

SERHII HNATIUK,
OLEKSANDR KORNEIKO

SIMULATION SOFTWARE PERFORMANCE RELIABILITY SPECIAL PROVISION OF COMMUNICATION

A mathematical model of reliability of software that allows you to quantify and predict the values of reliability, which differs from the known increase in the accuracy of the results. The essence of the model is to obtain a quantitative assessment of the reliability of computer systems under given constraints and assumptions on the basis of new functional relationships that reflect the change of values the indicators from time.

Keywords: computer system, software, a model of reliability, recovery time.

Сергій Євгенович Гнатюк, аспірант, Державний заклад «Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут», Київ, Україна.

E-mail: Gnatyk-2@i.ua.

Олександр Васильович Корнейко, кандидат технічних наук, професор, професор кафедри кібербезпеки та застосування автоматизованих інформаційних систем і технологій, Державний заклад «Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут», Київ, Україна.

E-mail: alex_korneiko@meta.ua.

Сергей Евгеньевич Гнатюк, аспирант, Государственное учреждение «Институт специальной связи и защиты информации Национального технического университета Украины «Киевский политехнический институт», Киев, Украина.

Александр Васильевич Корнейко, кандидат технических наук, профессор, профессор кафедры кибербезопасности и применения автоматизированных информационных систем и технологий, Государственное учреждение «Институт специальной связи и защиты информации Национального технического университета Украины «Киевский политехнический институт», Киев, Украина.

Serhii Hnatiuk, postgraduate student, State institution «Institute of special communication and information security of National technical university of Ukraine «Kyiv polytechnic institute», Kyiv, Ukraine.

Oleksandr Korneiko, candidate of technical sciences, professor, professor of cybersecurity and application of information systems and technologies academic department, State institution «Institute of special communication and information security of National technical university of Ukraine «Kyiv polytechnic institute», Kyiv, Ukraine.

УДК 004.93'11

ВИКТОР ЕВЕЦКИЙ

ОЦЕНКА ЭФФЕКТИВНОСТИ ОТДЕЛЬНЫХ ПРИЗНАКОВ И ИХ СОВОКУПНОСТЕЙ ДЛЯ ИНДИВИДУАЛЬНОГО РАСПОЗНАВАНИЯ ОБЪЕКТОВ

В статье рассмотрены подходы к оценке возможностей признаков для целей индивидуального распознавания объектов. На основе предложенного параметра

распознавания, позволяющего количественно оценить информативность признака, а также использовании допущения о нормальных законах распределения признаков и критерия Неймана-Пирсона при принятии решений были получены соотношения, позволяющие оценить результат использования различных совокупностей признаков для индивидуального распознавания объекта. В качестве примера приведен расчет вероятностей правильного распознавания объекта для различных сочетаний нескольких признаков.

Ключевые слова: распознавание объектов, индивидуальное распознавание объектов, признак объекта, пространство признаков, параметр распознавания, вероятность правильного распознавания, вероятность ложного распознавания.

Задача распознавания в настоящее время обычно решается следующим образом. Информацию о связи между значениями признаков объекта и его принадлежностью к определенному классу алгоритм распознавания должен извлечь из обучающей совокупности объектов. Эти объекты в совокупности являются обучающим множеством или выборкой.

Можно выделить следующие основные задачи теории распознавания образов:

1. Математическое описание образов. Наиболее удобным математическим описанием обычно считается векторное описание образов. В этом случае каждому образу x ставится в соответствие некоторый вектор признаков этого образа – элемент векторного пространства. Такое векторное пространство называется пространством признаков.

2. Выбор наиболее информативных признаков, описывающих данный образ. Это одна из основных и важных задач в теории распознавания образов – найти минимальное количество признаков, наиболее информативно описывающих образы в данной системе распознавания.

3. Описание классов распознаваемых образов. Эта задача сводится к определению границ классов. Границы классов могут быть заданы явно на этапе разработки системы распознавания или система сама должна их найти в процессе своей работы.

4. Нахождение оптимальных решающих процедур (методов классификации), т.е. методов соотнесения вектора признаков образа некоторому классу.

5. Оценка достоверности классификации образов.

В зависимости от особенностей информационных компонент системы выделяют три подхода к задаче распознавания образов:

- 1) принцип сравнения с эталоном;
- 2) принцип кластеризации;
- 3) принцип общности свойств [2].

