

Научная статья
УДК 519.2

Применение последовательной байесовской процедуры в задачах экспертного оценивания

Александр Иванович Буравлев

Аннотация. В статье рассмотрен статистический метод адаптивного экспертного оценивания сложных многопараметрических объектов, для которых отсутствуют возможности использования виртуальных (с помощью моделей) или натурных (экспериментальных) испытаний и соответствующих информационно-измерительных систем. Проведение экспертизы, по сути, представляет собой эксперимент, только в качестве информационно-измерительной системы в нем используется человек, обладающий знанием, опытом и профессиональной компетентностью. Результаты, выдаваемые экспертами, также содержат случайные ошибки и для анализа их достоверности (точности и надежности) могут быть использованы известные методы статистического анализа. Автором разработана методика и алгоритм адаптивного экспертного оценивания на основе байесовской процедуры. Приведен иллюстративный пример ее использования. Предлагаемый метод может быть использован при оценке качества и пригодности активно разрабатываемых в настоящее время систем искусственного интеллекта.

Ключевые слова: гипотезы и шкала решений экспертов; байесовские оценки вероятностей гипотез по результатам экспертизы; учет ошибок экспертов при оценке вероятностей их решений; критерии останова экспертной процедуры

Для цитирования: Буравлев А.И. Применение последовательной байесовской процедуры в задачах экспертного оценивания // Вооружение и экономика. 2026. №3(77). С. 5-11.

Original article

Application of a Consistent Bayesian Procedure in Expert Evaluation Tasks

Aleksandr I. Buravlev

Abstract. The article considers a statistical method of adaptive expert evaluation of complex multiparametric objects for which there are no possibilities of virtual (by means of models) or field (experimental) tests usage and corresponding information and measurement systems. Carrying out an expert examination is essentially an experiment, but it includes a human input with knowledge, experience, and professional competence in an information and measurement system. The results provided by experts also contain random errors, and well-known statistical analysis methods can be used to analyze their reliability (accuracy and reliability). The author has developed a methodology and algorithm for adaptive expert evaluation based on the Bayesian procedure. An illustrative example of their usage is given. The proposed method can be used for the quality and suitability evaluation of currently actively developed artificial intelligence systems.

Keywords: hypotheses and the scale of expert decisions; Bayesian estimates of hypothesis probabilities based on examination results; accounting for expert errors in the course of the decisions probabilities evaluation; criteria for the expert procedure stopping

For citation: Buravlev A.I. Application of a Consistent Bayesian Procedure in Expert Evaluation Tasks. *Vooruzhenie i ekonomika = Armament and Economics*. 2026; 77(3): 5-11. (In Russ.).

Экспертное оценивание находит широкое применение в задачах программно-целевого планирования развития систем вооружения и военной техники (ВВТ) на всех его этапах (целеполагания, анализа состояний объектов и условий их применения, выбора вариантов и стратегий развития ВВТ, вариантов их практической реализации и оценки эффективности) [1-3]. Это связано с высокой степенью неопределенности и динамичности процесса планирования и необходимостью в использовании знаний и опыта высококвалифицированных специалистов (экспертов) из разных областей (военное дело, военная техника и технологии, военное производство и экономика, управления организацией и др.). Эксперт как носитель знаний и опыта их практического применения обладает не только компетентностью, но и определенными социально-психологическими качествами (интуицией, предвидением, творческим озарением, этикой и моралью), которые присущи только Человеческому интеллекту. Практика применения различного рода экспертиз показывает, что в условиях ограниченной информации и высокой степенью ее неопределенности, решения экспертов оказываются

объективнее и точнее, чем получаемые на сложных имитационных и оптимизационных моделях, особенно в условиях многокритериального выбора [3-6].

При организации экспертизы важно правильно сформулировать ее цель, определить перечень факторов (признаков, параметров), подлежащих оцениванию, выбрать шкалы оценивания и разработать тесты для оценки уровня компетентности экспертов и степени их согласованности для предстоящей экспертизы [2]. На данном этапе в целях повышения уровня согласованности целесообразно использовать методы адаптации суждений экспертов к рассматриваемой проблеме с учетом результатов их тестирования, а также в ходе последующей работы. Ознакомление эксперта с результатами оценки других экспертов повышает уровень информированности эксперта и качество его экспертизы [2; 4].

Далее мы рассмотрим применение байесовской процедуры для экспертного оценивания, как одного из методов повышения уровня информированности экспертов с использованием вероятностно-статистического подхода.

