УДК 004.032.26

Махнист Л.П.

АНАЛИЗ СХОДИМОСТИ АЛГОРИТМОВ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ МЕТОДА НАИСКОРЕЙШЕГО СПУСКА

Рассмотрим нейронную сеть, состоящую из n нейронных элементов распределительного слоя и m - выходного слоя (рис. 1).

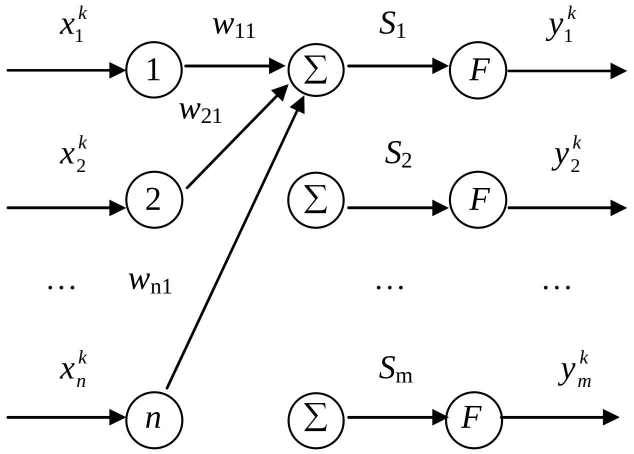

Рис. 1. Схема функционирования нейронной сети.

Для данной сети каждый нейрон распределительного слоя имеет синаптические связи w_{ij}, $\left(i=\overline{1,n},\ \ j=\overline{1,m}\right)$ со всеми нейронами обрабатывающего слоя. В качестве нейронов выходного слоя используются элементы с некоторой функцией активации F [1, 2]. На вход сети подаются входные образы – векторы $\overline{x^k}=\left(x_1^k,\ldots,x_n^k\right)$ $\left(k=\overline{1,L}\right)$.

Выходное значение j-ого нейрона сети для k-ого образа определяется выражением:

$$y_j^k=F\left(S_j^k\right),$$

где $S_j^k=\sum\limits_{i=1}^{n}w_{ij}x_i^k-T_j,\ \ j=\overline{1,m},\ \ k=\overline{1,L}$.

Задача обучения нейронной сети с фиксированной функцией активации F состоит в нахождении весовых коэффициентов w_{ij} $\left(i=\overline{1,n},\ \ j=\overline{1,m}\right)$ и порогов нейронных элементов T_j $\left(j=\overline{1,m}\right)$, которые минимизируют некоторую ошибку сети E_S, как отклонение выходных значений y_j^k от эталонных значений t_j^k – j-ого нейрона сети для k-ого образа. В качестве ошибки сети можно рассмотреть "квадратичное отклонение" $E_S=\frac{1}{2}\sum\limits_{k=1}^{L}\sum\limits_{j=1}^{m}\left(y_j^k-t_j^k\right)^2$, которое будем называть квадратичной ошибкой сети.

Столбец

$$\overline{W}=\left(w_{11},w_{21},\ldots,w_{n1},T_1,w_{12},w_{22},\ldots,w_{n2},\right.$$
$$\left.T_2,\ldots,w_{1m},w_{2m},\ldots,w_{nm},T_m\right)^T$$

будем называть приближенным решением или просто решением системы (по методу наименьших квадратов):

$$F\left(\sum_{i=1}^{n}w_{ij}x_i^k-T_j\right)=t_j^k,\ \ j=\overline{1,m},\ \ k=\overline{1,L},$$

если "квадратичное отклонение" $E_S=\frac{1}{2}\sum\limits_{k=1}^{L}\sum\limits_{j=1}^{m}\left(F\left(\sum\limits_{i=1}^{n}w_{ij}x_i^k-T_j\right)-t_j^k\right)^2$ достигает своего наименьшего значения. Для нахождения такого решения можно применять различные градиентные методы, например, метод наискорейшего спуска, метод сопряженных градиентов и их модификации [1, 2, 3].

Рассмотрим метод наискорейшего спуска $\overline{W}(t+1)=\overline{W}(t)-\alpha(t)\cdot\nabla E_S\left(\overline{W}(t)\right)$ минимизации функции ошибки сети

$$E_S\left(\overline{W}\right)=\frac{1}{2}\sum_{k=1}^{L}\sum_{j=1}^{m}\left(F\left(\sum_{i=1}^{n}w_{ij}x_i^k-T_j\right)-t_j^k\right)^2.$$

Нас будет интересовать поведение этого метода при различных предположениях относительно $E_S\left(\overline{W}\right)$ и $\alpha(t)$.