Используемые математические методы опираются на использование классического математического аппарата: методов линейного программирования, корреляционного анализа, теории статистических решений и т.п. Они применяются в тех случаях, когда признаки представлены измерениями (параметрами), а их связи могут быть описаны в виде аналитических зависимостей. В той или иной степени математические методы используются во всех трех подходах к распознаванию.

Индивидуальное распознавание объектов предполагает распознавание данного экземпляра объекта в совокупности точно таких же объектов. Объекты могут иметь самую различную природу, например, конкретный экземпляр радиолокационной станции среди таких же станций, конкретная особь животного в популяции этих животных, конкретный пользователь компьютера и т. д.

В таких условиях трудоемкая задача описания классов и определения границ классов фактически отсутствует, а наибольшее значение приобретает выбор необходимого (минимального) количества признаков обеспечивающего необходимое качество распознавания.

При индивидуальном распознавании объектов наиболее приемлем принцип сравнения с эталоном, поскольку класс объектов заранее известен.

В этом случае процесс распознавания заключается в простом сопоставлении образов, поступающих на вход распознающего устройства или алгоритма, с эталонами на основе выбранной меры сходства. Принцип сравнения с эталоном – один из первых подходов, возникших при построении технических систем распознавания.

При реализации процесса распознавания вначале необходимо определить, по возможности, наиболее полный перечень признаков, в той или иной степени характеризующих распознаваемые элементы среды. Затем следует проанализировать полученный состав признаков на предмет возможности их использования на языке признаков и возможности технической реализации получения информации о них в процессе функционирования распознающей системы. Отвечающие указанным требованиям признаки включаются в состав, так называемого, априорного словаря признаков. Разработанный таким образом априорный словарь признаков используется для предварительного описания задачи на языке признаков и разработки алгоритмов распознавания.

Подавляющее число доступных публикаций посвящены исследованию процедур селекции признаков уже на этапе обучения по конечному числу наблюдений с известными состояниями объекта [1].

Главная цель селекции признаков обычно состоит в переходе от исходного множества признаков N_0 к подмножеству¹ из N признаков ($N < N_0$) без практического уменьшения достоверности распознавания.

После обучения системы распознавания становится возможным также определить для каждого признака его ценность для решения задачи распознавания. Далее, наименее ценные признаки могут быть удалены из системы признаков. Затем система распознавания должна быть обучена заново, так как в результате удаления некоторых признаков статистика распределения оставшихся признаков может изменяться. Этот процесс может повторяться, т.е. быть итерационным.

Выбор адекватного множества признаков представляет собой одну из наиболее трудных проблем, с которой приходится сталкиваться при решении задач распознавания. Исходная совокупность признаков в настоящее время обычно задается субъективно или на основе мнений экспертов соответствующей предметной области, или в исходное описание стремятся включить любой доступный для измерения признак, который может быть хоть как-то полезным.

В теории распознавания образов оценку информативности признака обычно получают как отношение результатов распознавания объектов контрольной выборки в полном пространстве признаков к результатам распознавания, проводимого без учета оцениваемого признака. При этом, очевидно, что оценка информативности признака, может зависеть как от решающего правила, так и от объема обучающей выборки.

Следует отметить, что для устранения зависимости от способов реализации и обучения устройства распознавания важно иметь средства, позволяющие обоснованно подходить к выбору эффективных для распознавания признаков еще до этапа обучения.

Байесовский подход к анализу относится к наиболее разработанным в современной статистике так называемым параметрическим методам, для которых считается известным аналитическое выражение закона распределения (как правило, это нормальный закон) и требуется оценить лишь относительно небольшое количество параметров.

В статье [3] для оценки информативности признака при индивидуальном распознавании предлагалось ввести так называемый параметр распознавания q .

Параметр распознавания

$$q = \frac{\sigma_{po}}{\sigma_{oo}}, \quad (1)$$

где σ_{po} – среднеквадратическое отклонение разброса данного признака у разных объектов в группе однотипных объектов;

σ_{oo} – среднее квадратическое отклонение разброса данного признака у отдельного объекта.

Представляется очевидным, что данный признак тем более ценен для индивидуального распознавания, чем больший разброс он имеет у разных объектов и чем более он стабилен у отдельного объекта.

