ВЕРОЯТНОСТНЫЕ СВОЙСТВА НАЧАЛЬНЫХ ЗНАЧЕНИЙ ВЕСОВЫХ КОЭФФИЦИЕНТОВ В СИНХРОНИЗИРУЕМЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЯХ КИНЦЕЛЯ

докт. технических наук, проф. ГОЛИКОВ В.Ф.
магистр технических наук, асп. БРИЧ Н.В.

АННОТАЦИЯ

Использование методов нейрокриптографии является перспективным способом для формирования идентичных бинарных последовательностей. В статье проанализировано влияние начальных значений весовых коэффициентов двух сетей на скорость достижения синхронности ИНС. Вывлено положительное свойство – равновероятность направлений движения начальных значений весовых коэффициентов. На основе проведенного исследования авторами предполагается возможность по совершенствованию архитектуры сети и алгоритма коррекции весовых коэффициентов.

ABSTRACT

One of the most efficient ways for identical binary se quences generation is using methods of neural cryptography. The initial weight vestors values influence on speed of synchronization is analized. Equal probability of initial weight vestors motion directions is great advantage. On this base authors suppose new line of research conserned with improvement of network architecture and correction algorithm.

Введение

Одним из ведущих направлений развития информационных технологий является обеспечение информационной безопасности. Основой большинства криптографических средств защиты информации является шифрование. Актуальность решения задачи конфиденциальной доставки ключевой информации в симметричных (одноключевых) криптографических системах очевидна. На первый взгляд указанная проблема решена в асимметричных криптосистемах, где используются разные ключи для шифрования и расшифрования. Однако, во-первых, на сегодняшний день отсутствуют математические доказательства необратимости односторонних функций, используемых в асимметричных алгоритмах. Во-вторых, возникает проблема защиты открытых ключей от подмены.

И.Кантер и В.Кинцель предлагают идею использования синхронизируемых искусственных нейронных сетей (ИНС) [1].

ИНС представляет собой сеть элементов (искусственных нейронов), связанных между собой синаптическими соединениями. Математической моделью нейросети является персептрон.

ИНС считаются синхронизированными, если совпадают значения векторов весовых коэффициентов персептронов сетей (изначально значения принимаются различными и случайными). Подавая на входы персептронов одинаковые случайные вектора и сравнивая между собой выходные значения, можно корректировать значения весов. В результате многократного повторения эти величины в некоторый момент времени станут равными. Таким образом, обеспечив секретность начальных значений весов, в качестве общего криптографического ключа можно принять итоговые веса персептронов [2].

Следовательно, использование методов нейрокриптографии для формирования идентичных бинарных последовательностей является перспективным способом, но нуждается в более глубоких исследованиях, подтверждаю-

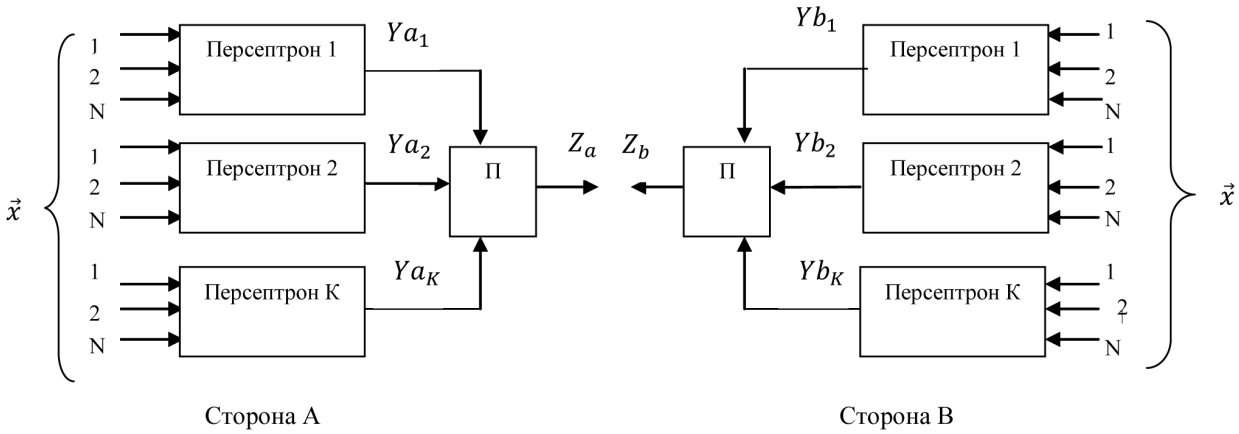

Рис. 1. Синхронизируемые ИНС.

