

*Is designed the computer simulator of fault message diagnostic on the interference background in MATLAB/Simulink environment. The detection, localization and identification of faults are illustrated by marine turbine control system.*

**Vibrodiagnostics, fault model, signal graphs, frequency response, fault localization, the regularization of ill-posed problems**

УДК 681.518.54:572.983

*А. Ю. Дорогов, В. С. Абатуров*

## **ИЗМЕРЕНИЕ ОБОБЩАЮЩЕЙ СПОСОБНОСТИ НЕЙРОННЫХ СЕТЕЙ**

*Представлена системная модель процесса распознавания образов. Определено понятие невырожденной схемы распознавания. Предложен метод измерения обобщающей способности нейронных сетей, пригодный при малых объемах обучающих выборок. Приведена методика проведения эксперимента. Показаны результаты измерения обобщающей способности для нейронных сетей прямого распространения.*

**Нейронная сеть, обобщающая способность, разделяющая мощность, системная модель**

Определяющей оценкой качества нейронных сетей прямого распространения является их способность к распознаванию образов. Нередко приходится сталкиваться с тем обстоятельством, что нейронная сеть безошибочно распознает образы обучающей выборки, но ошибается на образах контрольной выборки. В этом случае говорят, что нейронная сеть имеет недостаточную обобщающую способность. Вопросы оценки обобщающей способности для задач распознавания образов исследовались в [1],\* где обобщающая способность оценивалась вероятностью правильного распознавания образа, не принадлежащего обучающей выборке. Полученные результаты оказались весьма приближенными и для экспериментального подтверждения требовали больших объемов обучающих выборок. В данной статье рассматриваются детерминистские методы экспериментальной оценки обобщающей способности нейронных сетей с использованием обучающих выборок малого объема.

**Системная модель процесса распознавания образов.** Обобщенная модель системы распознавания образов [2] представлена на рис. 1. Элементам модели сопоставляются следующие смысловые понятия:  $U$  – множество распознаваемых образов (генеральная совокупность);  $D$  – множество информативных признаков;  $C$  – множество классов образов. Тогда отображение  $\mu$  можно интерпретировать как систему измерения информативных признаков, отношение  $\rho$  – как решающее правило, а отношение  $\varepsilon$  – как экспертную классификацию образов.

В практических реализациях система распознавания обучается на конечном наборе образов  $t \subset U$ , который является некоторой выборкой из генеральной совокупности. В этом случае мнение экспертов определяется отношением  $\varepsilon_t \subset \varepsilon$ . Будем полагать, что

---

\* Воронцов К. В. Лекции по теории обобщающей способности. 21 декабря 2007 г. <http://www.ccas.ru/voron/download/Generalization.pdf>.

система идеально обучается на выборке, т. е.  $\varepsilon_t = \mu r_t$ , где  $r_t \subset r$ . Когда обученной системе предъявляется образ, не вошедший в обучающую выборку, и система распознает этот образ согласно глобальному мнению экспертов  $\varepsilon$ , то говорят, что обученная система обладает свойством обобщения. Максимальное обобщение достигается при  $r_t = r$ .

Системная модель позволяет формализовать требования к полноте и достаточности схемы распознавания образов. Нотацию  $\mu^\#, \rho^\#, \varepsilon^\#$  будем использовать для обозначения обратных отношений, а  $1_U, 1_C, 1_D$  – для обозначения тождественных отношений. Предположим, что для генеральной совокупности образов экспертиза является идеальной (т. е. является полной проекцией  $\varepsilon \varepsilon^\# \supset 1_U, \varepsilon^\# \varepsilon = 1_C$ ), а решающее правило – безошибочным:  $\varepsilon = \mu r$ . Сделав естественные предположения об однозначности измеряющего отображения ( $\mu^\# \mu \subset 1_D$ ) и полноте использования признакового пространства решающим правилом ( $r r^\# \supset 1_D$ ) при идеальной экспертизе получим:

$$1_U \subset \varepsilon \varepsilon^\# = \mu r r^\# \mu^\# \Rightarrow \mu \mu^\# \supset 1_U,$$

$$1_C = \varepsilon^\# \varepsilon = r^\# \mu^\# \mu r \Rightarrow r^\# r = 1_C.$$

