

УДК 004.032.26:621.362

## ИССЛЕДОВАНИЕ АППРОКСИМАЦИИ ВОЛЬТАМПЕРНЫХ ХАРАКТЕРИСТИК ТЕРМОЭМИССИОННОГО ПРЕОБРАЗОВАТЕЛЯ ИСКУССТВЕННЫМИ НЕЙРОННЫМИ СЕТЯМИ

В.П. Зимин, М.М. Филиппов

Томский политехнический университет

E-mail: zimin@ido.tpu.ru

Поставлена и исследована задача аппроксимации семейства экспериментальных вольтамперных характеристик термоэмиссионного преобразователя слоистыми искусственными нейронными сетями. Получены закономерности и зависимости, позволяющие оценить качество аппроксимации при изменении параметров нейронных сетей, обучающих и тестирующих выборок.

### Постановка задачи

Искусственные нейронные сети (ИНС) обладают нелинейными аппроксимационными свойствами. Их используют для построения моделей в различных научных областях [1, 2]. При калибровке модели, первой фазы вычислительного эксперимента, одним из сложных этапов является сравнение результатов экспериментальных и теоретических исследований с целью проверки адекватности модели. Этот этап важен потому, что результативность следующей фазы вычислительного эксперимента – фазы прогноза, полностью определяется степенью адекватности модели. Поэтому ИНС, как одна из технологий математического моделирования, наряду с технологией вычислительного эксперимента, может эффективно использоваться как для построения модели, так и для проверки ее адекватности [3].

Рассмотрим постановку задачи применения нейросетевой технологии в области исследования прямого преобразования тепловой энергии в электрическую.

Для исследования эффективности преобразования тепловой энергии в электрическую в лабораторных условиях используют цезиевые термоэмиссионные преобразователи (ТЭП). Одним из основных параметров таких преобразователей являются их вольтамперные характеристики (ВАХ)  $J=J(V_H)$ , рис. 1 [4]. В общем случае ВАХ ТЭП зависят как от параметров преобразователя: температур эмиттера  $T_E$  и коллектора  $T_C$ , давления насыщенных паров в резервуаре с цезием  $P_{Cs}$ , межэлектродного зазора  $d$ , работ выхода эмиттера  $F_E$  и коллектора  $F_C$ , так и от состояния (характеристик) межэлектродной плазмы. Генерируемая мощность ТЭП определяется произведением  $JV_H$ .

Достаточно надежно, с небольшой погрешностью, измеряются четыре первых параметра ТЭП. Параметры плазмы измеряются, во-первых, только в зазорах значительно превышающие те, при которых преобразование энергии в ТЭП эффективно и, во-вторых, только для плазменного столба. Так, например, в приэлектродных областях погрешность измерения параметров плазмы может составлять 100 % и более. Наконец, работы выхода электродов  $F_E$ ,  $F_C$  измеряются только косвенно, при условиях,

значительно отличающихся от рабочих условий преобразователя. Параметром подобия в этом случае являются отношения  $T_E/T_{Cs}$ ,  $T_C/T_{Cs}$  ( $T_{Cs}$  – температура резервуара с цезием), т. к. экспериментальные данные с хорошей точностью описываются кривыми Рейзора вида  $F_E=F_E(T_E/T_{Cs})$  и  $F_C=F_C(T_C/T_{Cs})$ . Эмиссионные токи электродов экспоненциально зависят от работ выхода электродов ТЭП, и поэтому существенным образом влияют на его ВАХ. Отметим, что ВАХ ТЭП зависят только от некоторых обобщенных конфигураций параметров плазмы, т. е. интегрально от распределений параметров плазмы.