щих сходимость процесса синхронизации.

В частности, необходимо проанализировать влияние начальных значений весовых коэффициентов двух сетей на скорость достижения синхронности ИНС.

Основная часть

Пусть имеется ИНС, состоящая из К внутренних персептронов (рис.1). Каждый персептрон имеет n входов.

До начала синхронизации абонент *A* должен сформировать вектор весовых коэффициентов (1)

$$\overline{wa} = wa_{11}, wa_{12}, \dots, wa_{1N}, wa_{21}, wa_{22}, \dots, \dots, wa_{2N}, \dots, wa_{K1}, wa_{K2}, \dots, wa_{KN},$$

где $wa_{ij} \in [-L, L],\ i = 1, 2, \dots, K;\ j = 1,$

Величину L целесообразно выбирать как 2^q, где q- целое положительное число. Каждый элемент вектора $\overline{wa}$ есть случайное целое число с дискретным равномерным законом распределения (рис.2)

$$P(wa_{ij} = sa_{ij}) = \frac{1}{2L+1},$$

где $sa_{ij} = -L, -L+1, \dots, 0, 1, \dots, L.$

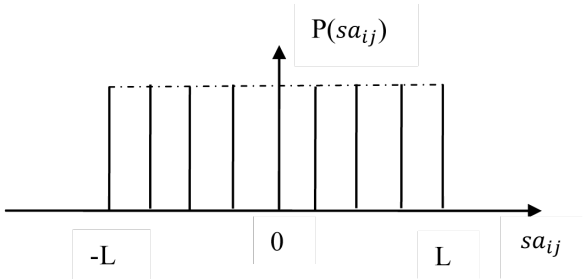

Рис.2. Закон распределения начальных значений весовых коэффициентов.

Абонент *B*, имеющий такую же сеть, независимо от *A* формирует вектор весовых коэффициентов своей сети (2)

$$\overline{wb} = wb_{11}, wb_{12}, \dots, wb_{1N}, wb_{21}, wb_{22}, \dots, \dots, wb_{2N}, \dots, wb_{K1}, wb_{K2}, \dots, wb_{KN},$$

с аналогичными вероятностными характеристиками.

На входы сетей *A* и *B* подается вектор синхронизирующих чисел

$$\vec{x} = x_{11}, x_{12}, \dots, x_{1N}, x_{21}, x_{22}, \dots, \dots, x_{2N}, \dots, x_{K1}, x_{K2}, \dots, x_{KN},$$

где $x_{ij} \in [-1,1]$ – дискретная случайная величина с равномерным распределением. Корректируя весовые коэффициенты сетей в ходе синхронизации по определенному правилу, абоненты добиваются их идентичности.

Каждый последующий шаг синхронизации начинается с подачи на входы обеих сетей выбранного случайным образом вектора $\vec{x}$. Затем вычисляется выходная величина *Z* для каждой из сетей

$$Za/b = \prod_{i=1}^{K} Ya/b_i = \prod_{i=1}^{K} \sigma\left(\sum_{j=1}^{N} w_{ij}^{A/B} x_{ij}\right)$$

Индекс *a/b* означает, что операция касается обеих сетей *A* и *B*, а единичный индекс – что операция касается одной сети соответственно. Модифицированная функция знака

$$\sigma(*) = \begin{cases} 1, & \sigma(*) \geq 0, \\ -1, & \sigma(*) < 0. \end{cases}$$

Формально это не является классическим методом вычисления выходного значения нейронных сетей, так как выходное значение целой архитектуры вычисляется как произведение выходных величин каждого скрытого нейрона.

На основании сравнения обоих полученных выходных величин реализован процесс синхронизации. Коррекция векторов весов обеих сетей происходит только тогда, когда обе выходные величины равны друг другу ($Z^A = Z^B$). Внутри данной сети корректируются веса только тех персептронов, выходная величина которых равна величине Z всей сети. Процесс синхронизации идет по правилу Хэбба

$$wa/b_{ij} = \begin{cases} wa/b_{ij} + Za/b * x_{ij}, Za = \\ = Zb \ \ a \ \ nd \ \ Za/b = Ya/b_i) \geq 0, \\ wa/b_{ij}, \qquad \text{в противном случае.} \end{cases}$$

Также каждый этап процесса синхронизации требует выполнения операции нормализации

$$wa/b_{ij} = \begin{cases} sign\left(wa/b_{ij}\right)L, & \left|wa/b_{ij}\right| > L, \\ wa/b_{ij}, & \text{в противном случае.} \end{cases}$$