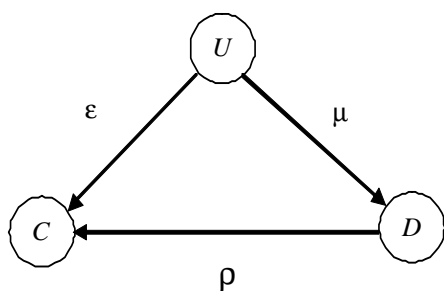


Рис. 1

Таким образом, при обучении на генеральной совокупности образов измеряющее отображение является функциональным, а решающее правило – полной проекцией. Такая схема распознавания называется невырожденной, или типичной [2].

**Методика измерений.** Существует проблема выбора тестового набора образов для испытания систем распознавания. Прежде всего следует ответить на вопрос:

«Что является объектом распознавания?» В зависимости от ответа выбирается модель пространства образов. В окружающем человека мире объектами распознавания, как правило, являются геометрические фигуры. В этом случае наиболее подходящей математической моделью мира является аффинное пространство. В аффинном пространстве учитывается только взаимное размещение точек фигуры и игнорируются абсолютные координаты ее размещения. Когда образы представлены точками аффинного пространства, то в качестве тестового набора стандартно используют образы в виде точек «общего положения» [3]. По определению, множество, состоящее из  $N$  точек  $n$ -мерного аффинного пространства, называют множеством точек общего положения, если ни одно из его подмножеств, состоящее из  $m+1$  точек ( $m \leq N$ ), не лежит на  $(m-1)$ -мерной гиперплоскости. Если одну из точек множества выбрать в качестве начала координат (обозначим эту точку через  $O$ ), то положение остальных точек можно задать векторами  $\{x^1, x^2, \dots, x^m\}$ , соединяющими их с началом координат. Другими словами, система векторов образует  $m$ -мерный репер с началом в точке  $O$ . Несложно показать, что векторы репера являются линейно независимыми, а множество образов – линейно разделимым.

Для оценки обобщающей способности образуем классы, окружив каждую точку общего положения сферическим облаком точек (рис. 2). Центр сферы будем считать эталонным классом. Выберем радиус сферы для каждого эталона таким образом, чтобы обеспечить экспертную разделимость классов, полагая, что мнение эксперта определяется по минимальному расстоянию между образом генеральной совокупности и центрами классов. В этом случае радиус сферы, окружающей эталон, не превышает половины минимального расстояния до ближайшего соседнего эталона. Это значение используется как база нормирования для радиуса сферы выбранного эталона. При экспериментальном исследовании относительный радиус всех окружающих сфер принимается одинаковым и изменяется в диапазоне от 0 до 1.

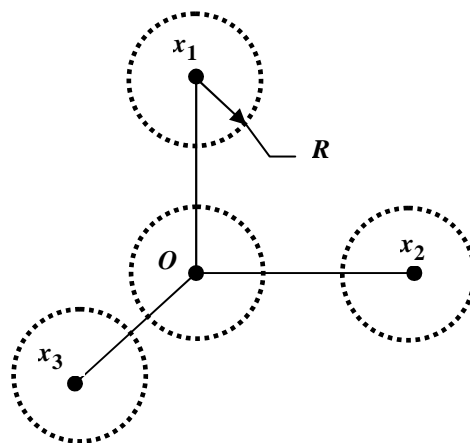


Рис. 2

Реализуем невырожденную систему распознавания на основе нейронной сети и решающего правила. В системной модели нейронная сеть выполняет функциональное преобразование образа в аффинное пространство признаков, выполняя роль схемы измерения. Измерение в аффинном пространстве признаков заключается в выборе начала координат, задании репера и определении значений координат образов в выбранном репере. Репер в пространстве признаков может быть произвольным, но для обеспечения однородности экспериментов целесообразно выбрать его ортонормированным с привязкой к началу координат (в простейшем случае это может быть унитарный код). Нейронная сеть обучается на подмножестве точек общего положения пространства образов таким образом, чтобы каждому эталону класса соответствовал один вектор репера. Размер подмножества определяется из условия безошибочного обучения нейронной сети (максимальный размер обучающего подмножества равен разделяющей мощности испытываемой нейронной сети).