1 Байесовский алгоритм адаптации экспертного оценивания

Пусть эксперту предложено сравнить пару альтернатив A_1, A_2 для некоторой ситуации выбора. Каждая альтернатива A характеризуется набором факторов (параметров) $U_A = (u_1, u_2, \dots, u_m)$, отражающих ее свойства (качество). По этим параметрам эксперт производит оценивание и сравнение альтернатив A_1, A_2 , используя известные методы (балльных оценок, парных сравнений, голосованием). Относительно альтернатив A_1, A_2 возможны следующие гипотезы:

$$H_1 = \{A_1 > A_2\}; H_2 = \{A_1 < A_2\}; H_0 = \{A_1 \sim A_2\},$$

где: $A_1 > A_2$ означает, что альтернатива A_1 предпочтительнее альтернативы A_2 или наоборот $A_2 > A_1$; $A_1 \sim A_2$ – альтернативы несравнимы между собой. Эксперт может принять решение относительно сравниваемых альтернатив путем формирования индикатора $X = 1$ в пользу альтернативы A_1 или $X = -1$ в пользу альтернативы A_2 , но может и отказаться от сравнения $X = 0$, если посчитает указанные альтернативы несравнимыми $A_1 \sim A_2$. Таким образом, результаты сравнения могут быть выражены в следующей шкале:

$$X = \begin{cases} 1, & \text{если } A_1 > A_2 \\ -1 & \text{если } A_1 < A_2 \\ 0 & \text{если } A_1 \sim A_2 \end{cases}. \quad (1)$$

Признак несравнимости $X = 0$ никак не влияет на оценки альтернатив, данные другими экспертами, но исключает результат данного эксперта из подсчета голосов. Отказ эксперта, с одной стороны, снижает его рейтинг компетентности, а с другой, – служит поводом для исследования причин отказа, в том числе, для выяснения достаточности представленных на экспертизу характеристик объектов U_A .

Возможна и другая схема оценки, когда альтернативы практически не различимы для эксперта $A_1 \sim A_2$ и он может считать их эквивалентными (равновозможными, равноценными). Какое значения в этом случае нужно приписать признаку решения эксперта $X = ?$. Здесь целесообразно использовать принцип «недостаточного основания» Лапласа при оценке результата эксперта. При $A_1 \sim A_2$ вполне допустимо принять $X = 1$, потому, как нет оснований считать, что $A_1 < A_2$. В этом случае нужно просмотреть результаты оценки других экспертов для данной ситуации и использовать вероятностно-статистический анализ этих решений.

Будем рассматривать процедуру экспертного оценивания исходных альтернатив как последовательную схему испытаний, начиная с одного наугад выбранного эксперта $n = 1$. По результатам опроса n экспертов рассчитываются статистические оценки вероятностей принятия гипотез:

$$P_n(H_1) = \sum_{i=1}^n (X_i = 1) / n = n_1 / n; P_n(H_2) = \sum_{i=1}^n (X_i = -1) / n = n_2 / n;$$

$$P_n(H_0) = \sum_{i=1}^n (X_i = 0)/n = n_0/n; P_n(H_1) + P_n(H_2) + P_n(H_0) = 1, \quad (2)$$

где: n_1, n_2, n_0 – число экспертов, принявших решение в пользу рассматриваемых гипотез.

Тогда для очередного $(n + 1)$ -го эксперта результаты экспертизы принимают следующие значения:

$$\begin{aligned} P_{n+1}(H_1|X_{n+1} = 1) &= (n_1 + 1)/(n + 1); \\ P_{n+1}(H_2|X_{n+1} = -1) &= (n_2 + 1)/(n + 1); \\ P_{n+1}(H_0|X_{n+1} = 0) &= (n_0 + 1)/(n + 1). \end{aligned} \quad (3)$$

Из выражений (2), (3) вытекают неравенства:

$$\begin{aligned} P_{n+1}(H_1|X_{n+1} = 1) &= (n_1 + 1)/(n + 1) > n_1/n = P_n(H_1); \\ P_{n+1}(H_2|X_{n+1} = -1) &= (n_2 + 1)/(n + 1) > n_2/n = P_n(H_2). \end{aligned}$$

Таким образом, при увеличении числа экспертов апостериорные вероятности гипотез увеличиваются, если решения экспертов отражают истинное содержание гипотез и уменьшаются при их несовпадении. Это полностью соответствует содержательному смыслу статистического подхода, когда увеличение числа экспериментов повышает достоверность оценки свойств исследуемого объекта.

