

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ COMPUTER SIMULATION HISTORY

УДК 004.09

DOI: 10.18413/2518-1092-2017-2-3-3-10

Никитин В.В.¹
Гунченко Ю.И.²
Басов О.О.¹

**ОЦЕНКА УСЛОВНЫХ ВЕРОЯТНОСТЕЙ БАЙЕСОВСКОЙ СЕТИ
ДОВЕРИЯ ПРИ АПРИОРНОЙ ИНФОРМАЦИИ
О ВЗАИМОДЕЙСТВИИ МЕЖДУ ЕЕ УЗЛАМИ В СИСТЕМЕ
МНОГОМОДАЛЬНОЙ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЯ**

¹ Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики (Университет ИТМО), Кронверкский пр., д. 49, г. Санкт-Петербург, 197101, Россия

² Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук, 14 линия, д. 39, Санкт-Петербург, 199178, Россия

e-mail: nikitin.aktash@mail.ru, gunchenko.yui@gmail.com, oobasov@mail.ru

Аннотация

Рассмотрен подход, реализованный в разработанной математической модели системы многомодальной аутентификации пользователя, получения оценки условных вероятностей байесовской сети доверия на основе использования нечисловой экспертной, неточной и неполной информации о биометрических параметрах пользователя. Найден усредненные оценки вероятностных параметров и соответствующие им стандартные отклонения, дающие возможность решить задачу оценки вероятности легитимности пользователя, работающего с системой.

Ключевые слова: несанкционированный доступ; аутентификация; многомодальность; байесовская сеть доверия; вероятность.

UDC 004.09

Nikitin V.V.¹
Gunchenko Yu.I.²
Basov O.O.¹

**ESTIMATION OF CONDITIONAL PROBABILITIES OF A BAYESIAN
NETWORK OF TRUST WITH A PRIORI INFORMATION ON THE
INTERACTION BETWEEN ITS NODES IN THE MULTIMODAL USER
AUTHENTICATION SYSTEM**

¹ St. Petersburg National Research University of Information Technologies, Mechanics and Optics (University of ITMO), 49 Kronverksky Prospect, St. Petersburg, 197101, Russia

² Federal state Budgetary Institution of Science of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39 14-th Linia, St. Petersburg, 199178, Russia

e-mail: nikitin.aktash@mail.ru, gunchenko.yui@gmail.com, oobasov@mail.ru

Abstract

The approach realized in the developed mathematical model of the multimodal user authentication system, obtaining an estimate of the conditional probabilities of the Bayesian network of trust based on the use of non-numerical expert, inaccurate and incomplete information about the user's biometric parameters is considered. The averaged estimates of probabilistic parameters and the standard deviations corresponding to them are obtained, which make it possible to solve the problem of estimating the probability of legitimacy of the user working with the system.

Keywords: unauthorized access; authentication; multimodality; Bayesian network of trust; probability.

Неуклонный рост информационного пространства человека в процессе информатизации общества предъявляет особые требования к вопросам обеспечения защиты и конфиденциальности данных. Анализ [1] технических возможностей современных автоматизированных систем по обеспечению защиты от несанкционированного доступа (НСД) свидетельствует о том, что далеко не полностью реализованы возможности автоматизированных рабочих мест пользователя по осуществлению процедуры его аутентификации. Повсеместно отчеты различных организаций в сфере защиты информации показывают [2], что защищенность автоматизированных систем от воздействий компьютерных атак и различных программ скрытого информационного воздействия не отвечает текущим угрозам. Существующие потребности в повышении уровня контроля доступа к информации и эффективном использовании автоматизированных систем [3], а также развитие методов получения нелегального доступа к информации, определяют особую актуальность научного поиска в области разработки новых и совершенствования существующих методик аутентификации пользователей.