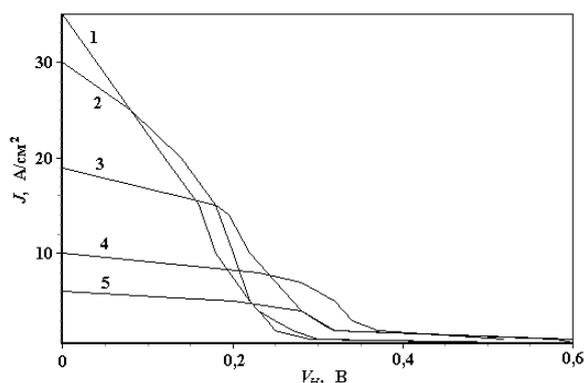


Рис. 1. Экспериментальные ВАХ ТЭП.  $T_E=1600$  К,  $T_C=900$  К,  $d=0,0127$  см,  $P_{Cs}$  (мм рт.ст.): 1) 4,00, 2) 2,80, 3) 2,00, 4) 1,42, 5) 1,00

Суммируя все это, в задаче проверки на адекватность модели процессов ТЭП можно выделить три уровня сложности. Первый уровень проверки на адекватность состоит в сравнении экспериментальных и модельных ВАХ при заданных параметрах преобразователя  $T_E$ ,  $T_C$ ,  $P_{Cs}$ ,  $d$ . На втором уровне дополнительно подключаются параметры  $F_E$ ,  $F_C$ . Третий уровень состоит в сравнении не только интегральных характеристик ВАХ ТЭП, но, главное, в сравнении экспериментальных и модельных распределений параметров плазмы. Для этого требуются дополнительные исследования и получение новых закономерностей, описывающих поведение обобщенных конфигураций параметров плазмы. Второй и третий уровни решения задачи адекватности модели значительно сложнее, чем первый. При решении задачи адекватности третьего уровня

автоматически решается задача адекватности первого уровня, обратное утверждение, вообще говоря, не верно. Включение в критерий адекватности модели всех параметров ТЭП (второй уровень) фактически приводит к решению задачи диагностики работ выхода электродов.

Настоящая статья посвящена исследованию задачи построения модели и проверки ее на адекватность первого уровня с применением нейросетевой технологии. В этом случае необходимо решать задачу аппроксимации ВАХ ТЭП, как функции многих переменных, с помощью ИНС.

#### Методика проведения численных экспериментов

При использовании нейросетевой технологии в первую очередь выбирают архитектуру нейронной сети. В нашем случае была использована слоистая сеть (многослойный перцептрон). Число нейронов во входном слое соответствовало количеству аргументов аппроксимируемых ВАХ ТЭП (от одного до пяти). В выходном слое брался один нейрон. При проведении численных экспериментов число скрытых слоев сети варьировалось. Но обычно брался один слой, так как такая ИНС способна успешно аппроксимировать любую многомерную функцию [1]. Выбор количества нейронов в скрытых слоях до настоящего времени является нерешенной задачей. Бралась интервальная оценка количества нейронов в одном скрытом слое, полученная эмпирическим путем [1]. Данная оценка зависит от количества нейронов во входном и выходном слоях, а так же от величины выборки обучения. Все имеющиеся примеры делились на выборки для обучения и тестирования сети. Отношение мощностей этих выборок  $M_{tr}:M_{ot}$  обычно было приблизительно как 2:1. Сами выборки формировались таким образом, чтобы примеры из них равномерно распределялись (перемешивались) в интервалах изменения компонент входного вектора.

Важную роль играет выбор функций активации формальных нейронов и методов обучения ИНС. Эти параметры и методы сети тесно связаны с используемыми нейроимитаторами. При решении нашей задачи использовался ряд доступных нейроимитаторов [5]. Наилучшие результаты были получены с использованием Neural Network Toolbox (NTT) Matlab [6]. Все излагаемые ниже результаты получены именно с его помощью.

Качество обучения ИНС контролировалось как по величине ошибки (квадраты отклонения между выходом сети и заданной величиной на выходе, суммированные для всех примеров выборки), так и по получаемому нейронной сетью виду ВАХ ТЭП. Последнее обуславливается тем, что при решении задачи важным является не только минимальное отклонение в каждой точке характеристики, но и поведение совокупности точек, т. е. зависимости  $J=J(V_H)$  для конкретных параметров преобразователя.