Можно показать, что условные вероятности правильного распознавания P_{np} и ложного распознавания $P_{л}$ (при принятии решения в соответствии с критерием Неймана-Пирсона и допущении о нормальном законе распределения признаков) определяются следующими соотношениями:

$$P_{np} = 2\Phi\left(\frac{\delta_n}{\sigma_{po}} q\right), \quad (2)$$

$$P_{л} = 2\Phi\left(\frac{\delta_n}{\sigma_{po}}\right), \quad (3)$$

где $\Phi(u) = \frac{2}{\sqrt{2\pi}} \int_0^u e^{-\frac{s^2}{2}} ds$ – функция Лапласа,

δ_n – пороговое значение при принятии решения по используемому признаку.

Для определения значения параметра распознавания q различных признаков необходимо проведение статистического анализа экспериментальных данных.

Практическая ценность предложенного для оценки информативности признака параметра распознавания q состоит в том, что знание его величины позволяет количественно оценить вероятность правильного распознавания объекта при использовании данного признака при заданном уровне ошибки – вероятности ложного распознавания. Это позволяет провести сравнительный анализ информативности отдельных признаков для индивидуального распознавания, а также определить соответствующий набор признаков для распознавания объекта с заданной степенью достоверности.

Рассчитанные вероятности правильного распознавания P_{np} при вероятностях ложного распознавания $P_{л} = 10^{-1}$, $P_{л} = 10^{-2}$, $P_{л} = 10^{-3}$ для разных значений параметра распознавания q приведены в табл. 1-3.

Таблица 1 – Значения вероятности P_{np} при $P_{л} = 10^{-1}$

q	2	5	0	20	30
P_{np}	0,2	0,47	0,79	0,998	0,9998

Таблица 2 – Значения вероятности P_{np} при $P_{л} = 10^{-2}$

q	2	5	10	20	50	70	90	100
P_{np}	0,02	0,05	0,1	0,20	0,47	0,62	0,74	0,77

Таблица 3 – Значения вероятности P_{np} при $P_{л} = 10^{-3}$

q	2	10	20	50	70	90	100
P_{np}	0,0005	0,0096	0,19	0,048	0,067	0,086	0,096

Для совокупности признаков могут быть получены следующие соотношения. Обозначим отношение $\frac{\delta_n}{\sigma_{po}}$ через a_i , где i – номер соответствующего признака в совокупности признаков. Тогда для оценки вероятности правильного распознавания данного объекта в группе однотипных при использовании двух признаков со значениями параметра

распознавания q_1 и q_2 , а также заданной условной вероятностью ложного распознавания $P_{\text{л}}$ можно получить:

$$\Phi(a_1) = \sqrt{\frac{P_{\text{л}} q_2}{4 q_1}}; \quad \Phi(a_2) = \sqrt{\frac{P_{\text{л}} q_1}{4 q_2}}; \quad (4)$$

$$P_{\text{np}} = 4 \Phi(a_1 \cdot q_1) \Phi(a_2 \cdot q_2) \quad (5)$$

Для оценки вероятности правильного распознавания данного объекта в группе при использовании трех признаков со значениями параметра распознавания q_1 , q_2 , q_3 и заданной условной вероятностью ложного распознавания $P_{\text{л}}$ можно также получить:

$$\Phi(a_2) = \sqrt[3]{\frac{P_{\text{л}} q_1 \cdot q_3}{8 q_2^2}} \quad \Phi(a_1) = \sqrt[3]{\frac{P_{\text{л}} q_2 \cdot q_3}{8 q_1^2}} \quad \Phi(a_1) = \sqrt[3]{\frac{P_{\text{л}} q_1 \cdot q_2}{8 q_3^2}} \quad (6)$$

$$P_{\text{np}} = 8 \Phi(a_1 \cdot q_1) \Phi(a_2 \cdot q_2) \Phi(a_3 \cdot q_3) \quad (7)$$

Оценим в качестве примера возможности для индивидуального распознавания объекта различных совокупностей признаков со следующими значениями параметров распознавания:

$$q_1 = 50, \quad q_2 = 4, \quad q_3 = 3.$$

Зададим вероятность ложного распознавания $P_{\text{л}} = 10^{-2}$. Используя полученные соотношения (1)...(7), имеем следующие результаты.

1. Используемые по отдельности признаки обеспечивают:

первый признак ($q_1 = 50$) $P_{\text{np}} = 0,47$;

второй признак ($q_2 = 4$) $P_{\text{np}} = 0,04$;

третий признак ($q_3 = 3$) $P_{\text{np}} = 0,03$.

