СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕЙРОСЕТЕВЫХ И СТАТИСТИЧЕСКИХ АЛГОРИТМОВ В ЗАДАЧАХ ОБНАРУЖЕНИЯ СИГНАЛОВ

Перов А.И., Соколов Г.Г.

Московский энергетический институт (технический университет)
Учебно-исследовательский центр “Радиоэлектронные технологии в ТЭК”
111250, Москва, ул. Красноказарменная, 14,

В классической теории цифровой обработки сигналов, для описания последних часто используется аппарат теории случайных процессов. Особенно широко такое описание используется в задачах радиолокации, радионавигации, радиоуправления, статистической теории связи, статистической теории оптимального управления. В настоящее время достаточно хорошо разработан математический аппарат синтеза оптимальных устройств обработки случайных сигналов [1], который широко используется в задачах обнаружения и различения временных и пространственно-временных сигналов, оценке параметров сигналов, фильтрации сглаживания, экстраполяции информационных процессов и др. С другой стороны, в последние годы интенсивно развиваются и пропагандируются нейросетевые методы и алгоритмы обработки сигналов, которые также рекомендуются для решения указанных выше задач. При этом часто не проводится детального и корректного сравнения характеристик качества решаемых задач при использовании предлагаемых нейросетевых алгоритмов с аналогичными характеристиками, полученными при использовании известных статистических алгоритмов, а если такое сравнение и проводится, то, как правило, оказывается, что нейросетевые алгоритмы имеют лучшие характеристики, чем соответствующие статистические алгоритмы обработки (например, [2]) .

Целью доклада является проведение сравнительного анализа характеристик нейросетевых алгоритмов и статистических алгоритмов цифровой обработки сигналов применительно к задаче обнаружения сигнала, наблюдаемого на фоне случайной помехи.

Ввиду ограниченности объема доклада рассматривается простейшая задача обнаружения детерминированного сигнала, наблюдаемого на фоне белого гауссовского шума. Данная задача взята потому, что для нее известно строгое аналитическое решение в рамках статистической теории и можно получить строгое решение для нейросетевого алгоритма. Не смотря на достаточно частную рассматриваемую задачу, из нее можно сделать ряд общих выводов относительно соответствия нейросетевых и статистических алгоритмов обработки сигналов, о тех областях, где тот или иной алгоритм может давать лучшие характеристики, о том, как надо обучать нейросеть при ее использовании в задачах обнаружения сигналов. Некоторые из общих выводов подтверждаются результатами моделирования более сложных задач обнаружения сигналов.

Пусть наблюдается выборка $\{y_k\}$, $k = 1 \ldots K$ вида

$$y_k = \vartheta S_k + n_k, \tag{1}$$

где y_k – k-й отсчет выборки, S_k – известные значения сигнала в k-м отчете, n_k – дискретный белый гауссовский шум с нулевым математическим ожиданием и дисперсией σ_n^2, ϑ – случайный параметр, принимающий значения 0 и 1 с априорными вероятностями p_0 и p_1 соответственно.

Статистическая теория обнаружения дает следующий оптимальный алгоритм обработки входных сигналов [1]

$$\xi_K = \frac{1}{\sigma_n^2} \sum_{i=1}^{K} y_k S_k > h, \tag{2}$$

где h - порог, значение которого зависит от выбранного критерия оптимизации. Например, для критерия максимума апостериорной вероятности $h = \sum_{i=1}^{K} S_i^2 \Big/ 2\sigma_n^2 + \ln\left(p_0/p_1\right)$. Для других критериев оптимизации алгоритм обработки (2) не меняется, а меняется лишь значение порога h. Таким образом, в рассматриваемой задаче обнаружения детерминированного сигнала и гауссовской помехи при любых критериях оптимизации оптимальный алгоритм обработки (2) включает линейную обработку входных отсчетов и сравнение с порогом.

Рассмотрим нейросетевые алгоритмы в задаче обнаружения сигнала (1). В общем случае нейросеть является нелинейной системой [3]. Однако, учитывая, что статистическая теория обработки сигналов дает линейный оптимальный алгоритм обработки, рассмотрим сначала линейную нейросеть, приведенную на рис. 1.

Для обучения нейросети используется N реализаций вида (1). При этом будем полагать, что в $N_0 = p_0 N$ реализациях параметр $\vartheta = 0$, а в $N_1 = p_1 N$ реализациях $\vartheta = 1$.