Теорема 1. Если градиент функции $E_S\left(\overline{W}\right)$ удовлетворяет условию Липшица: $\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|\le L\left\|\overline{W}-\overline{V}\right\|$ и $\alpha(t)$ удовлетворяет условию $0<\varepsilon<\alpha(t)<\frac{2}{L}-\varepsilon$, для любого t, то в методе $\overline{W}(t+1)=\overline{W}(t)-\alpha(t)\cdot\nabla E_S(t)$ ее градиент стремится к нулю: $\lim\limits_{t\to\infty}\nabla E_S\left(\overline{W}(t)\right)=0$, а функция монотонно убывает: $E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)$.

Доказательство: Учитывая, что

$$E_S\left(\overline{W}+\overline{V}\right)=E_S\left(\overline{W}\right)+\int_0^1\left(\nabla E_S\left(\overline{W}+\tau\overline{V}\right),\overline{V}\right)d\tau=$$
$$=E_S\left(\overline{W}\right)+\left(\nabla E_S\left(\overline{W}\right),\overline{V}\right)+\int_0^1\left(\nabla E_S\left(\overline{W}+\tau\overline{V}\right)-\nabla E_S\left(\overline{W}\right),\overline{V}\right)d\tau$$

и подставляя в это соотношение $\overline{W}=\overline{W}(t)$ и $\overline{V}=-\alpha(t)\nabla E_S\left(\overline{W}(t)\right)$, получим

$$E_S\left(\overline{W}(t+1)\right)=E_S\left(\overline{W}(t)\right)-\alpha(t)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2-$$

Махнист Леонид Петрович, *к.т.н., доцент каф. высшей математики Брестского государственного технического университета.*
Беларусь, БрГТУ, 224017, Беларусь, г. Брест, ул. Московская, 267.

$$-\alpha(t)\int_0^1\left(\nabla E_S\left(\overline{W}(t)-\tau\alpha(t)\nabla E_S\left(\overline{W}(t)\right)\right)-\nabla E_S\left(\overline{W}(t)\right),\nabla E_S\left(\overline{W}(t)\right)\right)d\tau$$

Так как выполняется условие $\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|\le L\left\|\overline{W}-\overline{V}\right\|$, то

$$E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)-\alpha(t)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2+$$

$$+L\alpha^2(t)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2\int_0^1\tau d\tau=$$

$$=E_S\left(\overline{W}(t)\right)-\alpha(t)\left(1-\alpha(t)\frac{L}{2}\right)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2.$$

Учитывая условие $0<\varepsilon<\alpha(t)<\frac{2}{L}-\varepsilon$, имеем

$$\alpha(t)\left(1-\alpha(t)\frac{L}{2}\right)=\frac{1}{2L}-\frac{L}{2}\left(\alpha(t)-\frac{1}{L}\right)^2\ge$$

$$\ge\frac{1}{2L}-\frac{L}{2}\left(\varepsilon-\frac{1}{L}\right)^2=\varepsilon-\frac{L}{2}\varepsilon^2=\frac{L}{2}\varepsilon\left(\frac{2}{L}-\varepsilon\right)>0$$

Тогда

$$E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)-\frac{L}{2}\varepsilon\left(\frac{2}{L}-\varepsilon\right)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2.$$

Суммируя неравенства $E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)-\gamma\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$, где $\gamma=\frac{L}{2}\varepsilon\left(\frac{2}{L}-\varepsilon\right)$ по t от 0 до k, получим

$$E_S\left(\overline{W}(k+1)\right)\le E_S\left(\overline{W}(0)\right)-\gamma\sum_{t=0}^{k}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$$ или

$$\sum_{t=0}^{k}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2\le\frac{1}{\gamma}\left(E_S\left(\overline{W}(0)\right)-E_S\left(\overline{W}(k+1)\right)\right)\le\frac{1}{\gamma}E_S\left(\overline{W}(0)\right)$$

Так как $E_S\left(\overline{W}\right)\ge 0$, то $\sum_{t=0}^{k}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2\le\frac{1}{\gamma}E_S\left(\overline{W}(0)\right)$ для любого k, т. е. $\sum_{t=0}^{\infty}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2<+\infty$. Следовательно, градиент функции $E_S\left(\overline{W}\right)$ стремится к нулю: $\lim_{t\to\infty}\nabla E_S\left(\overline{W}(t)\right)=0$, а функция $E_S\left(\overline{W}\right)$ монотонно убывает: $E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)$.