Значительное внимание при проведении численных экспериментов, особенно на первом эта-

пе, уделялось проверочным (калибровочным) расчетам. С этой целью брались одномерные как элементарных функции – линейные, квадратичные, экспоненциальные, так и более сложные композиции функций – экспоненциальные и тригонометрические. Последние позволяли моделировать совместно колебательные и быстрорастущие участки функциональных зависимостей.

#### Результаты моделирования с использованием ИНС

Анализ вида кривых семейства ВАХ ТЭП (рис. 1) позволил сформулировать несколько проверочных подзадач аппроксимации с помощью ИНС: аппроксимацию линейной функции  $y=ax+b$ , квадратичной –  $y=ax^2+bx+c$  и экспоненциальной –  $y=Ae^{ax}$ . Выбор такого вида функций обусловлен тем, что некоторые участки ВАХ описываются подобными зависимостями. Для функции вида  $y=Ae^{-ax}\sin(\omega_1x)+Be^{-bx}\sin(\omega_2x)$  были поставлены задачи аппроксимации ИНС и исследования эффекта переобучения. Указанные выше тестовые функции были оцифрованы (150 точек), а затем сформированы обучающая и тестирующая выборки ( $M_{tr}:M_{ot}=2:1$ ). Кроме этого, к примерам обучающей и тестирующей выборки добавлялся шум в соответствии с указаниями, изложенными в [7]. Брался один скрытый слой ИНС, количество нейронов в котором варьировалось от 5 до 100. Для скрытого и выходного слоев брались функции активации гиперболический тангенс – *tansig*.

Для успешной аппроксимации при отсутствии шума достаточно нескольких (5–15) нейронов в скрытом слое. Результат аппроксимации мало зависит от выбора функции обучения ИНС. В случае наложения шума рельеф функций существенно изменяется, становится менее гладким. Фактически появляется в разложении высокочастотные составляющие. Это приводит к тому, что при заданной ошибке обучения (на обучающей выборке) ошибка обобщения (на тестирующей выборке) зависит от числа нейронов в скрытом слое, немонотонным образом. Имеются минимумы этой ошибки как при малом, так и при большом количестве нейронов. Фактически появление шума усиливает эффект переобучения.

Изучалась аппроксимация семейства экспериментальных ВАХ ТЭП, полученных в [4] при  $T_E=1600, 1700, 1800, 1900, 2000$  К,  $E_C=900, 950$  К,  $P_{CS}=1,00; 1,42; 2,00; 2,83; 4,00; 5,67; 8,00; 11,00; 16,00$  мм рт.ст. и  $d=0,0013; 0,0038; 0,0068; 0,0127; 0,0254; 0,0517; 0,1016$  см. Характеристики были известны не при всех возможных сочетаниях параметров преобразователя, т. е., фактически, разбиение по параметрам преобразователя было с пропусками, нерегулярным. Серия из 183 ВАХ была оцифрована, и все точки разделены на обучающую ( $M_{tr}=820$ ) и тестирующую ( $M_{ot}=326$ ) выборки. Топология ИНС представлялась многослойным перцептроном, с пятью входными нейронами – по количеству переменных функции  $J=J(V_H, T_E, T_C, P_{CS}, d)$  и одним выходным ней-

роном. Задавался один или два скрытых слоя, количество нейронов в которых варьировалось [8].

Для обучения нейросети вначале бралась функция *traincfg*, которая использует метод оптимизации Флетчера – Пауэлла. Критериями остановки процесса обучения являлись достижения минимальной ошибки обучения  $\varepsilon_{\min}=10^{-7}$  или достижение максимального числа эпох обучения  $N_E=10^4$ . При такой величине  $N_E$  время обучения нейронной сети составляло несколько часов. Но для функции *traincfg* процесс обучения завершался вследствие достижения минимального шага алгоритма обучения.