Разработка эффективной многомодальной системы аутентификации пользователя, позволяющая осуществлять защиту от НСД и контроль доступа легитимного пользователя с учетом данных нескольких каналов коммуникативного взаимодействия пользователя и автоматизированной системы является выходом из сложившейся ситуации. В основе многомодальной аутентификации заложена имеющаяся возможность учета информации о функциональном (физиологическом, эмоциональном, психофизиологическом) состоянии пользователя [4], что способствует повышению достоверности реализуемой процедуры.

Непосредственно математическая модель многомодальной системы аутентификации пользователя на основе байесовской сети доверия (БСД) [5] представлена в работе [6]. Она позволяет комплексно учитывать информацию о различных биометрических признаков аутентификации пользователя, которые объединяются (группируются) на основе методов многомодального объединения [4, 7]. В процессе моделирования в среде GeNIe [8] также была найдена возможность решения задачи вероятностного прогнозирования достоверности аутентификации пользователя на основе данных о его биометрических признаках.

По получаемой информации (в процессе анализа работы пользователя) с помощью байесовского метода в разработанной модели появляется возможность вычисления вероятности истинной гипотезы (легитимность пользователя). Вновь получаемая информация многомодальной системой аутентификации выступает в качестве исходных данных для модификации априорных вероятностей, на основе которых вычисляются апостериорные вероятности. Также, еще одним его положительным моментом является возможность применения субъективных вероятностных оценок априорных гипотез при отсутствии эмпирических данных [9].

Уточнение таблиц условных вероятностей для моделируемой БСД производится на основе экспертной информации, имеющихся данных различных научных исследований, накопленных статистических данных, которые могут быть применимы в многомодальных системах аутентификации пользователя.

Однако, при анализе различных модальностей имеет место дефицит статистической числовой информации, приводящий к существенным трудностям в успешной идентификации объектов, описываемых наборами различных биометрических характеристик. Для устранения данного недостатка к оценке условных БСД при дефиците информации о взаимодействиях между узлами привлекают экспертов. Экспертная оценка, к сожалению, не может предоставить точную числовую оценку этих вероятностей, а мнения экспертов ограничены рядом сравнительных суждений о превосходстве или равенстве вероятностей принадлежности объекта к различным категориям.

Особенностью многомодальной аутентификации пользователя является учет информации, получаемую из разных каналов (источников) и обладающую различной значимостью (по разным параметрам – важности, информативности, надежности и т.п.). В этом случае, оценка относительной значимости источников информации с применением весовых коэффициентов

затруднительна, и, как правило, используются только сравнительные суждения о значениях весовых коэффициентов сопоставимых сравнительным суждениям о вероятностях.

Стоит отметить, что для экспертов, привлекаемых к оценке, именно такая нечисловая (ординальная) форма представления информации о вероятностях и весовых коэффициентах является наиболее приемлемой, что подтверждают различные психологические эксперименты [10, 11], а представление экспертных знаний в точной числовой форме вызывают появление смещенных оценок, и как итог – неверные выводы [12].

Интервалы возможных значений параметров также могут быть использованы в работе экспертов для оценок вероятностей [13]. Соответственно, появляется возможность объединения данной интервальной (неточной) с нечисловой информацией. Однако, для определения числовых значений всех вероятностей, порой и указанная объединенная информация оказывается недостаточно полной. Таким образом, далее каждый привлекаемый эксперт имеет нечисловую экспертную, неточную и неполную (ННН) информацию по оцениваемым вероятностям, а «супер»-экспертом выступает непосредственно сам исследователь, который обладает ННН-информацией по весовым коэффициентам соответствующих экспертных мнений.

Рассмотрим данный случай более детально, когда имеется необходимость вычисления оценки вероятности p_i , характеризующей принадлежность исследуемого объекта к определенному классу A_i из совокупности r альтернативных классов, с использованием информации I . Данная информация (I), исходит от некоторого эксперта (формируется из некоторого источника), и согласно описанию выше, подразделяется на две категории – нечисловая OI , (например, «вероятность альтернативы A_i меньше, чем вероятность альтернативы A_s » и т.п.), и интервальная информация II (например, диапазоны $[a_i, b_i]$, $0 \leq a_i \leq b_i \leq 1$, $i = 1, \dots, r$, возможного варьирования вероятностей p_1, \dots, p_r).