При выборе своих решений эксперт может допускать ошибки 1-го рода (не признавая действительное) с вероятностью $\alpha = P(X = -1|H_1)$ и 2-го рода (выдавая желаемое за действительное) $\beta = P(X = 1|H_2)$. Вероятность этих ошибок характеризует уровень компетентности эксперта их можно определить посредством предварительного тестирования экспертов. Методы тестирования широко применяются в настоящее время при оценке уровня знаний и профессиональной компетенции специалистов¹.

Оценки для вероятностей ошибок эксперта по результатам тестирования определяются следующими формулами:

$$\alpha = \sum_{k=1}^N (X_k = -1|H_1)/N; \beta = \sum_{k=1}^N (X_k = 1|H_2)/N,$$

где: N – число тестовых заданий с двумя ответами (истинным и ложным).

В ходе проведения реальной экспертизы эти оценки могут уточняться.

В процессе экспертизы экспертам сообщаются результаты других участников процедуры, что позволяет корректировать их индивидуальные оценки. Для каждого последующего эксперта ($j > n$) можно оценить ожидаемую вероятность (прогноз) принятия его решений по формуле полной вероятности:

$$\begin{aligned} P(X_j = 1) &= P(X_j = 1|H_1)P_n(H_1) + P(X_j = 1|H_2)P_n(H_2) = (1 - \alpha)P_n(H_1) + \beta P_n(H_2); \\ P(X_j = -1) &= P(X_j = -1|H_1)P_n(H_1) + P(X_j = -1|H_2)P_n(H_2) = \alpha P_n(H_1) + (1 - \beta)P_n(H_2). \end{aligned} \quad (4)$$

Такая обратная связь является основой адаптивного метода оценивания.

Далее, используя байесовскую процедуру, можно оценить вероятности принятия гипотез с учетом решений очередного $(n + 1)$ -го эксперта²:

$$\begin{aligned} P_n(H_2|X_{n+1} = -1) &= \\ &= P_n(X_{n+1} = 1|H_1)P_n(H_1)/\{P_n(X_{n+1} = 1|H_1)P_n(H_1) + P_n(X_{n+1} = 1|H_2)P_n(H_2)\} = \\ &= (1 - \alpha_{n+1})P_n(H_1)/\{(1 - \alpha_{n+1})P_n(H_1) + \beta_{n+1}P_n(H_2)\}; \\ P_n(H_2|X_{n+1} = -1) &= \\ &= P_n(X_{n+1} = -1|H_2)P_n(H_2)/\{P_n(X_{n+1} = -1|H_1)P_n(H_1) + P_n(X_{n+1} = -1|H_2)P_n(H_2)\} = \\ &= (1 - \beta_{n+1})P_n(H_2)/\{\alpha_n P_n(H_1) + (1 - \beta_{n+1})P_n(H_2)\}. \end{aligned} \quad (5)$$

¹ Чельшева М.Б. Теория и практика конструирования педагогических тестов: учеб. пособие. М.: Логос, 2002. 432 с.

² Там же.

Апостериорные вероятности гипотез также характеризуют степень достоверности решений эксперта относительно оцениваемых гипотез:

$$D_n(H_1) = P_n(H_1|X_{n+1}); D_n(H_2) = P_n(H_2|X_{n+1}).$$

Адаптация результатов экспертного оценивания завершается путем приравнивания априорных и апостериорных вероятностей гипотез с учетом решений очередного $(n + 1)$ -го эксперта:

$$P_{n+1}(H_1) = P_n(H_1|X_{n+1} = 1); P_{n+1}(H_2) = P_n(H_2|X_{n+1} = -1). \quad (6)$$

В качестве априорного распределения для $n = 1$ необходимо принять:

$$P_1(H_1) = P_1(H_2) = 0,5.$$

Возможна также схема самообучения эксперта по уровню различения гипотез $\lambda_n = P_n(H_1)/P_n(H_2)$:

$$X_{n+1} = \begin{cases} 1, & \text{если } \lambda_n \geq \hat{\lambda} \\ -1 & \text{если } \lambda_n < \hat{\lambda} \\ 0 & \text{если } A_1 \propto A_2. \end{cases} \quad (7)$$

где: $\hat{\lambda} > 1$ – заданный уровень различения гипотез.