Рассмотрим линейную нейросеть с K входами. Обычно в нейросети входные сигналы подвергаются масштабированию. Такая процедура существенная для нелинейной системы. Для линейной системы масштабирование не играет принципиального значения, но несколько усложняет математические выкладки. Поэтому ниже будем рассматривать нейросеть без предварительного масштабирования входных сигналов.

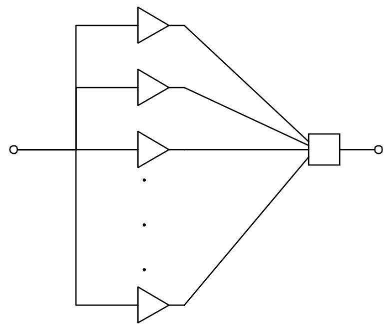

Рис.1. Линейная нейросеть

В процессе обучения нейросети по i -й реализации на k -й вход подается отсчет (2). Сигнал на выходе нейросети описывается соотношением

$$u = \sum_{k=1}^{K} c_k y_k + c_0 , \tag{3}$$

где $\mathbf{c} = \left| c_0 \; c_1 \; \ldots \; c_K \right|^{\mathrm{T}}$ – вектор коэффициентов, подлежащих оптимизации в процессе обучения.

Введем входной вектор $\mathbf{y} = \left| 1 \; y_1 \; y_2 \; \ldots \; y_K \right|^{\mathrm{T}}$. Тогда (3) может быть записано в форме скалярного произведения

$$u = \hat{\vartheta} = \mathbf{c}^{\mathrm{T}} \mathbf{y} . \tag{4}$$

В процессе обучения нейросети минимизируется среднеквадратическая функция

$$R = \frac{1}{N} \sum_{i=1}^{N} \left(\vartheta - \hat{\vartheta} \right)^2 = . \tag{5}$$

Так как в обучении присутствуют два типа реализаций (4), для которых соответственно $\vartheta = 1$ и $\vartheta = 0$, то (5) конкретизируется следующим образом

$$R = \frac{1}{N} \left[\sum_{i=1}^{N_0} \left(\mathbf{c}^{\mathrm{T}} \mathbf{y}_i^{(0)} \right)^2 + \sum_{j=1}^{N_1} \left(1 - \mathbf{c}^{\mathrm{T}} \mathbf{y}_j^{(1)} \right)^2 \right], \tag{6}$$

где $\mathbf{y}_i^{(1)}$, $\mathbf{y}_i^{(0)}$ соответствуют обучающим выборкам при наличии и отсутствии сигнала S_k, число которых равно N_1 и N_0.

Для линейной нейросети минимизацию (6) можно выполнить аналитически, что дает следующие соотношения для оптимальных коэффициентов

$$c_{0,opt} = \tilde{p}_1 \left(1 - \left(\mathbf{S}^{\mathrm{T}} \tilde{\mathbf{c}}_{opt} \right) \right), \;\; \tilde{\mathbf{c}}_{opt} = \frac{\tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right)}{\left(1 + \tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right) \right) E_S} \mathbf{S} , \tag{7}$$

где использованы обозначения

$$\tilde{\mathbf{c}} = \left| c_1 \; \ldots \; c_K \right|^{\mathrm{T}}, \;\; \mathbf{S} = \left| S_1 \; S_2 \; \ldots \; S_K \right|^{\mathrm{T}}, \;\; E_S = \sum_{k=1}^{K} S_k^2 , \;\; \tilde{p}_1 = \frac{N_1}{N} , \;\; \tilde{q} = \sum_{k=1}^{K} S_k^2 \Big/ \tilde{\sigma}_n^2 , \;\; \tilde{\sigma}_n^2 = \frac{1}{m} \sum_{i=1}^{m} n_i^2 .$$

При получении формул (7) использованы допущения

$$\tilde{m}_n = \sum_{i=1}^{m} n_i \Big/ m \approx 0 , , \;\; \tilde{r}_n^2 = \sum_{i=1}^{m} n_i \, n_{j \neq i} \Big/ m \approx 0 ,$$

Определим алгоритм работы обнаружителя следующим образом: решение о наличии сигнала принимается, если $\hat{\vartheta} \geq \frac{1}{2}$, что с учетом (4), (7) дает

$$\hat{\vartheta} = \tilde{\mathbf{y}}^{\mathrm{T}} \tilde{\mathbf{c}}_{opt} + c_{0,opt} = \tilde{\mathbf{y}}^{\mathrm{T}} \mathbf{S} \frac{\tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right)}{\left(1 + \tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right) \right) E_S} + \frac{\tilde{p}_1}{\left(1 + \tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right) \right)} \geq \frac{1}{2} , \tag{8}$$

где $\tilde{\mathbf{y}} = \left| y_1 \; y_2 \; \ldots \; y_K \right|^{\mathrm{T}}$.

