

УДК 519.71

Д.В. ПАВЛОВСКАЯ

*Черниговский государственный технологический университет, Украина***ПРОГНОЗИРОВАНИЕ КАК СПОСОБ ОБЕСПЕЧЕНИЯ НАДЕЖНОСТИ  
ВЫПОЛНЕНИЯ ЗАДАЧ В РАСПРЕДЕЛЕННОЙ СРЕДЕ**

Рассмотрена задача прогнозирования состояний распределенной среды, как способ обеспечения надежности выполнения задач большой размерности. Рассмотрен нейросетевой подход к прогнозированию состояний распределенной среды. Приведена архитектура нейронной сети прогнозирования состояний распределенной среды.

**распределенная среда, искусственная нейронная сеть, нейрон, прогнозирование****Введение**

Для обеспечения надежности выполнения задач в распределенной среде (РС) предложена модель прогнозирования состояния РС в заданный момент времени.

Целью прогнозирования состояния РС является определение с заданной вероятностью интервала времени, на протяжении которого сохраняется производительность РС, или вероятность конкретного значения производительности РС на заданный интервал времени. Итак, под основной целью понимаем обеспечение надежности работы вычислительных узлов (ВУ), которая получена расчетом вероятности сохранения производительности объекта прогнозирования (ОП) на заданном интервале времени. Заметим, что интервал времени в этом случае задается предварительно, а не определяется в процессе прогнозирования.

При решении задачи распределения нагрузки между ВУ РС возникает проблема расчёта предполагаемых ресурсов РС. Это приводит к необходимости использования системы прогнозирования.

Обычно в прогнозах указывается вероятная степень отклонения от тех или иных целей в зависимости от способа будущих действий и влияния различных объективных и субъективных факторов.

Известные методы прогнозирования (аналитические, вероятностные и статистические) основываются

на математическом аппарате и являются довольно громоздкими. Сложность методов повышается с увеличением количества параметров прогнозирования, но только использование достаточного количества параметров гарантирует высокую точность прогноза. В настоящее время использование искусственных нейронных сетей (ИНС) является наиболее актуальным и перспективным средством для решения задач прогнозирования состояний технических объектов [1].

Поэтому для упрощения механизма прогнозирования, а также для улучшения результатов прогнозирования состояний РС, выбран метод решения с использованием искусственных нейронных сетей.

Под ОП будем понимать РС, производительность которой прогнозируется. Все необходимые данные для процесса прогнозирования будут получены соответствующими средствами с использованием необходимых математических моделей и методов, а средством прогнозирования является ИНС.

Классическим подходом к прогнозированию производительности РС является использование параметров ВУ. Под использованием понимается контроль и наблюдение динамики их изменения. Именно динамика изменения является основанием для вынесения окончательного решения о будущем состоянии ОП на основе известных математических методов и моделей.

Поскольку принятым раньше средством прогнозирования есть ИНС, то необходимо определить ее архитектуру. На сегодня существуют лишь самые общие рекомендации относительно выбора начальной архитектуры сети. Архитектура ИНС должна давать возможность использования одной и той же сети для прогнозирования технического состояния РС различной конфигурации. Так как в РС возможны ситуации, когда необходимо конкретизировать будущее состояние ненадежных узлов, то архитектура сети должна обеспечивать прогнозирование состояния не только РС, а и отдельно каждого ВУ.

Предметом исследования является процесс прогнозирования работоспособности и производительности РС, поэтому важными задачами являются разработка модели среды, которая должна иметь способность отображения характерных особенностей этого процесса, и разработка математических моделей РС, как ОП.

### Результаты исследований

Весь процесс прогнозирования состояния РС разделим на две основных части:

- синтез архитектуры и определение параметров ИНС;
- непосредственное прогнозирование работоспособности и производительности РС по входным параметрам с использованием полученной на предыдущем шаге ИНС.

Синтез архитектуры и определение параметров ИНС происходит одновременно и в дальнейшем сеть многократно используется для прогнозирования состояния РС.