Предлагаемые адаптивные алгоритмы обеспечивают при увеличении числа экспертов сходимость вероятностей принятия гипотез $P_{n+1}(H_1)$, $P_{n+1}(H_2)$ к стационарным значениям. Останов процесса экспертного оценивания осуществляется при достижении заданной точности:

$$\max\{|P_{n+1}(H_1) - P_n(H_1)|; |P_{n+1}(H_2) - P_n(H_2)|\} \leq \delta,$$

где: $\delta = 0,05$ – заданный уровень погрешности оценивания или при достижении заданного порога различения гипотез, например, $\lambda_n \geq 3,0$.

Предлагаемый подход позволяет учесть главную особенность экспертного оценивания – это наличие субъективной составляющей в оценке альтернатив, которые имеют как случайный, так и неопределенный характер в силу социально-психологических особенностей данного вида «измерений». Точность и надежность таких измерений обеспечивается повышением уровня компетентности экспертов и их численности для решения «неизмеримых» в классическом понимании проблем.

2 Иллюстративный пример

Группе экспертов, компетентных в области вооружения и военной техники (ВВТ), предлагается провести экспертизу военно-технического уровня двух образцов A_1 , A_2 . Каждый образец характеризуется набором военно-технических характеристик $U_A = (u_1, u_2, \dots, u_m)$. Экспертам предлагается произвести многопараметрическое сравнение образцов ВВТ, в том числе с применением различных методов [4; 5]. Результаты экспертизы приведены в таблице 1. Решения, принимаемые экспертами, соответствуют схеме (1). Эксперты прошли предварительное тестирование с оценкой уровня их компетентности и численных значений их возможных ошибок 1-го и 2-го рода, которые также приведены в таблице. Начальные значения вероятностей оцениваемых гипотез. Начальные значения вероятностей принятия гипотез принимаются на уровне $P_1(H_1) = P_1(H_2) = 0,5$, т.е. с максимальной неопределенностью.

На рисунке 1 приведены графики изменения вероятностей гипотез без учета ошибок экспертов, а на рисунке 2 – с учетом ошибок экспертов, полученных по данным их тестирования. Из приведенных далее в статье таблиц следует, что по мере увеличения числа экспертов (экспериментов) вероятности гипотез стабилизируются и стремятся к своим стационарным значениям $\bar{P}(H_1)$, $\bar{P}(H_2)$.

Таблица 1 – Данные экспертного оценивания изделий ВВТ

№ эксперта	0	1	2	3	4	5	6	7	8	9	10
Решения эксперта X	0	1	1	-1	1	1	0	-1	1	1	1
Вероятности ошибок экспертов 1-го и 2-го рода											
α	0	0,1	0,05	0,07	0,09	0,01	0,03	0,08	0,1	0,15	0,03
β	0	0,15	0,1	0,1	0,05	0,12	0,05	0,05	0,12	0,12	0,09
Оценки вероятности гипотез для сравниваемых альтернатив без учета ошибок ($\alpha = \beta = 0$)											
$P(H_1)$	0,50	0,67	0,75	0,60	0,67	0,71	0,71	0,63	0,67	0,70	0,73
$P(H_2)$	0,50	0,33	0,25	0,40	0,33	0,29	0,29	0,38	0,33	0,30	0,27
Оценки вероятности гипотез для сравниваемых альтернатив с учетом ошибок экспертов											
$P(H_1)$	0,50	0,65	0,74	0,60	0,62	0,74	0,71	0,59	0,64	0,63	0,73
$P(H_2)$	0,50	0,35	0,26	0,40	0,38	0,26	0,29	0,41	0,36	0,37	0,27
Оценки достоверности решений экспертов											
$D(H_1)$	0,50	0,92	0,97	0,93	0,97	0,95	0,98	0,97	0,94	0,94	0,97
$D(H_2)$	0,5	0,81	0,86	0,90	0,84	0,97	0,93	0,88	0,81	0,72	0,92