Алгоритм (8) можно преобразовать к виду, аналогичному (2)

$$\frac{\tilde{\mathbf{y}}^{\mathrm{T}}\mathbf{S}}{\sigma_n^2} \ge \frac{q}{2} + \frac{\tilde{p}_0 - \tilde{p}_1}{2\tilde{p}_1\left(1-\tilde{p}_1\right)} \cdot \frac{q}{\tilde{q}} = h, \qquad (9)$$

где $q = \dfrac{E_S}{\sigma_n^2}$ – отношение сигнал/шум на входе системы обработки, $\tilde{p}_0 = \dfrac{N_0}{N}$.

Таким образом получаем, что нейросетевой алгоритм при принятой процедуре обучения (6) фактически реализует классическое правило обнаружения. Однако величина порога h несколько иная. Причем, первое слагаемое в соотношении для порога $q/2$ такое же как и в обнаружителях, синтезированных по критерию максимума апостериорной вероятности. Различие лишь во втором слагаемом, зависящем от вероятностей наличия или отсутствия сигнала. Для нейросети эти вероятности определяются процессом обучения сети.

Моделирование нейросетевого алгоритма с использованием пакета *STATISTICA Neural Networks* фирмы *StatSoft*, проведенное для сигнала $S_k = A\cos\left(2\pi k/K + \varphi_0\right)$, подтвердило факт настройки весовых коэффициентов в соответствии с формулами (7), а, следовательно и реализацию в итоговом нейросетевом алгоритме правила обнаружения (9). При этом характеристики обнаружения (вероятность правильного обнаружения, вероятность ложной тревоги) в нейросетевом алгоритме не лучше, чем в статистическом алгоритме при одной и той же величине порога h. Термин “не лучше” здесь употреблен потому, что в процессе реального обучения нейросети весовые коэффициенты после обучения соответствуют формулам (7) с некоторой погрешностью, которая зависит от объема обучающих данных и приводит к некоторой неоптимальной нейросети и, как следствие, к ухудшению точностных характеристик.

Следует отметить одну принципиальную разницу между нейросетевым и статистическим алгоритмом. В статистическом алгоритме величина порога h задается априорно, а в нейросетевом алгоритме она определяется процессом обучения, т.к. в (9) правая часть зависит от $\tilde{p}_0, \tilde{p}_1, \tilde{q}$, которые определяются режимом обучения. Если в режиме обучения менять характеристики N_0, N_1, или вид распределения помехи, то будет меняться и значение порога, а, следовательно, и характеристики обнаружения. Отсюда следует важный вывод о том, что априорная информация о статистических характеристиках входных сигналов, которая используется в статистической теории обработки сигналов как исходная в процессе синтеза оптимального алгоритма, в нейросетевых алгоритмах обработки должна корректно использоваться на стадии обучения нейросети. Другим следствием является то, что, если априорная информация о статистических характеристиках входных сигналов отсутствует, то неопределенным становится и корректная процедура обучения сети. Конечно, сеть в процессе любого обучения чему-то научится. Однако, насколько это соответствует поставленным целям и решаемой задаче остается неопределенным. Поэтому точностные и иные количественные характеристики нейросетевых алгоритмов не всегда могут быть лучше характеристик известных алгоритмов.

Использование в рассматриваемой задаче нелинейных нейросетей приводит к ухудшению характеристик таких устройств.

В докладе рассматривается еще одна задача обнаружения - обнаружение сигнала $S_k = A\cos\left(2\pi k/K + \varphi_0\right)$ с неизвестной (случайной) начальной фазой φ_0, наблюдаемого на фоне белого гауссовского шума. Статистическая теория для такой задачи дает известный алгоритм линейного детектирования огибающей сигнала и сравнение с порогом [1]. Такой алгоритм обработки является нелинейным, поэтому нейросетевой алгоритм также искался в классе нелинейных. В результате статистического обучения получена структура нейросетевого алгоритма и исследованы его характеристики, которые оказались хуже, чем для статистического алгоритма обработки. Отсюда можно сделать вывод о том, что, если входной сигнал имеет действительно статистический характер и его статистическое описание известно и таково, что допускает решение задачи синтеза оптимального алгоритма обработки статистическими методами, то данные методы и следует использовать, т.к. нейросетевые методы и алгоритмы в таких задачах не дадут лучшего результата.

