множеству показателей, к задаче рейтингового оценивания множества объектов группой экспертов, и выполнена его адаптация к задаче рейтингового оценивания.

Данный подход позволяет гибко и без существенного увеличения вычислительной сложности устранять коллизии, возможные при групповом оценивании объектов.

Для подтверждения применимости предлагаемого метода разработана демонстрационная программа на языке Java, которая при дальнейшей доработке может быть использована для быстрого решения задачи распределения рейтинговых мест среди участников различных мероприятий, например, студенческих конкурсов. Отметим, что изначально для решения рассмотренной задачи планировалось использовать библиотеку Figaro для языка Scala [6], но ввиду относительной вычислительной простоты задачи было решено использовать при создании демонстрационной программы рейтингового оценивания язык базовой для Scala платформы − Java.

Список литературы

[1] Моррис У.Т. Наука об управлении: байесовский подход; [пер. с англ.]. М.: Мир, 1971. 304 с.

[2] Райфа Г. Анализ решений (введение в проблему выбора в условиях неопределенности); [пер. с англ.]. М.: Наука, 1977. 408 с.

[3] Уткин Л.В. Анализ риска и принятие решений при неполной информации. СПб: Наука, 2007. 404 с.

[4] Прокопчина С.В. Мягкие измерения и управление сложными системами на основе регуляризирующего байесовского подхода // Экономика и управление: проблемы, решения. 2015. Т. 5. № 12. С. 16–25.

[5] Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. М.: Наука, 1988. 480 с.

[6] Пфеффер А. Вероятностное программирование на практике; [пер. с англ.]. М.: ДМК Пресс, 2017. 462 с.

[7] Knowledge Plus Statistics: Understanding the Emerging World of Deep Probabilistic Programming Languages [Электронный ресурс]. URL: https://towardsdatascience.com/knowledge-plus-statistics-understanding-the-emerging-world-of-deep-probabilistic-programming-504baf9d47f3 (дата обращения 24.03.2019)

[8] Кожомбердиева Г.И., Красковский А.Е. Способ определения условных вероятностей при байесовском оценивании качества управленческих решений на железнодорожном транспорте // Интеллектуальные системы на транспорте: матер. IV Междунар. науч.-практич. конф. «ИнтеллектТранс-2014». СПб: Изд-во ПГУПС, 2014. С. 412–418.

[9] Кожомбердиева Г.И., Бураков Д.П., Гарина М.И. Использование формулы Байеса при оценивании выполнения практик модели CMMI® // Программные продукты и системы. 2017. Т. 30. № 4. С. 67–74.

[10] Кожомбердиева Г.И., Бураков Д.П. Об использовании формулы Байеса в задачах оценивания качества // Международная конференция по мягким вычислениям и измерениям. 2017. Т. 1. С. 31–34.

[11] Кожомбердиева Г.И., Бураков Д.П. Байесовский подход к решению задач оценивания качества // Мягкие измерения и вычисления. 2018. № 5. С. 15–26.

[12] Бураков Д.П., Кожомбердиева Г.И. Использование формулы Байеса при оценивании качества программного обеспечения согласно стандарту ISO/IEC 9126 // Программные продукты и системы. 2019. Т. 32. № 1. С. 34–41.

[13] Микони С.В. Многокритериальный выбор на конечном множестве альтернатив: Учебное пособие. СПб: Издательство «Лань», 2009. 272 с.

[14] Колбин В.В. Математические методы коллективного принятия решений: Учебное пособие. СПб: Издательство «Лань», 2015. 256 с.