Процесс синтеза архитектуры и определение параметров ИНС состоит из таких этапов (рис. 1):

$E_1$  – определение класса ОП, для которого будет разрабатываться ИНС прогнозирования;

$E_2$  – формирование множества параметров класса ОП по критериям оптимальной информативности и доступности;

$E_3$  – синтез архитектуры ИНС прогнозирования в соответствии с введенным классом и избранными параметрами;

$E_4$  – формирование учебной выборки необходимого объема для полученной на предыдущем шаге ИНС;

$E_5$  – обучение ИНС;

$E_6$  – верификация процесса обучения;

$E_7$  – завершение процесса синтеза ИНС, определение параметров сети и их представление для сохранения и передачи на следующие этапы.

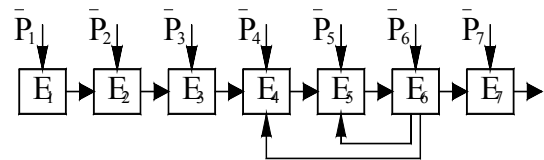


Рис. 1. Формализованная схема процесса синтеза архитектуры и определение параметров ИНС прогнозирования состояния

Для реализации каждого из этапов необходимо определить соответствующие параметры, которые представляются множеством векторов

$$P_{син} = \{\bar{P}_1, \bar{P}_2, \dots, \bar{P}_7\},$$

где каждый вектор состоит из множества параметров соответствующего этапа:  $\bar{P}_i = \{p_1^i, p_2^i, \dots, p_{n_i}^i\}$ ,  $n_i$  – количество параметров в этапе,  $i = 1..7$ .

На этапе  $E_7$  получаем ИНС прогнозирования, готовую к определению состояния РС.

Вторая часть процесса прогнозирования состоит из таких этапов (рис. 2):

$E_8$  – сбор исходной информации об ОП и представление ее в необходимом формате в соответствии с этапами  $E_1$  и  $E_2$ ;

$E_9$  – формирование входного вектора для прогнозирующей ИНС на основе полученной информации на предыдущем этапе и заданного интервала времени, на который необходимо получить прогноз;

$E_{10}$  – представление входного вектора и снятие с выходов ИНС соответствующих реакций;

$E_{11}$  – накопление информации для дообучения ИНС. Этап необходимый в том случае, когда есть возможность кроме сбора параметров делать вывод о текущем состоянии и уровне производительности ОП;

$E_{12}$  – интерпретация полученных результатов для пользователя в зависимости от уровня его квалификации и по необходимости.

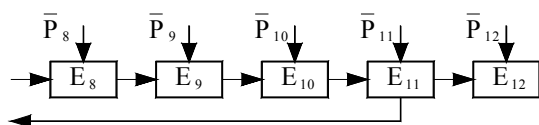


Рис. 2. Формализованная схема процесса прогнозирования состояния и производительности РС при помощи ИНС

Аналогично процессу синтеза, для реализации второго этапа процесса прогнозирования необходимые параметры, представляются множеством векторов  $P_{np} = \{\bar{P}_8, \bar{P}_9, \dots, \bar{P}_{12}\}$ , где каждый вектор состоит из множества параметров соответствующего этапа:  $\bar{P}_i = \{p_1^i, p_2^i, \dots, p_{n_i}^i\}$ ,  $n_i$  – количество параметров в этапе,  $i = 8..12$ .

Таким образом, обобщенный процесс прогнозирования содержит две основные отдельные части [2]: синтез ИНС прогнозирования состояния и производительности РС и непосредственно прогнозирование. Связь между двумя частями осуществляется с помощью данных, которые передаются из этапа  $E_7$  к  $E_8$  (архитектура ИНС, формат параметров ОП и тому подобное) и данных, которые могут передаваться в случае необходимости из этапа  $E_{11}$  к  $E_5$  (дополнительная учебная выборка).

Основным принципом прогнозирования является принцип непрерывности. Он является основой эффективного решения задачи прогнозирования. Источником объективной информации в этом случае

считается получение параметров объекта, которые определяют его текущее состояние. Именно эта информация используется для прогнозирования состояния РС в период ее функционирования [3].

Состояние ряда составляющих РС изменяется со временем, соответственно, использование устаревшей информации может привести к тому, что модель прогнозирования станет неадекватной действительности.

В период функционирования РС существенным образом увеличивается количество прецедентов, которые приводят к значительному увеличению размерности ИНС. Эта ситуация требует использования соответствующих архитектур ИНС, которые бы разрешали получать прогноз состояния РС с необходимой точностью и достоверностью.