Представляет интерес рассмотреть вероятностные свойства начальных значений весовых коэффициентов и процесса синхронизации. Введем в рассмотрение случайную величину $\partial_{ij} = wa_{ij} - wb_{ij}$, равную начальному рассогласованию соответствующих весовых коэффициентов сетей. Найдем закон распределения этой величины. Область возможных значений этой величины с учетом того, что $wa_{ij}, wb_{ij} \in [-L, L]$, равна $[-2L, 2L]$. Вероятность того, что $\partial_{ij} = s_{ij}$, где $s_{ij} = -2L, -2L + + 1, \ldots, 0, 1, \ldots, 2L$, равна

$$P\left(\partial_{ij} = s_{ij}\right) = \sum_{sa_{ij}, sb_{ij} \in (s_{ij} = sa_{ij} - sb_{ij})} P\left(wa_{ij} = sa_{ij};\ wb_{ij} = sb_{ij}\right) \quad (3)$$

Поскольку все значения wa_{ij}, wb_{ij} равновероятны, то для нахождения $P(\partial_{ij} = S_{ij})$ достаточно определить количество значений пар (sa_{ij}, sb_{ij}), соответствующих каждому возможному значению S_{ij}. Для этого построим область возможных значений (sa_{ij}, sb_{ij}) и проведем в

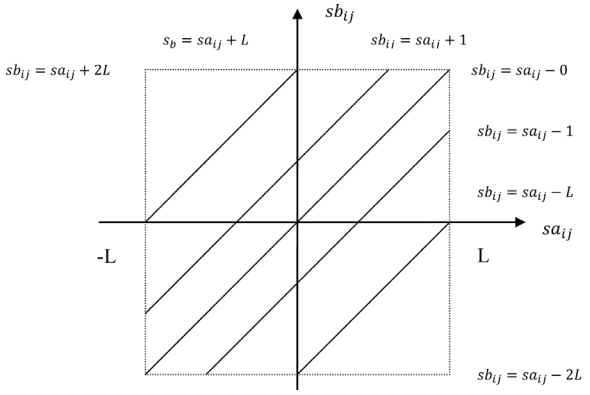

Рис.3. Область возможных значений (sa_{ij}, sb_{ij}).

В соответствии с рис.3 и с учетом того, что значения wa_{ij} и wb_{ij} до начала коррекции независимы между собой, можно записать:

для $s_{ij} = 0$

$$P\left(\partial_{ij} = 0\right) = \sum_{sa_{ij}=-L}^{L} P\left(wa_{ij} = sa_{ij}\right) * * P\left(wb_{ij} = sa_{ij} - 0\right) = \frac{2L+1}{(2L+1)^2},$$

для $s_{ij} > 0$

$$P\left(\partial_{ij} = s_{ij}\right) = \sum_{sa_{ij}=-L+s_{ij}}^{L} P\left(wa_{ij} = sa_{ij}\right) * * P\left(wb_{ij} = sa_{ij} - s_{ij}\right) = \frac{2L+1-s_{ij}}{(2L+1)^2},$$

для $s_{ij} < 0$

$$P\left(\partial_{ij} = s_{ij}\right) = \sum_{sa_{ij}=-L}^{L-s} P\left(wa_{ij} = sa_{ij}\right) * * P\left(wb_{ij} = sa_{ij} - s_{ij}\right) = \frac{2L+1+s_{ij}}{(2L+1)^2}.$$

Объединяя полученные результаты, запишем

$$P\left(\partial_{ij} = s_{ij}\right) = \frac{2L + 1 - ABS(s_{ij})}{(2L+1)^2} \quad (4)$$

где $ABS(S_{ij})$ – абсолютное значение величины s_{ij}, $s_{ij} \in [-2L, 2L]$. Зависимость (4) изображена на рис. 5.

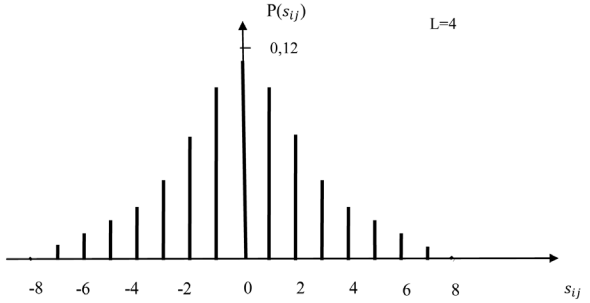

Рис.5. Закон распределения разности весовых коэффициентов

Из рисунка видно, что наибольшей вероятностью обладают совпадающие значения коэффициентов: примерно – 0,12 и значения, отличающиеся на 1, примерно – 0,20. Всего в интервале различий [-2,2] лежит до 60% весовых коэффициентов. Следует ожидать, что именно эти коэффициенты в процессе синхронизации будут согласовываться быстрее остальных. На рис.6 изображена зависимость (4) для $L = 5$. Из сравнения рис.5 и рис.6 видно, что с ростом L закон распределения ∂_{ij} становится более равномерным, и что следует ожидать увеличения времени полной синхронизации.