Зависимости ошибок обучения  $\varepsilon_{tr}$  и обобщения  $\varepsilon_{tot}$  от количества нейронов в одном скрытом слое представлены на рис. 2, кривые 1 и 2. Данные зависимости немонотонные, причем минимум одновременно достигается при  $N_{H1}=800$ . По теоретическим представлениям [1, 2] количество нейронов не должно быть больше количества обучающих примеров, т. к. проявляется эффект переобучения. Но в данном случае при  $N_{H1}=800$  нейронах получено хорошее качество обобщения. Удовлетворительные результаты, минимальные ошибки обучения и обобщения, удалось получить для ИНС, содержащей  $N_{H1}=200$  нейронов.

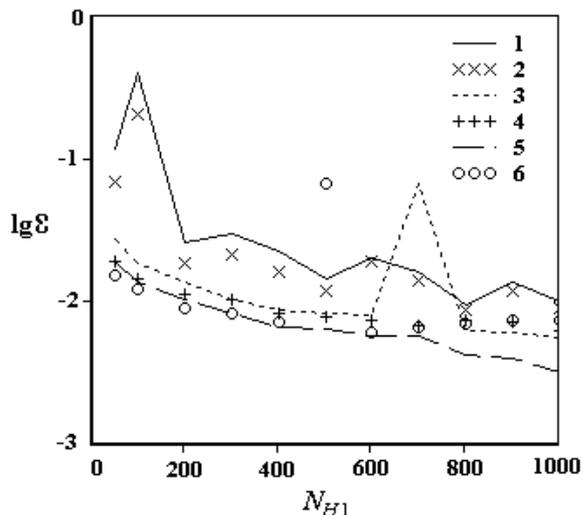


Рис. 2. Зависимость ошибок обучения (кривые 1, 3, 5) и обобщения (кривые 2, 4, 6) от числа нейронов в скрытом слое

Для второй функции обучения *trainscg* (алгоритм масштабируемых сопряженных градиентов) обучение завершалось по достижению максимального количества эпох обучения  $N_E=10^4$ . Использование *trainscg* позволило повысить качество аппроксимации. Зависимость ошибок обучения  $\varepsilon_{tr}$  и обобщения  $\varepsilon_{tot}$  от количества нейронов в скрытом слое представлена на рис. 2, кривые 3 и 4, соответственно.

Наилучшие результаты получены при использовании ИНС с  $N_{H1}=1000$ . Минимальное количество нейронов, при котором получены одновременно минимальные ошибки  $N_{H1}=600$ . Обученная ИНС с функцией с большей степенью точности способна аппроксимировать ВАХ ТЭП.

Таблица. Зависимости ошибок обучения и обобщения при аппроксимации ВАХ двухслойной ИНС

$N_{H1}$	$N_{H2}$	$\varepsilon_{tr}$	$\varepsilon_{tot}$
100	50	0,00764	0,00749
<b>100</b>	<b>100</b>	<b>0,00712</b>	<b>0,00691</b>
100	300	0,06842	0,00699
100	500	0,00830	0,00732
300	50	0,00188	0,00438
300	100	0,02363	0,00532
300	300	0,00309	0,05263
300	500	0,49184	0,20584
500	50	0,00089	0,00518
500	100	0,00078	0,00528
500	300	0,00149	0,00587
500	500	0,49184	0,20584
700	50	0,00075	0,00383
700	100	0,00057	0,00561
700	300	0,00057	0,00664
700	500	0,00080	0,00487
<b>1000</b>	<b>50</b>	<b>0,00025</b>	<b>0,00689</b>
1000	100	0,00028	0,00825
1000	300	0,00057	0,00750
1000	500	0,00049	0,00740