Теорема доказана.

Лемма 1. Если градиент функции $E_S\left(\overline{W}\right)$ удовлетворяет условию Липшица: $\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|\le L\left\|\overline{W}-\overline{V}\right\|$, то $\left\|\nabla E_S\left(\overline{W}\right)\right\|^2\le 2L\cdot E_S\left(\overline{W}\right)$.

Доказательство: Полагая в методе $\overline{W}(t+1)=\overline{W}(t)-\alpha(t)\cdot\nabla E_S\left(\overline{W}(t)\right)$ величину шага $\alpha(t)=\frac{1}{L}$, который, очевидно, удовлетворяет условию $0<\varepsilon<\alpha(t)<\frac{2}{L}-\varepsilon$ и, учитывая неравенство $E_S\left(\overline{W}(t+1)\right)\le E_S\left(\overline{W}(t)\right)-\alpha(t)\left(1-\alpha(t)\frac{L}{2}\right)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$ (см. теорему 1), получим $E_S\left(\overline{W}(t)-\frac{1}{L}\nabla E_S\left(\overline{W}(t)\right)\right)\le E_S\left(\overline{W}(t)\right)-\frac{1}{2L}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$. Так как $E_S\left(\overline{W}\right)\ge 0$, то $0\le E_S\left(\overline{W}(t)\right)-\frac{1}{2L}\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$ или $\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2\le 2L\cdot E_S\left(\overline{W}(t)\right)$ для любого t, т. е. $\left\|\nabla E_S\left(\overline{W}\right)\right\|^2\le 2L\cdot E_S\left(\overline{W}\right)$, что и требовалось доказать.

Замечание. Очевидно, что результат, полученный в лемме, выполняется для любой дифференцируемой функции $f\left(\overline{x}\right)$, градиент которой удовлетворяет условию Липшица: $\left\|\nabla f\left(\overline{x}\right)-\nabla f\left(\overline{y}\right)\right\|\le L\left\|\overline{x}-\overline{y}\right\|$, и для которой $f\left(\overline{x}\right)\ge 0$ для любого $\overline{x}$. Тогда выполняется $\left\|\nabla f\left(\overline{x}\right)\right\|^2\le 2L\cdot f\left(\overline{x}\right)$.

Рассмотрим в качестве функции $f\left(\overline{x}\right)=\frac{1}{2}\left(A\overline{x},\overline{x}\right)$, где матрица A является симметричной и неотрицательно определенной: $A\ge 0$. Учитывая, что функция $f\left(\overline{x}\right)=\frac{1}{2}\left(A\overline{x},\overline{x}\right)$ удовлетворяет условиям леммы и то, что $\nabla f\left(\overline{x}\right)=A\overline{x}$ и, очевидно, $\left\|\nabla f\left(\overline{x}\right)-\nabla f\left(\overline{y}\right)\right\|=\left\|A\overline{x}-A\overline{y}\right\|\le\left\|A\right\|\left\|\overline{x}-\overline{y}\right\|$, т. е. для функции $f\left(\overline{x}\right)$ выполняется условие Липшица с константой $L=\left\|A\right\|$, получим

$$\left\|A\overline{x}\right\|^2=\left\|\nabla f\left(\overline{x}\right)\right\|^2\le 2L\cdot f\left(\overline{x}\right)=2\left\|A\right\|\cdot\frac{1}{2}\left(A\overline{x},\overline{x}\right)=\left\|A\right\|\cdot\left(A\overline{x},\overline{x}\right)$$

. Таким образом, для симметричной и неотрицательно определенной матрицы A выполняется $\left(A\overline{x},\overline{x}\right)\ge\frac{1}{\left\|A\right\|}\left\|A\overline{x}\right\|^2$ для любого $\overline{x}$.

Лемма 2. Если $E_S\left(\overline{W}\right)$ является дважды дифференцируемой выпуклой функцией: $E_S\left(\lambda\overline{W}+(1-\lambda)\overline{V}\right)\le\lambda E_S\left(\overline{W}\right)+(1-\lambda)E_S\left(\overline{V}\right)$, для любых $\overline{W}$, $\overline{V}$, $0\le\lambda\le 1$, а ее градиент $\nabla E_S\left(\overline{W}\right)$ удовлетворяет условию Липшица: $\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|\le L\left\|\overline{W}-\overline{V}\right\|$, то

$$\left(\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right),\overline{W}-\overline{V}\right)\geq\frac{1}{L}\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|^2$$

.