С целью выбора необходимой структуры проведено сравнение трех базовых архитектур ИНС, которые, наиболее подходят для разработки модели прогнозирования состояния РС. Каждая из архитектур имеет свои особенности, которые накладывают определенные ограничения. В частности, к недостаткам ИНС с обратными связями можно отнести непредсказуемость структуры фазового пространства сети. Недостатком самоорганизующихся карт есть их продолжительное обучение. В ИНС прямого распространения невозможно спрогнозировать время их обучения.

В результате проведенного сравнения, как наиболее оптимальные, для решения задачи прогнозирования состояния РС на этапе функционирования определены полносвязные ИНС прямого распространения. Полносвязные ИНС прямого распространения обеспечивают возможность запоминания существенного количества образов, а так же реализуют распознавание обобщением на основе сходства и преимуществ. Сети прямого распространения способны формировать правильный вывод, относительно ситуаций, которые ранее не наблюдались. Большой объем учебной выборки, необходимый для

полносвязных ИНС прямого распространения, достаточно быстро накапливается на этапе эксплуатации РС. В свою очередь, время обучения ИНС для использования их в системах прогнозирования состояния РС на этапе функционирования не является критическим, поскольку обучение на большой выборке будет проводиться одновременно, а со временем будет проводиться лишь дообучение, при условиях формирования достаточно объемной и информативной учебной выборки.

В общем виде реализация прогнозирования состояния РС, с использованием ИНС прямого распространения, состоит из пяти основных этапов: сбор исходных данных и их представление в упорядоченной форме; синтез архитектуры ИНС; синтез модели прогнозирования состояния РС путем обучения ИНС; верификация процесса обучения; получение прогноза.

Первый этап состоит из двух основных задач: сбор, накопление исходных данных и представление их в упорядоченной форме.

Второй этап целесообразно разделить на следующие подзадачи, решение которых приведет к необходимому результату: определение количества слоев ИНС прямого распространения; определение количества нейронов в каждом слое; выбор функции активации нейронов каждого слоя.

Параметры архитектуры выбираются, исходя из уже известных и достаточно изученных конфигураций для синтеза ИНС с наилучшей прогнозирующей способностью.

Наиболее приемлемой можно считать ИНС с двумя скрытыми слоями, поскольку увеличение количества скрытых слоев больше двух не дает важного улучшения точности прогнозирования, увеличивая при этом трудоемкость обучения. Первый скрытый слой выполняет функцию «грубой» аппроксимации, вместе с тем второй скрытый слой есть уточняющим, в нем конкретизируются значение основных признаков, которые являются решающими для конечного прогноза.

Количество элементов первого и второго скрытых слоев ИНС уточняется экспериментально. В общем, количество нейронов зависит от объема учебной выборки, ее характера, необходимой точности восстановления ситуации. Поэтому она решается тогда, когда известен конкретный класс ОП.

Для определения необходимых параметров прогнозной модели можно использовать как алгоритм обратного распространения ошибки, так и его модификации. Целью обучения ИНС, которая прогнозирует состояние РС, есть подбор таких весовых коэффициентов и векторов смещения, которые бы при подаче на вход вектора причин минимизировали среднеквадратичную погрешность обучения.

Верификация при прогнозировании трактуется как этап, на котором выполняется оценка достоверности или точности прогнозирования. Методы верификации направлены больше на оценку метода, чем на оценку качества самого результата прогнозирования.

Процедура получения результатов прогноза заключается в подаче на входной слой обученной ИНС вектора причин определенной выборки, которая используется для обучения и верификации, а также получения соответствующих сигналов (отклика) в выходном слое, который характеризует оценку прогнозирования.

Процесс синтеза архитектуры целесообразно разделить на следующие этапы, решение которых приведет к необходимому результату:

- 1) определение количества слоев  $L$  ИНС прямого распространения;
- 2) определение количества нейронов  $z_i$  в каждом слое,  $i = 1, \dots, L$ ;
- 3) выбор функции  $f_i$  активации нейронов каждого слоя,  $i = 1, \dots, L$ .

Как указано выше, базовой архитектурой есть полносвязная ИНС прямого распространения. Обобщенный вид архитектуры сети представлен на рис. 3.