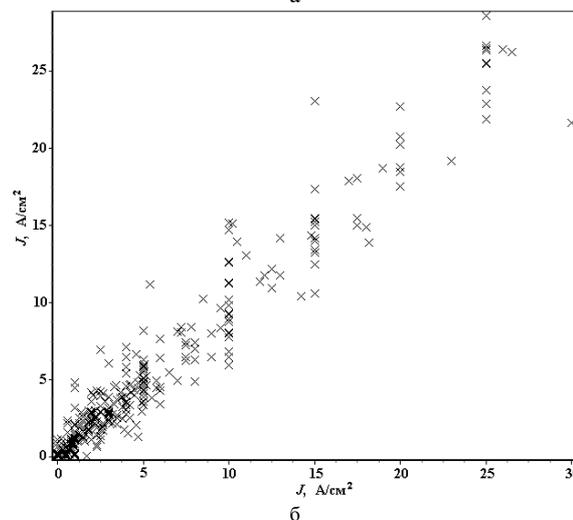
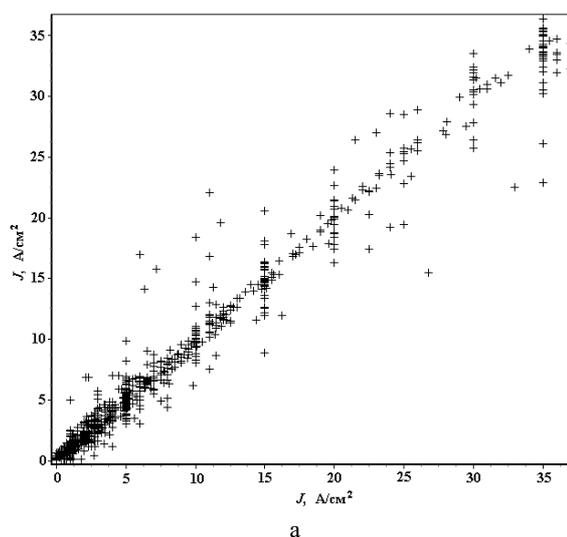
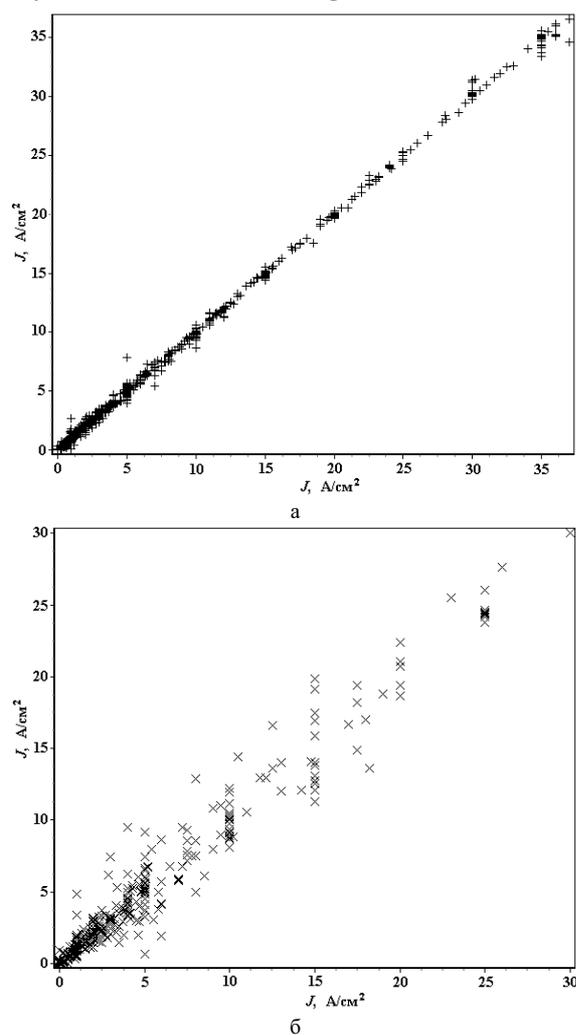


Рис. 3. Характеристика качества аппроксимации ИНС с одним скрытым слоем: а) для выборки обучения; б) тестирования

Значительное увеличение количества эпох обучения  $N_E=5 \cdot 10^4$  привело к некоторому уменьшению ошибки обучения  $\varepsilon_{tr}$  (рис. 2, кривая 5). Точность аппроксимации тестовой выборки практически изменилась мало, скорость обучения заметно уменьшилась. Зависимость ошибки обобщения  $\varepsilon_{tot}$  (рис. 2, кривая 6) имеет выраженный пологий минимум в районе  $N_{H1}=600$ . На рис. 3 для  $N_{H1}=600$  представлены результаты сравнения экспериментальных значений (абсцисса) и значений аппроксимации ИНС (ордината) в каждой точке ВАХ для обучающей и тестовой выборок.