№ эксперта	11	12	13	14	15	16	17	18	19	20	21
Решения эксперта X	0	-1	-1	1	1	0	1	-1	0	1	0
Вероятности ошибок экспертов 1-го и 2-го рода											
α	0,07	0,12	0,07	0,11	0,03	0,12	0,05	0,1	0,08	0,1	0
β	0,15	0,1	0,03	0,1	0,05	0,07	0,04	0,05	0,1	0,08	0
Оценки вероятности гипотез для сравниваемых альтернатив без учета ошибок ($\alpha = \beta = 0$)											
$P(H_1)$	0,73	0,67	0,62	0,64	0,67	0,67	0,69	0,65	0,65	0,67	0
$P(H_2)$	0,27	0,33	0,38	0,36	0,33	0,33	0,31	0,35	0,35	0,33	0
Оценки вероятности гипотез для сравниваемых альтернатив с учетом ошибок экспертов											
$P(H_1)$	0,72	0,62	0,58	0,61	0,66	0,61	0,67	0,60	0,63	0,63	0
$P(H_2)$	0,28	0,38	0,42	0,39	0,34	0,39	0,33	0,40	0,37	0,33	0
Оценки достоверности решений экспертов											
$D(H_1)$	0,94	0,95	0,98	0,94	0,97	0,96	0,98	0,97	0,94	0,96	0
$D(H_2)$	0,82	0,79	0,90	0,82	0,94	0,79	0,90	0,84	0,86	0,82	0

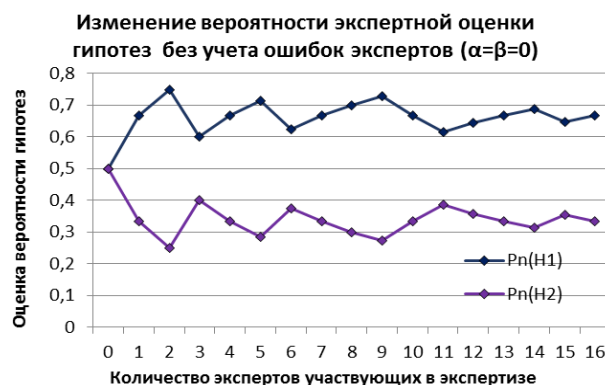


Рисунок 1 – Динамика вероятностей принятия гипотез в зависимости от числа экспертов

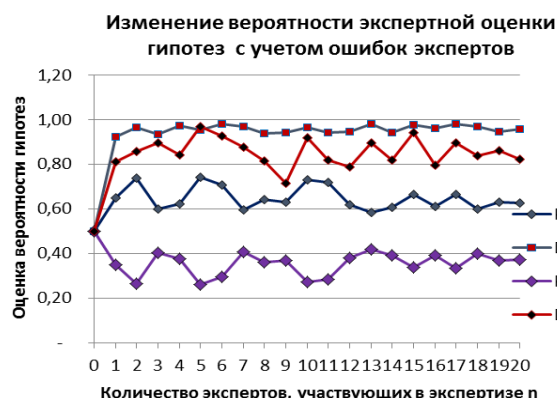


Рисунок 2 – Динамика вероятностей принятия гипотез с учетом вероятностей ошибок

Наличие ошибок экспертов, безусловно, снижает качество экспертизы, однако достоверность принимаемых решений остается достаточно высокой $D_{2n}(H_1) = 0,96$; $D_{2n}(H_2) = 0,82$.

Применение алгоритмов адаптации экспертов на основе доведения им информации о достигнутых результатах позволяет улучшить качество экспертизы. В таблице 2 приведены результаты расчета при адаптации экспертов по критерию порога различения гипотез (7). Здесь каждому последующему эксперту сообщалась величина достигнутого порога $\lambda_n = P_n(H_1)/P_n(H_2)_n$. В таблице они отмечены серым фоном ячейки.

Таблица 2 – Данные экспертного оценивания с адаптацией

№ эксперта	0	1	2	3	4	5	6	7	8
Решения эксперта X	0	1	1	1	-1	1	1	-1	1
$P_n(H_1)$	0,50	0,67	0,75	0,80	0,67	0,71	0,75	0,67	0,70
$P_n(H_2)$	0,50	0,33	0,25	0,20	0,33	0,29	0,25	0,33	0,30
λ_n	1	1,86	1,96	2,94	1,83	2,10	2,36	1,54	1,71
$D_n(H_1)$	0,50	0,92	0,96	0,97	0,94	0,95	0,98	0,93	0,50
$D_n(H_2)$	0,94	0,97	0,96	0,97	0,95	0,98	0,97	0,95	0,96