Литература

1. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. - М.: Радио и связь, 1991. - 608 с.
2. Истратов А.Ю., Мельник А.В., Грибков В.Ф. Эмпирический нейроалгоритм обработки радиолокационной информации//5 Всероссийская конференция “Нейрокомпьютеры и их применение”. 1999. С. 228-233.
3. Галушкин А.И. Теория нейронных сетей. - М.: ИПРЖР, 2000. - 416 с.

THE COMPARATIVE ANALYSIS OF NEURAL AND STATISTICAL ALGORITHMS IN PROBLEN OF SIGNAL DETECTION

Perov A.I., Sokolov G.G.

Moscow power energy institute (technical university)
111250, Moscow, Krasnokazarmennaya Str., 14,Training-recerch centre
"Radioelectronic technologies in FEB"

The purpose of the report is the realization of the comparative analysis of characteristics of neural and statistical algorithms of digital signals processing with reference to the problem of detection of a signal observable on a background of a random noise.

We shall consider the elementary problem of detection of the deterministic signal observable on a background of white gaussian noise. There is observed a sample $\{y_k\}$, $k = 1 \ldots K$ of the kind

$$y_k = \vartheta S_k + n_k \,, \tag{1}$$

where y_k - the sample readout with number k, S_k - known value of the signal in the sample readout with number k, n_k - discrete white gaussian noise with zero mathematical expectation and dispersion σ_n^2, ϑ - parameter accepting values 0 and 1 with apriory probabilities p_0 and p_1 accordingly.

The statistical theory of detection gives the following optimal algorithm of processing of input signal [1]

$$\xi_K = \frac{1}{\sigma_n^2} \sum_{i=1}^{K} y_k S_k > h \,, \tag{2}$$

where h - threshold, which value depends on the chosen criterion of optimization. Thus, in the examined problem the optimal algorithm (2) includes linear processing of input readout and comparison with a threshold.

Taking into account, that the statistical theory of signal processing gives linear optimal algorithm (2), at first we consider the linear neural network, given in Fig. 1. For training this network is used N realizations of the kind (1). Thus we shall believe, that in $N_0 = p_0 N$ realizations parameter $\vartheta = 0$, and in realizations $N_1 = p_1 N$ we have $\vartheta = 1$.

Let the linear neural network has K inputs. Let's consider the network without preliminary scaling of input signals, since for linear system it does not play basic role. During training of the neural network by i sample the readout (2) is fed on the k input. The signal on the output of the neural network is described by the formula

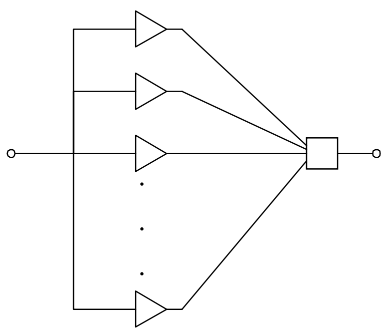

Fig.1. Linear neural network

$$u = \hat{\vartheta} = \mathbf{c}^{\mathrm{T}} \mathbf{y} \,, \tag{3}$$

where $\mathbf{y} = \left| 1 \; y_1 \; y_2 \; \ldots \; y_K \right|^{\mathrm{T}}$, $\mathbf{c} = \left| c_0 \; c_1 \; \ldots \; c_K \right|^{\mathrm{T}}$ – factors subject to optimization.

During training of the neural network the square function is minimized

$$R = \frac{1}{N} \sum_{i=1}^{N} \left(\vartheta - \hat{\vartheta} \right)^2 = R = \frac{1}{N} \left[\sum_{i=1}^{N_0} \left(\mathbf{c}^{\mathrm{T}} \mathbf{y}_i^{(0)} \right)^2 + \sum_{j=1}^{N_1} \left(1 - \mathbf{c}^{\mathrm{T}} \mathbf{y}_j^{(1)} \right)^2 \right], \tag{4}$$

where $\mathbf{y}_i^{(1)}$, $\mathbf{y}_i^{(0)}$ correspond to training samples at presence and absence of the signal S_k, which number is equal N_1 and N_0. For the linear neural network minimization of (4) can be executed analytically, that gives the following expression for optimal factors

$$c_{0,opt} = \tilde{p}_1 \left(1 - \left(\mathbf{S}^{\mathrm{T}} \tilde{\mathbf{c}}_{opt} \right) \right), \;\; \tilde{\mathbf{c}}_{opt} = \frac{\tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right)}{\left(1 + \tilde{p}_1 \tilde{q} \left(1 - \tilde{p}_1 \right) \right) E_S} \mathbf{S} \,, \tag{5}$$