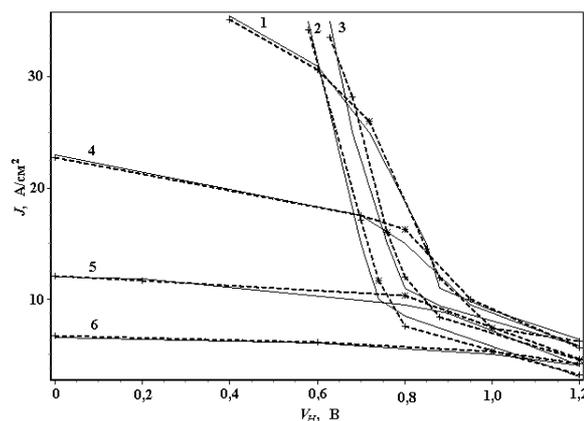
**Рис. 4.** Характеристика качества аппроксимации ИНС с двумя скрытыми слоями: а) для выборки обучения; б) тестирования

Для предыдущих описанных численных экспериментов с ВАХ ТЭП отклонение от прямой  $y=x$  было еще большим, особенно для тестовой выборки.

Дальнейшие исследования проводились с добавлением второго скрытого слоя ИНС и с функцией обучения *trainscg*. Для  $N_E=10^4$  процесс обуче-

ния оканчивался по достижении максимального количества эпох, но точность аппроксимации повысилась незначительно. Наилучшие результаты,  $\varepsilon_{tr}=0,003524$  и  $\varepsilon_{tot}=0,004733$ , получены для ИНС, имеющей  $N_{H1}=700$ ,  $N_{H2}=300$ . Увеличение эпох обучения  $N_E=5 \cdot 10^4$  привело к усилению обобщающих способностей ИНС (см. табл.). Значение  $\varepsilon_{tr}$  стало на порядок ниже, значение  $\varepsilon_{tot}$  практически не изменилось. Графические результаты сравнения аппроксимации обучающей и тестовой выборок с истинными значениями для каждой точки ВАХ при  $N_{H1}=100$ ,  $N_{H2}=100$  представлены на рис. 4. При сравнении с рис. 3 видно, что разброс как для тестирующей, так и обучающей выборок уменьшился, причем для последней – значительно.

На рис. 5 представлены характерные результаты аппроксимации (пунктирные линии) семейства экспериментальных ВАХ ТЭП нейронной сетью, имеющей два скрытых слоя  $N_{H1}=100$ ,  $N_{H2}=100$ . Аппроксимация остальных ВАХ имеет подобную картину. Из рис. 5 видно, что имеется хорошее совпадение не только для значений зависимостей  $J=J(V_H)$ , но и удовлетворительная аппроксимация дифференциальных сопротивлений.



**Рис. 5.** Результаты аппроксимации ВАХ ИНС с двумя скрытыми слоями. Параметры ТЭП:  $T_e=2000$  К,  $T_c=900$  К,  $d=0,0127$  см,  $R_c$  (мм рт.ст.): 1) 16; 2) 11; 3) 8; 4) 5; 5) 4; 6) 2; 8

### Обсуждение результатов

Из численных экспериментов с проверочными (калибровочными) функциями следует, что перед использованием ИНС необходимо исключить имеющийся в экспериментальных данных шум. Результаты аппроксимации проверочных функций с шумом в некоторой мере противоречат ранее полученным [1], согласно которым ИНС успешно аппроксимирует зашумленные данные. Поэтому требуется дальнейшее исследование этого вопроса.