Соответственно, объединенная нечисловая и неточная информация $I = OI \cup II$ описывается следующей системой неравенств и равенств для вероятностей p_1, \dots, p_r различных альтернатив A_1, \dots, A_r :

$$I = \begin{cases} p_i > p_l, \\ p_u = p_v, \\ a_s \leq p_s \leq b_s, \\ i, l, u, v, s \in \{1, \dots, r\}. \end{cases} \quad (1)$$

В связи с тем, что даже объединенная информация I (1) не позволяет найти числовые значения вероятностей p_1, \dots, p_r , можно сделать вывод о том, что информация I – неполна, а эксперт обладает ННН-информацией I о вероятностях p_1, \dots, p_r альтернатив A_1, \dots, A_r .

Комплексный учет ННН-информации I формирует множество $P(r; I)$ всех векторов вероятностей, которые с геометрической точки зрения представляют собой $(r - 1)$ -мерный полиэдр $P(r; I)$. Сама ННН-информация I с точностью до множества $P(r; I)$ как раз и определяет вектор вероятностей $p = (p_1, \dots, p_r)$.

Получить случайный вектор, распределенный на полиэдре $P(r; I)$:

$$\begin{aligned} \tilde{p}(I) &= (\tilde{p}_1(I), \dots, \tilde{p}_r(I)), \\ \tilde{p}_i(I) &\geq 0, \\ \tilde{p}_1(I) + \dots + \tilde{p}_r(I) &= 1, \end{aligned} \quad (2)$$

можно путем моделирования неопределенности выбора вектора вероятностей $p = (p_1, \dots, p_r)$ из множества $P(r; I)$ согласно [14].

Рассмотрим его более детально: стохастической оценкой вероятности альтернативы A_i , учитывающей ННН-информацию I является компонента $\tilde{p}_i(I)$ случайного вектора $\tilde{p}(I)$, а получаемые при этом математические ожидания:

$$\bar{p}_i(I_j) = E \tilde{p}_i(I_j) \quad (3)$$

выступают в качестве усредненных оценок вероятностей p_i , $i=1, \dots, r$.

Стандартные отклонения, рассчитываемые как:

$$\delta_i(I) = \sqrt{D \tilde{p}_i(I)} \quad (4)$$

задают меру разброса стохастических оценок $\tilde{p}_i(I)$ от согласованных с ними усредненных значений $\bar{p}_i(I)$. Таким образом, вектор:

$$\bar{p}(I) = (\bar{p}_1(I), \dots, \bar{p}_r(I)) \quad (5)$$

усредненных оценок вероятностей представляет собой числовой образ ННН-информации I .

При многомодальной аутентификации ННН-информацию, получаемую по коммуникативным каналам взаимодействия пользователя и АС, удобно задать кортежем:

$$I = \langle I_1, \dots, I_m \rangle, \quad (6)$$

компонента I_j которого представляет какую-либо систему неравенств (или равенств) для вероятностей p_1, \dots, p_r , определяющая с точки зрения ННН-информации I_j множество $P(r; I_j)$ всех допустимых векторов $p = (p_1, \dots, p_r)$ вероятностей различных альтернатив A_1, \dots, A_r .

По аналогии с [14] путем моделирования неопределенность выбора вектора $p = (p_1, \dots, p_r)$ из множества $P(r; I_j)$, можно определить случайный вектор вероятностей:

$$\begin{aligned} \tilde{p}(I_j) &= (\tilde{p}_1(I_j), \dots, p_r(I_j)), \\ \tilde{p}_i(I_j) &\geq 0, \\ \tilde{p}_1(I_j) + \dots + \tilde{p}_r(I_j) &= 1. \end{aligned} \quad (7)$$

При анализе его компонент, можно сделать вывод, что каждая из них представляет собой стохастическую оценку неизвестной вероятности p_i , полученную на основе ННН-информации I_j от соответствующего эксперта.