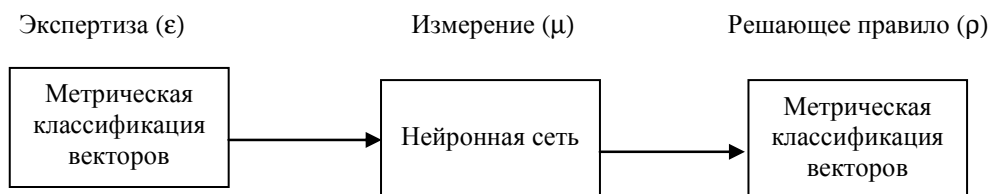


Рис. 3

Схема измерения обобщающей способности показана на рис. 3.

Решающее правило поддерживается метрической схемой классификации, которая по минимальному расстоянию между паттерном и точками репера формирует результат распознавания. Обобщающая способность оценивается максимальным относительным радиусом сферических облаков, при котором обеспечивается безошибочное распознавание всех точек классов. Максимальная обобщающая способность не превышает единицы, минимальная – равна нулю.

**Результаты экспериментов.** Для формирования входного обучающего набора использовались векторы ортогонального косинусного преобразования. Ортогональный репер привязывался к началу координат. Нулевая точка использовалась в обучающем множестве как один из образов. Векторы выходного репера были представлены унитарным кодом. Облака классов содержали по 2000 точек, которые случайно размещались на сферах заданного радиуса. Количество точек в облаке выбиралось из условия надежности фиксации факта безошибочного распознавания. Относительные радиусы сфер совпадали для всех классов. Эксперименты проводились для однослойной и двухслойной нейронных сетей.

1. *Однослойная нейронная сеть.* Размерность сети по входу и выходу равнялась восьми (рис. 4). Размерность обучающего репера изменялась от 1 до 8. Обучение нейронной сети по реперу проводилось при помощи алгоритма ErrorBackPropagation без накопления градиента. Обученной сети предъявлялось полное множество точек окружающих сфер, и выполнялся контроль ошибок распознавания. В процессе проведения экспериментов радиусы сфер уменьшались до исчезновения ошибок распознавания. Найденное значение радиуса принималось за оценку обобщающей способности при заданном размере обучающего репера.

При повторном обучении в тех же условиях значение обобщающей способности могло изменяться в пределах 10 %. Таким образом, наблюдалась небольшая зависимость обобщающей способности от стартовой точки процедуры обучения.

Эксперименты проводились для сети с линейной, тангенциальной и сигмоидной функциями активации. Результаты для одного из экспериментов представлены на рис. 5. Можно отметить, что обобщающая способность нейронной сети с линейными функциями активации оказалась существенно выше, чем сети с функциями активации в виде гиперболического тангенса и сигмоида.