Аппроксимация ИНС семейства экспериментальных ВАХ ТЭП как нелинейной функции пяти аргументов позволила создать модель и изучить ее характерное поведение.

Зависимости ошибок обучения и обобщения от числа нейронов в скрытых слоях имеют немонотонный характер вследствие сложного характера оптимизируемой функции (множество локальных минимумов) и большого числа параметров оптимизации. Из-за сложного характера областей притяжения этих минимумов практически невозможно достичь глобального оптимума.

Существенное влияние на аппроксимацию ВАХ ИНС оказывает индивидуальная оцифровка каждой ВАХ и нерегулярность разбиения по параметрам преобразователя. В последнем случае фактически решается задача с пропусками данных. Требуется дальнейшее изучение аппроксимации ВАХ для регулярной сетки.

Время численного эксперимента составляло часы работы высокопроизводительной ЭВМ. Ускорение этого процесса возможно при применении специализированных нейроимитаторов или нейрокомпьютеров.

Подходы, использованные авторами при аппроксимации ВАХ ТЭП, могут быть применены для построения моделей характеристик тепловых, гидродинамических и иных процессов.

## Выводы

Поставлена задача аппроксимации ВАХ термоэмиссионных преобразователей искусственными нейронными сетями, из решения которой следует:

- Перед аппроксимацией сложных нелинейных зависимостей необходимо путем их предварительного анализа выделять характерные зависимости, составлять тестовые функции и проводить калибровочные численные эксперименты.
- Выбор нейроимитатора, методов обучения и количества скрытых слоев нейросети значительно влияет на точность аппроксимации. Это позволяет достичь высокой точности аппроксимации без изменения топологии сети.
- В связи с наличием множественных локальных минимумов функции оптимизации зависимость ошибок обучения и обобщения от параметров нейросети носит немонотонный характер.
- Увеличение количества эпох обучения нейросети приводит к медленному уменьшению ошибки обучения, хотя ошибка обобщения остается практически неизменной.

## СПИСОК ЛИТЕРАТУРЫ

1. Хайкин С. Нейронные сети: полный курс. – 2-е изд. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
2. Осовский С. Нейронные сети для обработки информации: Пер. с польск. – М.: Финансы и статистика, 2002. – 344 с.
3. Зимин В.П. Применение искусственных нейронных сетей в технологии вычислительного эксперимента // Вычислительные технологии. – 2004. – Т. 9. – Новосибирск; Вестник КазНУ им. Аль-Фараби. Серия математика, механика, информатика. – 2004. – № 3 (42). – Алматы. – С. 231–234.
4. Ruffe F., Lieb D. The output characteristics of a vapor-deposited tungsten emitter and a molybdenum collector // Report TE 4086-3-70. – NY, 1969. – 112 p.
5. Зимин В.П., Филиппов М.М. Исследование аппроксимационных возможностей многослойных нейронных сетей в некоторых нейроимитаторах // Молодежь и современные информационные технологии: Труды IV Всерос. научно-практ. конф. студентов, аспирантов и молодых ученых. – Томск: Изд-во ТПУ, 2006. – С. 157–158.
6. Математический пакет Matlab. – Режим доступа: <http://www.mathworks.com/>, вход свободный.
7. Тютерев В.В. Метод эволюционного накопления признаков для автоматического построения нейронных сетей // Вычислительные методы и программирование. – 2001. – № 2. – С. 88–108.
8. Зимин В.П., Филиппов М.М. Исследование влияния параметров ИНС на аппроксимацию ВАХ ТЭП // Молодежь и современные информационные технологии: Труды V Всерос. научно-практ. конф. студентов, аспирантов и молодых ученых. – Томск: Изд-во ТПУ, 2007. – С. 183–187.

*Поступила 04.04.2008 г.*

## Ключевые слова:

Вольтамперная характеристика, термоэмиссионный преобразователь энергии, искусственная нейронная сеть, аппроксимация, выборка обучения и тестирования.