ННН-информацию, имеющуюся у исследователя можно представить в виде следующей системы:

$$J = \begin{cases} w_i > w_r, \\ w_u = w_v, \dots; \\ A_l \leq w_l \leq B_l \dots \end{cases} \quad (8)$$

которая использует весовые коэффициенты о сравнительной значимости отдельных источников:

$$\begin{aligned} w_1, \dots, w_m, \\ w_j \geq 0, \\ w_1 + \dots + w_m = 1. \end{aligned} \quad (9)$$

Соответственно, доступную исследователю всю ННН-информацию, можно описать в виде следующего кортежа:

$$\langle I, J \rangle = \langle I_1, \dots, I_m, J \rangle, \quad (10)$$

последняя компонента которого представляет собой определенную систему равенств/неравенств используемых весовых коэффициентов w_1, \dots, w_m .

С точки зрения ННН-информации, введенная информация J , характеризует множество $W(m; J)$ допустимых векторов $w = (w_1, \dots, w_m)$ весовых коэффициентов. Путем моделирования

неопределенности выбора вектора $w = (w_1, \dots, w_m)$ из множества $W(m; J)$ согласно [14], можно определить следующий случайный вектор весовых коэффициентов:

$$\begin{aligned} \tilde{w}(J) &= (\tilde{w}_1(J), \dots, \tilde{w}_m(J)), \\ \tilde{w}_j(J) &\geq 0, \\ \tilde{w}_1(J) + \dots + \tilde{w}_m(J) &= 1, \end{aligned} \quad (11)$$

в котором стохастической оценкой некоторого весового коэффициента w_j , на основе ННН-информации J выступает каждая компонента $\tilde{w}_j(J)$. Получить искомые усредненные оценки весов w_j , $j = 1, \dots, m$ можно путем расчета математического ожидания:

$$\bar{w}_j(J) = E \tilde{w}_j(J) \quad (12)$$

Разброс стохастических оценок компонент вектора $\tilde{w}_j(J)$ справа и слева от их соответствующих усредненных значений $\bar{w}_j(J) = E \tilde{w}_j(J)$ определяется стандартными отклонениями, рассчитываемыми как:

$$\varepsilon_j(J) = \sqrt{D \tilde{w}_j(J)} \quad (13)$$

Таким образом, относительно ННН-информации J вектор усредненных оценок весовых коэффициентов представляет ее есть числовой образ.

При анализе полученных данных можно составить матрицу, строки которой сформированы из стохастических оценок вероятностей альтернатив – $(\tilde{p}_i(I_j))$, $i = 1, \dots, r$, $j = 1, \dots, m$, и представляют собой случайные вектора $\tilde{p}(I_j)$, компоненты которых характеризуют оценки вероятностей альтернатив, по ННН-информации I_j из определенного источника.

Многокритериальную оценку вероятности p_i альтернативы A_i можно получить путем транспонирования столбца матрицы $(\tilde{p}_i(I_j))$ и определения случайного вектора:

$$\tilde{p}^{(i)} = (\tilde{p}_i(I_1), \dots, \tilde{p}_i(I_m)). \quad (14)$$

Сводная оценка вероятности альтернативы A_i , построенная путем линейного взвешенного смоделированного неопределенного выбора вектора агрегирования стохастических оценок $\tilde{p}_i(I_j)$ со случайными весовыми коэффициентами $\tilde{w}_j(J)$, $j = 1, \dots, m$, имеет вид:

$$\tilde{\tilde{p}}_i(I; J) = \sum_{j=1}^m \tilde{p}_i(I_j) \tilde{w}_j(J) \quad (15)$$

Отметим, что в оценке $\tilde{\tilde{p}}_i(I; J)$ вероятности p_i альтернативы A_i заложен комплексный учет всей ННН-информации $(I; J) = (I_1, \dots, I_m; J)$, которая имеется у исследователя.