where the designations are used

$$\tilde{\mathbf{c}} = \left| c_1 \; \ldots \; c_K \right|^{\mathrm{T}}, \; \mathbf{S} = \left| S_1 \; S_2 \; \ldots \; S_K \right|^{\mathrm{T}}, \;\; E_S = \sum_{k=1}^{K} S_k^2 \,, \;\; \tilde{p}_1 = \frac{N_1}{N} \,, \; \tilde{q} = \sum_{k=1}^{K} S_k^2 \Big/ \tilde{\sigma}_n^2 \,, \; \tilde{\sigma}_n^2 = \frac{1}{m} \sum_{i=1}^{m} n_i^2 \,,$$

and assumptions also are accepted $\tilde{m}_n = \sum_{i=1}^{m} n_i \Big/ m \approx 0$, , $\tilde{r}_n^2 = \sum_{i=1}^{m} n_i \, n_{j \neq i} \Big/ m \approx 0$,

Let's define the algorithm of detection as follows: the decision of signal presence is accepted, if $\hat{\vartheta} \geq \frac{1}{2}$. Then from (3), (5) it is possible to receive the algorithm

$$\frac{\tilde{\mathbf{y}}^{\mathrm{T}}\mathbf{S}}{\sigma_n^2} \geq \frac{q}{2} + \frac{\tilde{p}_0 - \tilde{p}_1}{2\tilde{p}_1\left(1-\tilde{p}_1\right)} \cdot \frac{q}{\tilde{q}} = h\,, \tag{6}$$

where $\tilde{\mathbf{y}} = \left| y_1\, y_2\, \ldots\, y_K \right|^{\mathrm{T}}$, $q = E_S / \sigma_n^2$ – signal-to-noise ratio, $\tilde{p}_0 = N_0 / N$.

Thus we receive, that at the accepted procedure of training (4) the neural algorithm realizes a classical rule of detection. However value of the threshold differs from some, that reflects the fact of other optimization criterion.

The modeling of the neural algorithm with use of the soft *STATISTICA Neural Networks* of *StatSoft* firm, carried out for the signal $S_k = A\cos\left(2\pi k/K + \varphi_0\right)$, has confirmed the fact of adjustment of weight factors according to (5) with accuracy $\approx 10\%$. Thus of the characteristic of detection (probability of correct detection, probability of false alarm) in the neural algorithm is not better, than in statistical algorithm at the same value of threshold. It is necessary to note a difference between the neural algorithm and statistical algorithms. In statistical algorithms the value of threshold is set apriory and in the neural algorithm it is defined by the process of training, since in (6) the right part depends from $\tilde{p}_0, \tilde{p}_1, \tilde{q}$. From this it is followed the conclusion that apriory information about statistical characteristics of input signals, which is used in the statistical theory of signal processing as initial during synthesis, in neural algorithms should correctly be used at the stage of training of a neural network. An other consequence is that, at absence of apriory information about statistical characteristics of input signals a procedure of training of a network becomes also uncertain. Therefore accuracy and other quantitative characteristics of neural algorithms not always can be better than the characteristics of known statistical algorithms, as against the statement [2].

Use in examined problem a nonlinear neural network results in deterioration of the characteristics of such devices.

In the report one more problem of detection - detection of a signal $S_k = A\cos\left(2\pi k/K + \varphi_0\right)$ with a random initial phase φ_0 is considered. The statistical theory for the such problem gives known algorithm of linear detecting of signal envelope and comparison with a threshold [1], which is nonlinear. A neural algorithm was also searched as nonlinear. The researches of such algorithms have shown that their characteristics are worse, than for statistical optimal algorithm of signal processing. From here it is possible to make a conclusion that, if an input signal has really statistical property and its statistical description is known and is those, that supposes the decision of problem of synthesis of optimal algorithm by statistical methods, the given methods should be used, since neural network methods and the algorithms in such tasks will not give the best result.

References

1. Tichonov V.I., Harisov V.N. Statistical analysis and synthesis of radio engineering devices and systems. - M.: Radio and communication, 1991. - 608 p.

2. Istratov A.U., Melnik A.V., Gribkov V.F. Empirical neural network algorithm of processing of the radar-tracking information // 5 All-Russia conferences "Neural computers and their application". 1999. P.228-233.

3. Galushkin A.I. Theory of neural networks. - M.: IPRGR, 2000. - 416 p.