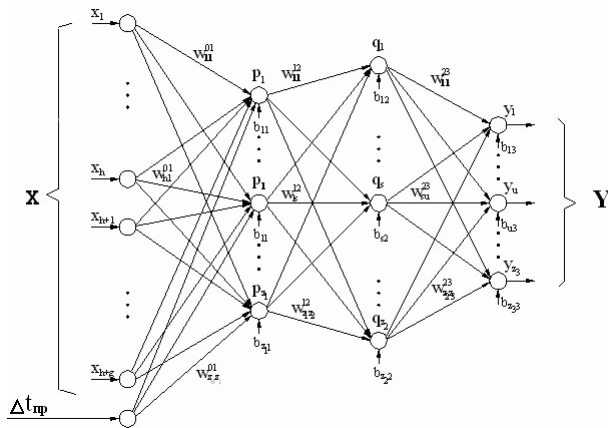


Рис. 3. Обобщенная архитектура ИНС прогнозирования состояния РС

Параметры архитектуры выбирались исходя из уже известных и достаточно изученных конфигураций для синтеза ИНС с наилучшей прогнозирующей способностью. Количество слоев нейронов выбиралось согласно требованиям, которые выдвигались к способностям ИНС. Как наиболее приемлемая была выбрана ИНС с двумя скрытыми слоями. Таким образом, общее количество слоев сети будет составлять  $L = 4$  входной слой, два скрытых и выходной. Первый скрытый слой выполняет функцию «грубой» аппроксимации, второй скрытый слой есть уточняющий.

В случае необходимости определения времени понижения производительности системы в целом до 0, предложенную модель необходимо дополнить выходным слоем, который состоит из одного искусственного нейрона [4]. В общем виде дополненная архитектура ИНС представлена на рис. 4.

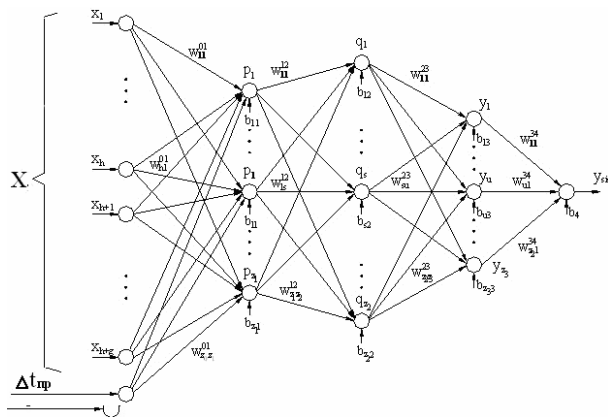


Рис. 4. Дополненная модель ИНС прогнозирования состояния РС

## Выводы

Таким образом, предложенная концепция прогнозирования состояния РС состоит в применении ИНС как средства моделирования прогноза, в которой расширен класс параметров прогнозирования, разделенных на общие и конкретные.

Разработанная нейросетевая модель процесса прогнозирования состояния РС дает возможность после обучения использовать ИНС для прогнозирования состояния систем определенного класса.

Определен вид параметра оценки достоверности прогнозирования, который базируется на достоверности функционирования РС, что определяется вероятностями нахождения системы в работоспособном и неработоспособном состояниях. Достоверность прогноза в таком случае зависима от точности прогнозирования ИНС и промежутка времени, через которое система перейдет в неработоспособное состояние. Показано, что достоверность прогнозирования увеличивается с улучшением точности обучения.

## Литература

1. Заець А.В., Піскун А.С., Павловська Д.В. Планування та розподілення ресурсів в системі паралельних обчислень. // Матеріали 1-ої Міжн. конф. молодих науковців CSE-2006. – Львів, 2006. – С. 2.
2. Руденко О.Г., Бодянский Е.В. Искусственные нейронные сети: Учебное пособие. – Х.: ООО «Компания СМІТ», 2005. – 408 с.
3. Галеев Р.І., Победа А.О.. Модель прогнозування станів розподіленого середовища // Матеріали 2-ої Міжн. конф. молодих науковців CSE-2007. – Львів. – С. 3.
4. Руденко О.Г. Использование нейросетей для прогнозирования состояний распределенной среды // Мат. межд. НПК "ИНФОТЕХ-2007". Ч. 2. – С. 3.

Поступила в редакцию 8.01.2008

**Рецензент:** канд. техн. наук, А.В. Горбенко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.