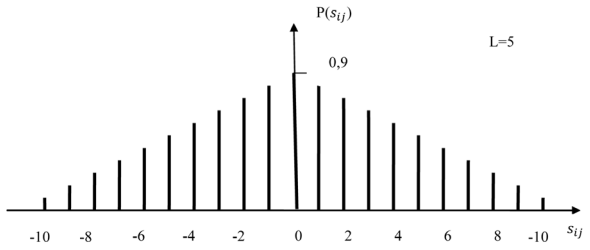

Рис.6. Закон распределения разности весовых коэффициентов

Рассмотрим теперь некоторые вероятностные характеристики процесса синхронизации. Нанесем на плоскость (sa_{ij}, sb_{ij}) случайную точку $(sa_{ij} = sa_{ij}^*, sb_{ij} = sb_{ij}^*)$ (рис.7) и зная правило коррекции весовых коэффициентов, проследим возможные траектории ее движения к полному согласованию (к прямой $sb_{ij} = sa_{ij}$).

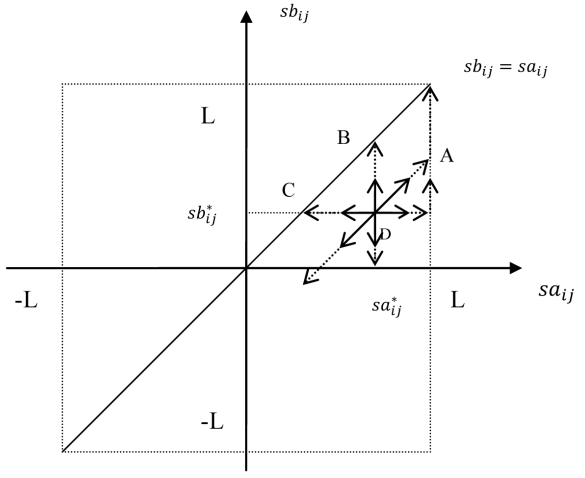

Рис.7. Возможные траектории движения значений весовых коэффициентов.

В соответствии с выбранным правилом коррекции в процессе синхронизации на каждом шаге $(t + 1)$ могут произойти следующие события:

$$A) wa_{ij}(t+1) = wa_{ij}(t) \pm 1,$$
$$wb_{ij}(t+1) = wb_{ij}(t) \pm 1.$$
$$B)\ wa_{ij}(t+1) = wa_{ij}(t) \pm 1,$$
$$wb_{ij}(t+1) = wb_{ij}(t),$$
$$C)\ wa_{ij}(t+1) = wa_{ij}(t),$$
$$wb_{ij}(t+1) = wb_{ij}(t) \pm 1.$$
$$D)\ wa_{ij}(t+1) = wa_{ij}(t),$$
$$wb_{ij}(t+1) = wb_{ij}(t).$$

Событие *A* наиболее благоприятное для согласования значений коэффициентов сетей. При этом событии точка (sa_{ij}^*, sb_{ij}^*) передвигается параллельно линии $sb_{ij} = sa_{ij}$ (вверх и вправо, если прибавляется 1, либо вниз и влево, если прибавляется -1). Это движение на рис.7 обозначено вектором *A*. Достигнув границы (L или $-L$), точка передвигается по границе вверх (вниз) до тех пор пока не достигнет линии $sb_{ij} = sa_{ij}$.

Событие *B* передвигает начальную точку параллельно оси sa_{ij} (вправо, если прибавляется 1, либо влево, если прибавляется -1). При этом движение влево ведет к линии $sb_{ij} = sa_{ij}$ по наиболее короткой траектории, если точка (sa_{ij}^*, sb_{ij}^*) лежит в правой полуплоскости.

Рассуждая аналогично, видим, что событие *C* обеспечивает движение параллельно оси sb_{ij} (вверх или вниз).

Событие *D* оставляет начальную точку без движения.