Математическое ожидание и дисперсию сводной оценки вероятности $\tilde{\tilde{p}}_i(I; J)$ можно рассчитать как:

$$\bar{\tilde{\tilde{p}}}_i(I; J) = E \tilde{\tilde{p}}_i(I; J) = \sum_{j=1}^m \bar{p}_i(I_j) \bar{w}_j(J), \quad (16)$$

$$\begin{aligned} S_i^2(I, J) &= D \tilde{\tilde{p}}_i(I, J) = \sum_{\substack{j, l=1 \\ j \neq l}}^m \bar{p}_i(I_j) \bar{p}_i(I_l) \varepsilon_{jl}(J) + \\ &+ \sum_{j=1}^m [\delta_i^2(I_j) \varepsilon_j^2(J) + \delta_i^2(I_j) \bar{w}_j^2(J) + \bar{p}_i^2(I_j) \varepsilon_j^2(J)], \end{aligned} \quad (17)$$

где $\varepsilon_{jl} = \text{cov}(\tilde{w}_j(J), \tilde{w}_l(J))$ – ковариация стохастических весовых коэффициентов $\tilde{w}_j(J)$, $\tilde{w}_l(J)$, $j \neq l$.

Таким образом, найденные усредненные оценки $\bar{p}_i(I, J)$ вероятностей альтернатив и соответствующие им стандартные отклонения $S_i(I, J) = \sqrt{S_i^2(I, J)} = \sqrt{D\bar{p}_i(I, J)}$, $i = 1, \dots, r$, дают возможность решить задачу оценки вероятностей p_1, \dots, p_r соответствия исследуемого объекта различным альтернативным классам A_1, \dots, A_r с помощью анализа всей имеющейся у исследователя ННН-информации $(I, J) = (I_1, \dots, I_m, J)$, относительно как вероятностей, так и весовых коэффициентов.

Применительно к оценке условных вероятностей в разработанной математической модели на основе БСД с учетом всей имеющейся ННН-информации возможна реализация и получение усредненных оценок с использованием программной среды APIS (рисунок) [15].