№ эксперта	9	10	11	12	13	14	15	16	17
Решения эксперта X	1	1	1	-1	1	1	1	1	0
$P_n(H_1)$	0,73	0,75	0,77	0,71	0,73	0,75	0,76	0,78	0
$P_n(H_2)$	0,27	0,25	0,23	0,29	0,27	0,25	0,24	0,22	0
λ_n	2,51	2,17	2,61	2,10	2,48	2,74	2,62	2,66	0
$D_n(H_1)$	0,97	0,96	0,97	0,95	0,98	0,97	0,95	0,96	0
$D_n(H_2)$	0,72	0,87	0,71	0,75	0,78	0,87	0,86	0,72	0

Из таблицы 2 видно, что два эксперта (№13, №16) поменяли свои оценки с учетом полученной информации о предыдущих результатах экспертизы. На рисунке 3 показаны изменения вероятностей экспертной оценки с учетом адаптации результатов экспертизы.

Из анализа таблиц 1, 2 и рисунков также видно, что при числе экспертов $n = 10$ достигается достаточно высокий порог различия гипотез $\lambda_n = 2,7$ с вероятностями $P_n(H_1) = 0,73$; $P_n(H_2) = 0,27$ и достоверностями $D_n(H_1) = 0,97$; $D_n(H_2) = 0,92$. С увеличением

числа экспертов в два раза порог различия гипотез снизился до уровня $\lambda_{2n} = 1,7$, также снизились вероятности принятия гипотез $P_{2n}(H_1) = 0,63$; $P_{2n}(H_2) = 0,33$ и их достоверности $D_{2n}(H_1) = 0,96$; $D_{2n}(H_2) = 0,82$. Данный результат свидетельствует о том, что увеличение числа экспертов с одинаковым или меньшим уровнем компетенции не способствует повышению достоверности принимаемых решений. Не случайно практика экспертного анализа показывает, что число экспертов должно находиться в пределах 8-10 человек. Адаптация экспертов с учетом группового мнения позволяет уменьшать их ошибки и повышать объективность экспертного оценивания.

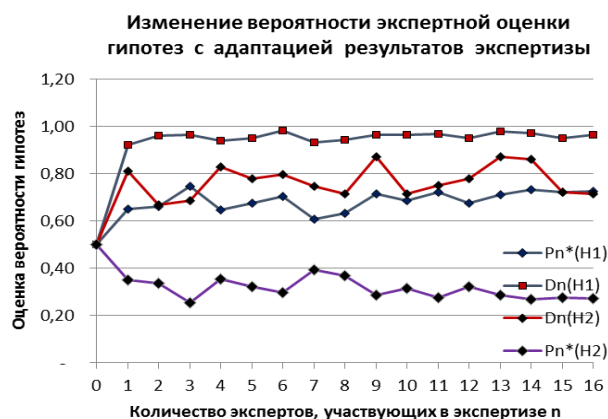


Рисунок 3 – Динамика вероятностей принятия гипотез с учетом адаптации результатов экспертизы

Вывод

Рассмотренный выше метод адаптивного экспертного оценивания позволяет без применения сложных алгоритмов многопараметрического анализа получить интегральную оценку сравниваемых альтернатив на основе знаний и опыта экспертов. Эти алгоритмы целесообразно использовать для обучения экспертных систем с применением искусственного интеллекта, так как здесь идет полная передача знаний и опыта человека машинному интеллекту [7].

Список источников

1. Методология программно-целевого планирования развития системы вооружения на современном этапе / Под общ. ред. В.М. Буренка. М.: Граница, 2013. 519 с.
2. Бешелев С.Д., Гурвич Ф.Г. Математико-статистические методы экспертных оценок. М.: Статистика, 1980. 263 с.
3. Литвак Б.Г. Экспертная информация. Методы получения и анализа. М.: Радио и связь, 1982. 184 с.
4. Орлов А.И. Организационно-экономическое моделирование. Ч.2: Экспертные оценки. М.: МГТУ им. Н.Э. Баумана, 2011. 486 с.
5. Семенов С.С. Оценка качества и технического уровня сложных систем: Практика применения метода экспертных оценок. М.: Ленанд, 2015. 352 с.
6. Саати Т.Л. Принятие решений: метод анализа иерархий. М.: Радио и связь, 1993. 320 с.
7. Буравлев А.И., Ветошкин В.М. Искусственный интеллект: сущность, принципы работы, области применения // Вооружение и экономика. 2024. №2(68). С. 33-42.

Информация об авторе

А.И. Буравлев – доктор технических наук, профессор.