Таким образом, множество точек (начальные значения весовых коэффициентов синхронизируемых сетей) в процессе синхронизации движется в поле $sb_{ij} = sa_{ij}$ по прямым линиям: диагоналям, горизонталям и вертикалям, переходя с одной линии на другую в сторону главной диагонали $sb_{ij} = sa_{ij}$. Процесс синхронизации заканчивается, когда все точки окажутся на этой диагонали.

События *A, B, C, D* являются случайными

и происходят с определенными вероятностями. Определим эти вероятности. Для простоты расчетов будем рассматривать ИНС, состоящую из трех персептронов ($K = 3$), а в качестве анализируемых весовых коэффициентов выберем весовые коэффициенты первого персептрона сетей A и B.

Событие A для этой пары персептронов имеет место, если одновременно выполняются следующие условия:

$$Ya_1 = Yb_1,\ Ya_2 = Ya_3,\ Yb_2 = Yb_3.$$

В этом случае $Z_a = Z_b$ и весовые коэффициенты обоих персептронов корректируются путем прибавления x_{ij}. Для сети, состоящей из трех персептронов и в силу независимости Ya_1, Yb_1, можно записать

$$P(Ya_1 = Yb_1) = P(Ya_1 = 1,\ Yb_1 = 1) + P(Ya_1 = -1,\ Yb_1 = -1) = P(Ya_1 = 1) * P(Yb_1 = 1) + P(Ya_1 = -1) * P(Yb_1 = -1).$$

Для *каждого персептрона* справедливо

$$P(Ya_1 = 1) = P(Yb_1 = 1) = P(Ya_1 = -1) = P(Yb_1 = -1) = 1/2.$$

Следовательно, $P(Ya_1 = P(Yb_1) = 1/2$. Аналогично можно получить: $P(Ya_2 = Ya_3) = P(Yb_2 = Yb_3) = 1/2$. Таким образом, искомая вероятность равна

$$P(A) = P(Ya_1 = Yb_1,\ Ya_2 = Ya_3,\ Yb_2 = Yb_3) = 1/8.$$

Событие B имеет место, если одновременно выполняются следующие условия:

$Ya_1 \neq Yb_1,\ Ya_2 = Ya_3,\ Yb_2 \neq Yb_3$. Тогда, с учетом ранее приведенных обоснований, справедливо

$$P(A) = P(Ya_1 \neq Yb_1,\ Ya_2 = Ya_3,\ Yb_2 \neq Yb_3) = P(Ya_1 = 1) * P(Yb_1 = -1) + P(Ya_1 = -1) * P(Yb_1 = 1)) * (P(Ya_2 = 1) * P(Ya_3 = 1) + P(Ya_2 = -1) * P(Yb_3 = -1)) * P(Yb_2 = 1) * P(Yb_3 = -1) + P(Yb_2 = -1) * P(Yb_3 = 1) = 1/8.$$

Аналогично для события C имеем $P(C) = P(Ya_1 \neq Yb_1,\ Ya_2 \neq Ya_3,\ Yb_2 = Yb_3) = 1/8$. Вероятность события D, найдем из условия, что события A, B, C, D составляют полную группу событий,

$$P(D) = 1 - (P(A) + P(B) + P(C)) = 5/8.$$

ВЫВОДЫ

1. Закон распределения разности весовых коэффициентов отличается от равномерного. Наиболее вероятны небольшие разности значений.

2. Направления движения начальных значений весовых коэффициентов равновероятны, что является положительным свойством, т.к. сохраняется максимальная неопределенность у внешнего наблюдателя относительно неизвестных ему значений весовых коэффициентов.

3. Более половины тактов синхронизации являются нерезультативными, что может составить предмет дальнейшего совершенствования архитектуры сети и алгоритма коррекции весовых коэффициентов.

4. Следует помнить, что данные выводы относятся к начальным значениям весовых коэффициентов двух сетей, поскольку они являются независимыми. В процессе синхронизации сетей A и B между соответствующими весовыми коэффициентами появляется корреляционная связь, которая усиливается с увеличением числа тактов синхронизации и приведенные соотношения становятся несправедливыми. При полной синхронизации, когда $wa_{ij} = wb_{ij}$, коэффициент линейной корреляции равен 1.

ЛИТЕРАТУРА

1. **Kanter**, I. The Theory of Neural Networks and Cryptography, Quantum Computers and Computing / I. Kanter, W.Kinzel.−2005. – Vol. 5, n.1. – P. 130−140.

2. **Kinzel,** W. Neural Cryptography / W.Kinzel, / I. Kanter // 9th International Conference on Neural Information Processing, Singapore, 2002.