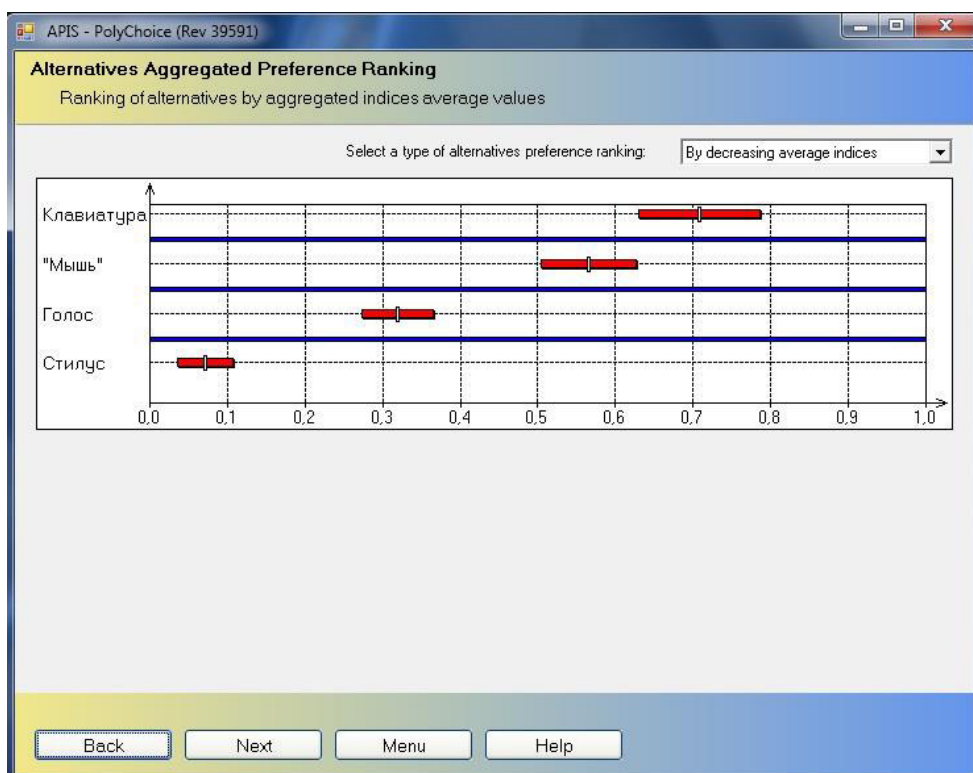


Рис. Пример расчета весовых коэффициентов в среде APIS для многомодальной аутентификации пользователя
Fig. Example of calculating weights in an APIS environment for multimodal user authentication

Заложенный функционал программной среды APIS позволяет использовать и нечисловую, и интервальную информацию по оцениваемым параметрам. В зависимости от используемых входных данных найдена возможность получения оценки условных вероятностей БСД используемой модели при априорной информации о взаимодействии между ее узлами.

Список литературы

1. Никитин В.В. Существующие системы аутентификации и идентификации пользователей: основные проблемы и направления их модернизации // Вестник Московского Университета МВД России, № 2, 2014. – С. 165-172.
2. Отчет об исследовании утечек конфиденциальной информации аналитического центра компании Info Watch за 2016 г. – <http://infowatch.ru/report2016>, дата обращения 23.03.2017 г.
3. Проскурин В.Г., Крутов С.В., Мацкевич И.В. Программно-аппаратные средства обеспечения информационной безопасности. Защита в операционных системах: уч. пособие для вузов. – М.: Радио и связь, 2009, 168.
4. Басов О.О., Карпов А.А., Сайтов И.А. Методологические основы синтеза полимодальных инфокоммуникационных систем государственного управления: монография. – Орёл: Академия ФСО России, 2015, 270.

5. Концептуальная модель многомодального интерфейса абонентского терминала / Струев Д.А., Бондарева Н.В., Будков В.Ю., Басов О.О., Ронжин А.Л. // Научные ведомости Белгородского государственного университета (Экономика, Информатика), № 23 (244), вып. 40, 12, 2016.
6. Korb K.B., Nicholson A.E. Bayesian Artificial Intelligence. New York: Chapman and Hall/CRC, 2004. 364 p.
7. Никитин В.В., Гунченко Ю.И. Модель системы многомодальной аутентификации пользователя на основе байесовской сети доверия // Экономика и менеджмент систем управления. – Воронеж, 2017, № 2.2 (24), 276-282.
8. GeNIe 2.0 Available at: <https://www.bayesfusion.com>.
9. Мусина В.Ф. Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков // Труды СПИИРАН. 2013. Вып. 24. С. 135-151.
10. Хованов Н.В., В.В. Корников, Серегин И.А. Многокритериальное оценивание финансовых рисков в условиях неопределенности. – СПб., СПбГУ, 2002. – 96 с.
11. Hovanov N.V. Discrete variable measurement using non-numerical, inexact, and incomplete information on probabilities' distribution // Proceedings of the International Seminar "Mathematics, Statistics and Computation to Support Measurement Quality". SPb., UNIIM, 2002. P. 20-21.
12. Moshkovich H., Mechitov A., Olson D. Ordinal judgments for comparison of multiattribute alternatives // European Journal of Operational Research. – 2002. – Vol. 137. – Pp. 625-641.
13. Engemann K. J., Yager R. R. A general approach to decision making with interval probabilities // International Journal of General Systems. – 2001. – Vol. 30. – Pp. 623-647.
14. Bayes T. An essay towards solving a problem in the doctrine of chances // Biometrika. – 1958. – Vol. 45. – P. 296-315.
15. Hovanov N. Imprecise Event Trees Framework for Evaluation of Russian Economy Perspectives / Hovanov N., Yudaeva M., Kolesov D. 7th International Conference of the ERCIM WG on Computational and Methodological Statistics University of Pisa, Italy, Dec 6-8, 2014.