2. Совокупность первого и второго признаков может обеспечить $P_{\text{np}} = 0,86$.

3. Совместное использование достаточно информативного первого признака в совокупности с малоинформативными вторым и третьим признаками позволяет получить $P_{\text{np}} = 0,95$.

Таким образом, могут количественно оцениваться возможности разных совокупностей признаков для целей индивидуального распознавания объектов.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Файнзильберг Л. С. Математические методы оценки полезности диагностических признаков / Л. С. Файнзильберг. – К. : Освита Украины, 2010. – 152 с.

2. Чабан Л. Н. Теория и алгоритмы распознавания образов : учебное пособие / Л. Н. Чабан. – М. : МИИГАиК, 2004. – 70 с.

3. Евецкий В. Л. Повышение эффективности защиты данных за счет биометрической идентификации пользователя / В. Л. Евецкий // Information Technology and Security. – 2012. – № 2 (2). – С. 14-16.

Статья поступила в редакцию 21.10.2015.

REFERENCE

1. Fainzilberg, L. S. (2010), *Matematicheskie metody otsenki poleznosti diagnosticheskikh priznakov* [Mathematical methods for evaluating the usefulness of diagnostic features], Osvita Ukrainy Publ., Kyiv, 152 p.

2. СHaban, L. N. (2004), *Teoriia i algoritmy raspoznavaniia obrazov* [Theory and algorithms of pattern recognition], МПГАiК Publ., Moscow, 70 p.

3. Yevetskyi, V.L. (2012), *Povyshenie effektivnosti zashchity dannykh za schet biometricheskoi identifikatsii polzovatel'ia* [Improvement of the efficiency of data protection by biometric identification], Information technology and security, No. 2 (2), pp. 14-16.

ВІКТОР ЛЕОНІДОВИЧ

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ОКРЕМИХ ОЗНАК ТА ЇХ СУКУПНОСТЕЙ ДЛЯ ІНДИВІДУАЛЬНОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

У статті розглянуті підходи до оцінки можливостей ознак для цілей індивідуального розпізнавання об'єктів. На основі запропонованого параметра розпізнавання, що дозволяє кількісно оцінити інформативність ознаки, а також використанні припущення про нормальні закони розподілу ознак і критерію Неймана-Пірсона при прийнятті рішень були отримані співвідношення, що дозволяють оцінити результат використання різних сукупностей ознак для індивідуального розпізнавання об'єкта. Як приклад наведено розрахунок ймовірностей правильного розпізнавання об'єкта для різних сполучень декількох ознак.

Ключові слова: розпізнавання об'єктів, індивідуальне розпізнавання об'єктів, ознака об'єктів, простір ознак, параметр розпізнавання, ймовірність правильного розпізнавання, ймовірність помилкового розпізнавання.

VIKTOR YEVETSKYI

EVALUATION OF CERTAIN SIGNS AND THEIR COLLECTIVELY FOR INDIVIDUAL RECOGNITION OBJECTS

The article discusses approaches to possibility estimation of various features for the individual object recognition. On the basis of the proposed recognition parameter, which allows to estimate informatively of feature, as well as the use of assumptions about normal distribution of features and Neyman-Pearson criterion the formulas were obtained to assess the result of using of different feature groups for the individual object recognition. As an example the results of calculation the probabilities of object correct recognition for variety combinations of several features are given in the article.

Keywords: object recognition, the individual object recognition, object attribute, feature space, recognition parameter, the probability of correct recognition, the probability of false detection.

Виктор Леонидович Евецкий, кандидат технических наук, доцент, доцент кафедры, Государственное учреждение «Институт специальной связи и защиты информации Национального технического университета Украины «Киевский политехнический институт», Киев, Украина.

E-mail: viktorevetsky@gmail.com.

Віктор Леонідович Евецкий, кандидат технічних наук, доцент, доцент кафедри, Державний заклад «Інститут спеціального зв'язку та захисту інформації Національний технічний університет України «Київський політехнічний інститут», Київ, Україна.

Viktor Yevetskyi, candidate of technical sciences, associate professor, associate professor of academic department, State institution «Institute of special communication and information security of National technical university of Ukraine «Kyiv polytechnic institute», Kyiv, Ukraine.