Доказательство: Учитывая, что функция $E_S\left(\overline{W}\right)$ дважды дифференцируема, имеем

$$\nabla E_S\left(\overline{V}\right)=\nabla E_S\left(\overline{W}\right)+\int_0^1\left(\nabla^2 E_S\left(\overline{W}+\tau(\overline{V}-\overline{W})\right),\overline{V}-\overline{W}\right)d\tau=$$

$$=\nabla E_S\left(\overline{W}\right)+A(\overline{V}-\overline{W})$$

где матрица $A=\int_0^1\nabla^2 E_S\left(\overline{W}+\tau(\overline{V}-\overline{W})\right)d\tau$ является симметричной и неотрицательно определенной: $A\geq 0$, так как выпуклость дважды дифференцируемой функции $E_S\left(\overline{W}\right)$ эквивалентна условию $\nabla^2 E_S\left(\overline{W}\right)\geq 0$. В силу того, что градиент $\nabla E_S\left(\overline{W}\right)$ удовлетворяет условию Липшица $\left\|\nabla^2 E_S\left(\overline{W}\right)\right\|\leq L$ и, очевидно, выполняется условие $\|A\|\leq L$. Так как $\nabla E_S\left(\overline{V}\right)-\nabla E_S\left(\overline{W}\right)=A(\overline{V}-\overline{W})$ или $\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)=A(\overline{W}-\overline{V})$, и, учитывая замечание к лемме 1, имеем

$$\left(\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right),\overline{W}-\overline{V}\right)=\left(A(\overline{W}-\overline{V}),\overline{W}-\overline{V}\right)\geq$$

$$\geq\frac{1}{\|A\|}\left\|A(\overline{W}-\overline{V})\right\|^2\geq\frac{1}{L}\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|^2$$

что и требовалось доказать.

Нетрудно доказать следующее утверждение.

Лемма 3. Если функция $E_S\left(\overline{W}\right)$ является сильно выпуклой с константой ℓ:

$$E_S\left(\overline{W}\right)\geq E_S\left(\overline{V}\right)+\left(\nabla E_S\left(\overline{V}\right),\overline{W}-\overline{V}\right)+\frac{\ell}{2}\left\|\overline{W}-\overline{V}\right\|^2,$$

$\overline{W}^*$ – ее точка минимума, то $\left\|\nabla E_S\left(\overline{W}\right)\right\|^2\geq 2\ell\left(E_S\left(\overline{W}\right)-E_S\left(\overline{W}^*\right)\right)$.

Теорема 2. Пусть функция $E_S\left(\overline{W}\right)$ дифференцируема, а ее градиент удовлетворяет условию Липшица с константой L: $\left\|\nabla E_S\left(\overline{W}\right)-\nabla E_S\left(\overline{V}\right)\right\|\leq L\left\|\overline{W}-\overline{V}\right\|$ и функция $E_S\left(\overline{W}\right)$ является сильно выпуклой с константой ℓ:

$$E_S\left(\overline{W}\right)\geq E_S\left(\overline{V}\right)+\left(\nabla E_S\left(\overline{V}\right),\overline{W}-\overline{V}\right)+\frac{\ell}{2}\left\|\overline{W}-\overline{V}\right\|^2.$$

Тогда если $\alpha(t)$ удовлетворяет условию $0<\varepsilon<\alpha(t)<\frac{2}{L}-\varepsilon$, для любого t, то метод $\overline{W}(t+1)=\overline{W}(t)-\alpha(t)\cdot\nabla E_S(t)$ сходится к единственной точке глобального минимума $\overline{W}^*$ со скоростью геометрической прогрессии: $\left\|\overline{W}(t)-\overline{W}^*\right\|\leq cq^t$, $0<q<1$.

Доказательство: Так как выполнены все условия теоремы 1, то справедливо неравенство:

$$E_S\left(\overline{W}(t+1)\right)\leq E_S\left(\overline{W}(t)\right)-\alpha(t)\left(1-\alpha(t)\frac{L}{2}\right)\left\|\nabla E_S\left(\overline{W}(t)\right)\right\|^2$$

. Учитывая, что функция $E_S\left(\overline{W}\right)$ является сильно выпуклой с константой ℓ, то $\left\|\nabla E_S\left(\overline{W}\right)\right\|^2\geq 2\ell\left(E_S\left(\overline{W}\right)-E_S\left(\overline{W}^*\right)\right)$, где $\overline{W}^*$ – точка глобального минимума.