References

1. Nikitin V.V. The existing system of authentication and human identification: the main problems and tasks of their modernization // Bulletin of the Moscow University of the Ministry of Internal Affairs of Russia, № 2, 2014. – p. 165-172.
2. Report on the investigation of leaks of confidential information from the analytical center of «InfoWatch», 2016. – <http://infowatch.ru/report2016>, 23.03.2017 г.
3. Proskurin V.G., Krutov S.V., Matskevich I.V. Software and hardware for information security. Protection in operating systems: manual for high schools – Moscow : Radio and communication, 2009, 168.
4. Basov O.O., Karpov A.A., Saitov I.A. Methodological foundations for the synthesis of multimodal infocommunication systems of public administration: monograph – Orel : Akademiya FSO Rossii, 2015, 270.
5. A conceptual model multimodal user interface of the subscriber terminal / Struev D.A., Bondareva N.V., Budkov V.Yu., Basov O.O., Ronzhin A.L. // Belgorod State University Scientific Bulletin (Economics, Information technologies), № 23 (244), Issue 40, December 2016.
6. Korb K.B., Nicholson A.E. Bayesian Artificial Intelligence. New York: Chapman and Hall/CRC, 2004. 364 p.
7. Nikitin V.V., Gunchenko Yu.I. Model of multimodal user authentication system based on Bayesian network of trust // Economics and Management Systems Management. – Voronezh, 2017, № 2.2 (24), 276-282.
8. GeNIe 2.0 Available at: <https://www.bayesfusion.com>.
9. Musina V.F. Bayesian networks of trust as a probabilistic graphic model for assessing medical risks // Proceedings of SPIIRAS. 2013. Vol. 24. p. 135-151.
10. Khovanov N.V., Kornikov V.V., Seregin I.A. Multicriteria evaluation of financial risks in conditions of uncertainty, St. Petersburg, St. Petersburg State University, 2002. – 96.
11. Hovanov N.V. Discrete variable measurement using non-numerical, inexact, and incomplete information on probabilities' distribution // Proceedings of the International Seminar "Mathematics, Statistics and Computation to Support Measurement Quality". SPb., UNIIM, 2002. P. 20-21.
12. Moshkovich H., Mechitov A., Olson D. Ordinal judgments for comparison of multiattribute alternatives // European Journal of Operational Research. – 2002. – Vol. 137. – p. 625-641.
13. Engemann K. J., Yager R. R. A general approach to decision making with interval probabilities // International Journal of General Systems. – 2001. – Vol. 30. – Pp. 623-647.
14. Bayes T. An essay towards solving a problem in the doctrine of chances // Biometrika. – 1958. – Vol. 45. – P. 296-315.
15. Hovanov N., Yudaeva M., Kolesov D. Imprecise Event Trees Framework for Evaluation of Russian Economy Perspectives, 7th International Conference of the ERCIM WG on Computational and Methodological Statistics University of Pisa, Italy, Dec 6-8, 2014.

Никитин Виктор Викторович, аспирант кафедры речевых информационных систем

Гунченко Юлия Игоревна, программист

Басов Олег Олегович, доктор технических наук, профессор кафедры речевых информационных систем

Nikitin Viktor Viktorovich, post-graduate student of Department of Speech Information Systems

Gunchenko Yulia Igorevna, programmer

Basov Oleg Olegovich, Doctor of Technical Sciences, Professor of Department of Speech Information Systems