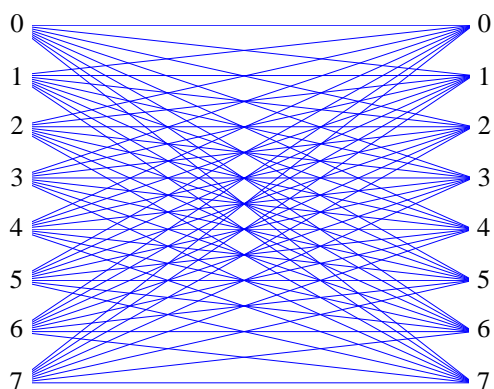


Рис. 4

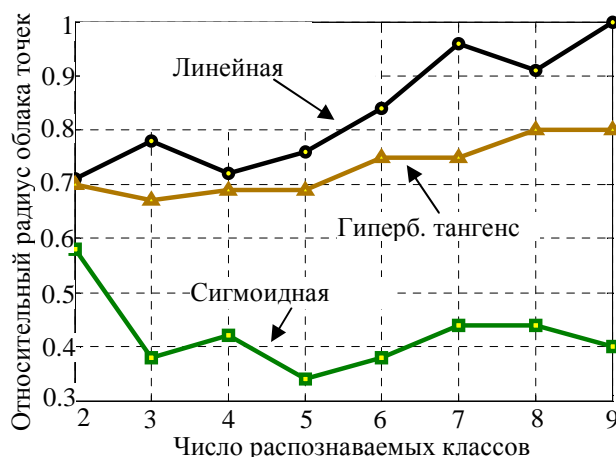


Рис. 5

2. *Двухслойная нейронная сеть.* Эксперименты проводились на двухслойной нейронной сети прямого распространения «с узким горлом» (рис.6, а). Размерность сети по входу и выходу равнялась 8. Исследовались 2 варианта нейронных сетей, отличающихся функциями активации. В первом варианте оба слоя сети имели линейные функции активации, а во втором – первый слой имел линейную функцию, второй – сигмоидную. Число нейронов в скрытом слое ( $D$ ) варьировалось от 1 до 9. Число классов в экспериментах выбиралось на верхней границе разделяющей мощности линейной нейронной сети (это на единицу больше числа нейронов в скрытом слое). Для обучения линейной нейронной сети

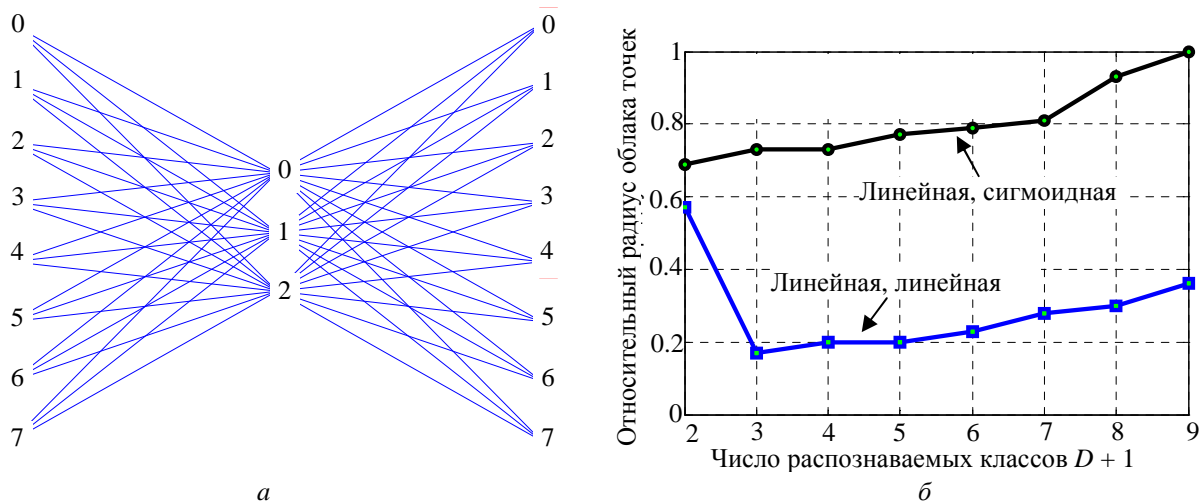


Рис. 6

использовался алгоритм ErrorBackPropagation без накопления градиента, а для нелинейной – алгоритм PROP [4]. На рис. 6, б показана зависимость обобщающей способности от числа достоверно распознаваемых классов.