Тогда

$$E_S\left(\overline{W}(t+1)\right)\leq E_S\left(\overline{W}(t)\right)-$$
$$-\ell\alpha(t)\left(2-L\alpha(t)\right)\left(E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\right).$$

Вычитая $E_S\left(\overline{W}^*\right)$ из левой и правой части последнего неравенства, получим

$$E_S\left(\overline{W}(t+1)\right)-E_S\left(\overline{W}^*\right)\leq E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)-$$
$$-\ell\alpha(t)\left(2-L\alpha(t)\right)\left(E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\right)=$$
$$=\left(1-\ell\alpha(t)\left(2-L\alpha(t)\right)\right)\left(E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\right)=$$
$$=\left(\ell L\alpha^2(t)-2\ell\alpha(t)+1\right)\left(E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\right).$$

Обозначим $q(t)=\ell L\alpha^2(t)-2\ell\alpha(t)+1$. Тогда

$$E_S\left(\overline{W}(t+1)\right)-E_S\left(\overline{W}^*\right)\leq q(t)\left(E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\right)\leq$$
$$\leq q(t)q(t-1)\left(E_S\left(\overline{W}(t-1)\right)-E_S\left(\overline{W}^*\right)\right)\leq$$
$$\leq\left(\prod_{i=0}^{t}q(i)\right)\left(E_S\left(\overline{W}(0)\right)-E_S\left(\overline{W}^*\right)\right).$$

Таким образом $E_S\left(\overline{W}(t)\right)-E_S\left(\overline{W}^*\right)\leq\left(\prod_{i=0}^{t-1}q(i)\right)\left(E_S\left(\overline{W}(0)\right)-E_S\left(\overline{W}^*\right)\right)$.

Так как $\alpha(t)$ удовлетворяет условию $0<\varepsilon<\alpha(t)<\frac{2}{L}-\varepsilon$, для любого t, то

$$q(t)=\ell L\alpha^2(t)-2\ell\alpha^2(t)+1=\ell L\left(\alpha(t)-\frac{1}{L}\right)^2+$$
$$+1-\frac{\ell}{L}\geq q(\frac{1}{L})=\frac{L-\ell}{L}\geq 0$$

,и

$$q(t)=\ell L\alpha^2(t)-2\ell\alpha^2(t)+1\leq q(\varepsilon t)=$$
$$=q(\frac{2}{L}-\varepsilon)=\ell L\varepsilon^2-2\ell\varepsilon+1=1-\ell L\varepsilon\left(\frac{2}{L}-\varepsilon\right)<1$$

Тогда $0 \le q(t) \le q_0 = 1 - \ell L\varepsilon\left(\frac{2}{L}-\varepsilon\right) < 1$ и, следовательно, $E_S\left(\overline{W}(t)\right) \to E_S\left(\overline{W}^*\right)$ при $t \to \infty$.

В силу того, что для сильно выпуклой функции с константой ℓ выполняется:

$$E_S\left(\overline{W}\right) \ge E_S\left(\overline{V}\right) + \left(\nabla E_S\left(\overline{V}\right), \overline{W} - \overline{V}\right) + \frac{\ell}{2}\left\|\overline{W} - \overline{V}\right\|^2,$$

то, полагая $\overline{V} = \overline{W}^*$ и учитывая, что $\nabla E_S\left(\overline{W}^*\right) = \overline{0}$, получим

$$E_S\left(\overline{W}\right) \ge E_S\left(\overline{W}^*\right) + \frac{\ell}{2}\left\|\overline{W} - \overline{W}^*\right\|^2$$

или

$$\left\|\overline{W} - \overline{W}^*\right\|^2 \le \frac{2}{\ell}\left(E_S\left(\overline{W}\right) - E_S\left(\overline{W}^*\right)\right).$$

Откуда

$$\begin{aligned}\left\|\overline{W}(t) - \overline{W}^*\right\|^2 &\le \frac{2}{\ell}\left(E_S\left(\overline{W}(t)\right) - E_S\left(\overline{W}^*\right)\right) \le \\ &\le \frac{2}{\ell} q_0^t \left(E_S\left(\overline{W}(0)\right) - E_S\left(\overline{W}^*\right)\right)\end{aligned}.$$

Таким образом, $\left\|\overline{W}(t) - \overline{W}^*\right\| \le c q^t$, где $c = \sqrt{\frac{2}{\ell}}$, $0 \le q = \sqrt{q_0} = \sqrt{1 - \ell L\varepsilon\left(\frac{2}{L}-\varepsilon\right)} < 1$, что и требовалось доказать.