Во втором эксперименте исследовалась зависимость обобщающей способности от числа классов для двухслойной нейронной сети с числом нейронов в скрытом слое, равным двум. Функции активации: первый слой – линейная, второй слой – сигмоидная. Нижняя граница разделяющей мощности для данной сети равна 3, но за счет нелинейных свойств сеть безошибочно обучалась на реперах вплоть до 8-го порядка. Для обучения нейронной сети использовался алгоритм PROP. На рис. 7, а изображена топология сети, а результаты эксперимента показаны на рис. 7, б. Из графика видно, что с увеличением числа достоверно распознаваемых классов обобщающая способность нейронной сети падает практически до нуля.

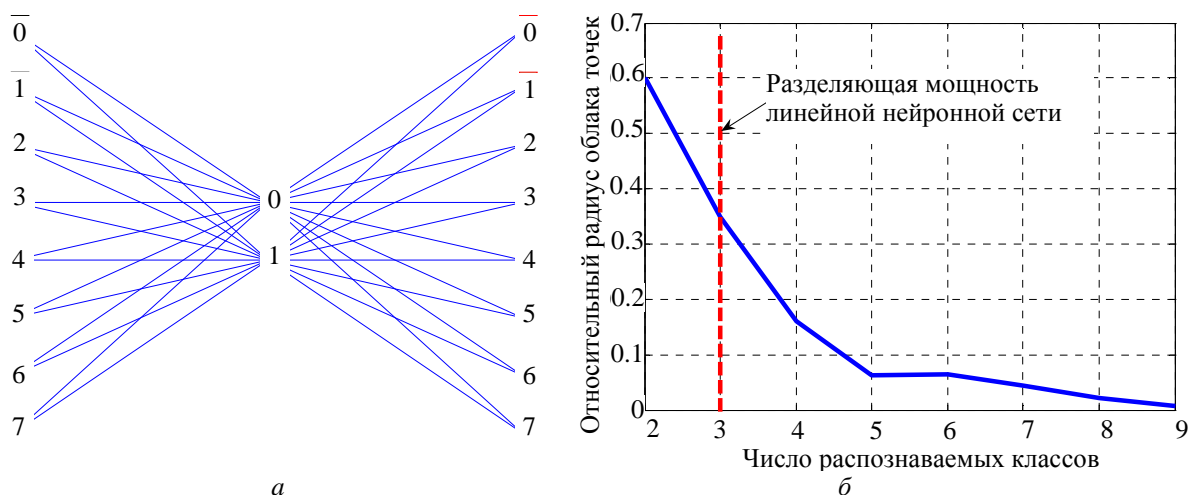


Рис. 7

Таким образом, предложенный метод оценки обобщающей способности не требует обучения нейронной сети на выборках большого объема. Размер обучающей выборки может быть ограничен нижним пределом разделяющей мощности нейронной сети. Результаты экспериментов обладают повторяемостью, достаточной для практических целей. Принцип измерения обобщающей способности имеет наглядную геометрическую интерпретацию, что позволяет получить априорную оценку конфигурации множества достоверно распознаваемых образов нейросетевого классификатора.

## СПИСОК ЛИТЕРАТУРЫ

1. Вапник В. Н., Червоненкис А. Я. Теория распознавания образов. М.: Наука, 1974.
2. Гисин В. Б., Цаленко М. Ш. Алгебраическая теория систем и ее приложения: Системные исследования. Методологические проблемы // Ежегодник. М.: Наука, 1984. С.130–151.
3. Гонсалес Р., Дж. Ту. Принципы распознавания образов. М.: Мир, 1978.
4. Riedmiller M., Braun H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm // In Proc. of the IEEE Intern. Conf. on Neural Networks (ICNN), San Francisco, 1993. P. 586–591.

*A. Yu. Dorogov, V. C. Abaturvov*

### *MEASURING OF GENERALIZATION CAPACITY FOR NEURAL NETWORKS*

*System model of recognition pattern process is presented. Conception of nondegenerated recognition scheme is introduced. Method of generalization capacity measuring for neural networks, which is suitable for small value of training set, is suggested. Methodology of the experiments is given. Results of the generalization capacity measuring for feed forward neural networks are shown.*

**Neural network, generalization capacity, separating power, system model**