Замечание. Использованная техника построения доказательств приведенных утверждений может быть распространена и на более широкий класс функций ошибки сети с получением локальных аналогов полученных теорем.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Головко В.А. Нейронные сети: обучение, организация и применение. Кн. 4: Учебное пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.: ил. (Нейрокомпьютеры и их применение).
2. Гладкий И.И., Головко В.А., Махнист Л.П. Обучение нейронных сетей с использованием метода наискорейшего спуска // Вестник Брестского государственного технического университета. Физика, математика, химия. – Брест: БГТУ, 2001. – № 5 – С. 47-55.
3. Махнист Л.П. Обучение нейронных сетей с использованием метода сопряженных градиентов // Вестник Брестского государственного технического университета. Машиностроение, автоматизация, ЭВМ. – Брест: БГТУ, 2002. – № 4 – С. 74-77.

УДК 004.8.032.26

Головко В.А., Войцехович Л.Ю.

НЕЙРОСЕТЕВЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ АТАК НА КОМПЬЮТЕРНЫЕ СЕТИ

1. ВВЕДЕНИЕ

Одной из форм глобализации мирового пространства является информационная глобализация, которая связана с широким распространением сети Интернет. Информационная глобализация увеличивает степень уязвимости компьютерных систем, что уменьшает их безопасность. Атакой на компьютерные сети называется совокупность определенных действий, приводящих к подрыву безопасности системы. В результате атаки злоумышленник может получить доступ к конфиденциальной информации или нарушить нормальное функционирование системы. Это приводит к большим материальным и социальным издержкам.

Важным этапом обеспечения безопасности компьютерных систем является проектирование систем обнаружения атак (Intrusion Detection System – IDS). Такие системы способны на основе анализа сетевого трафика автоматически обнаруживать атаки TCP/IP, что позволяет предпринять необходимые меры для нейтрализации угрозы.

В данной работе рассматриваются нейросетевые подходы для построения систем обнаружения атак. В качестве базы данных для тестирования системы используется KDD-99 [1], которая содержит почти 5 миллионов записей соединений и 41 параметр сетевого трафика. При этом атаки делятся на четыре основные категории: DoS, U2R, R2L и Probe.

Атака DoS – отказ в обслуживании, характеризуется генерацией большого объема трафика, что приводит к перегрузке и блокированию сервера.

Атака U2R предполагает получение зарегистрированным пользователем привилегий локального суперпользователя (администратора).

Атака R2L характеризуется получением доступа незарегистрированного пользователя к компьютеру со стороны удаленной машины.

Атака Probe заключается в сканировании портов с целью получения конфиденциальной информации.

В работе предлагаются различные варианты построения систем обнаружения атак, которые базируются на использовании рециркуляционных и многослойных нейронных сетей. Результаты экспериментов обсуждаются.

2. ГЕНЕРИРОВАНИЕ АРХИТЕКТУРНЫХ РЕШЕНИЙ

Рассмотрим различные архитектурные решения для построения систем обнаружения атак. В качестве входных данных используется 41-размерный вектор, который характеризует параметры соединения сети. Задачей IDS является обнаружение и распознавание атак. Поэтому в качестве выходных данных используется m-мерный вектор, где m равняется количеству атак плюс нормальное состояние.

На рис. 1 приведена система обнаружения атак, которая состоит из рециркуляционной нейронной сети (RNN) и многослойного персептрона (MLP).

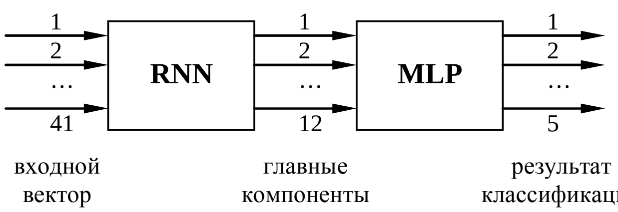

Рис. 1. 1-ый вариант IDS.

Задачей RNN является сжатие входного пространства образов с целью получения главных компонент [2]. Главные компоненты являются некоррелированными и содержат наиболее информативные признаки исходного пространства образов. Многослойный персептрон осуществляет обработку сжатого пространства входных образов (главных компонент) с целью распознавания типа атаки.