

С.В.Добрецов (асп., каф. АиВТ), Л.К.Птицына, д. т.н., проф.

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АПРИОРНОЙ ОЦЕНКИ ПОКАЗАТЕЛЕЙ ПРОИЗВОДИТЕЛЬНОСТИ АЛГОРИТМОВ НЕЛИНЕЙНОГО ПЛАНИРОВАНИЯ ДЕЙСТВИЙ

Как было показано в [1], динамический выбор оптимального алгоритма генерации плана действий, совершаемый в зависимости от специфики решаемой задачи планирования, способен в ряде случаев существенно повысить эффективность работы интеллектуального агента. Поскольку необходимым условием практического применения данного подхода является возможность выбора оптимального алгоритма до непосредственного этапа построения плана действий, предлагается использовать прогнозируемые значения показателей эффективности, входящих в критерий оптимальности решения. Прогнозирование (априорная оценка) осуществляется на основании информации, известной агенту в момент принятия решения.

Для проверки реализуемости данного подхода в качестве средства оценки значения отдельного показателя эффективности алгоритма планирования использовалась простейшая нейросетевая модель $R_N(x)$ на базе персептрона (сети прямого распространения) с одним скрытым слоем и единственным выходным элементом, состоящая из следующих компонент:

1. Входной слой $\{P_k\}$, выполняющий в случае необходимости преобработка исходных данных $\{I_k\}$. Количество узлов N_{in} равно количеству входных параметров.
2. Скрытый слой, содержащий узлы с заданной функцией активации $A(x)$. Ввиду специфики используемого программного обеспечения в качестве функции активации использовался гиперболический тангенс. Следует отметить, что в [2] доказано, что в общем случае единственным существенным требованием к виду функции активации искусственного нейрона является нелинейность; при выполнении этого условия для любой непрерывной функции нескольких переменных можно построить ИНС, которая вычисляет данную функцию с любой заданной точностью. При этом в соответствии с [3] использование нечетной активационной функции, какой и является $th(x)$, ускоряет процесс обучения.
3. Выходной слой, обеспечивающий постобработку результатов в зависимости от решаемой задачи (дискретизация, восстановление исходного масштаба и т.д.). В зависимости от типа регрессионной задачи (оценка непосредственно значений показателя или оценка выполнимости условия ϵ -оптимальности) в качестве активационной функции выбиралась линейная либо логистическая функция соответственно.

В качестве входного набора использованы полное множество информационных показателей I_A , характеризующих среду функционирования агента и условия задачи планирования, и подмножество значимых показателей $I_A^* \subset I_A$, выбранное с использованием статистического критерия хи-квадрат.

Одним из ключевых вопросов, возникающих при построении систем с использованием двухслойного персептрона, является выбор оптимального размера скрытого слоя N_{hid} . В данной работе была применена адаптивная процедура выбора размера скрытого слоя, заключающаяся в обучении набора сетей с разными размерами скрытого слоя и выборе N_{hid} из условия минимизации контрольной ошибки E'_{av} , вычисляемой на некотором сравнительном множестве задач. На рис. 1 представлена зависимость E'_{av} от узлов в скрытом слое для $1 \leq N_{hid} \leq 12$ (верхняя граница была выбрана как соответствующая максимальному рассматриваемому значению N_{in} , достигаемому при входном наборе I_A). Как видно из

полученных данных, в обоих случаях следует признать оптимальным значением $N_{hid}=5$, причем тот факт, что оптимальное значение N_{hid} не зависит от размера входного множества, может рассматриваться как дополнительное подтверждение корректности процедуры выбора множества I_A^* .

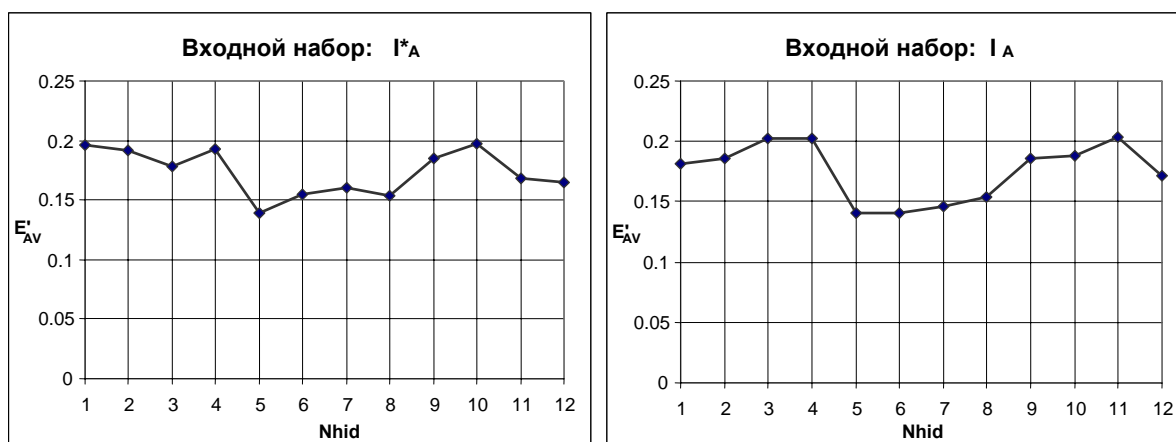


Рис. 1. Зависимость E'_{av} от размера скрытого слоя N_{hid}

Численное моделирование проводилось с использованием пакета Matlab версии 6.0 и библиотеки Netlab 3.2, предоставляющей набор базовых процедур для симуляции ИНС. Ввиду относительно небольших размеров сети и обучающего набора данных вместо стандартного для перцептрона метода обучения с обратным распространением ошибки применялась процедура обучения ИНС на основе минимизации функции ошибки с использованием обобщенных методов оптимизации, в частности, квази-ньютоновского алгоритма. Процесс обучения заканчивался в случае, когда относительное изменение функции ошибки не превосходило 10^{-8} на протяжении 10 последовательных итераций.

Качество результатов применения нейросетевого подхода оказалось сопоставимо с результатами, полученным при использовании квадратичной регрессии, несмотря на значительную разницу в количестве степеней свободы, что свидетельствует о более высокой эффективности нейросетевой модели. Как и следовало ожидать, ошибки обучения при использовании множества I^* заметно ниже, чем в случае I_A^* , что объясняется более чем 50% разницей в числе синаптических связей. Относительно большие значения максимальной ошибки на всем тестовом множестве свидетельствуют о существовании задач, в которых используемая модель не смогла произвести адекватную оценку показателя. В то же время сравнительно низкие значения средней ошибки говорят о незначительной доли подобных случаев, что позволяет признать данную модель удовлетворительной. Суммируя результаты решения обоих типов регрессионных задач, можно сделать следующие выводы:

1. Предпочтительной областью применения данного метода являются качественные оценки эффективности алгоритмов планирования (оценка ε -оптимальности).
2. Расширение набора входных данных с I_A^* до I_A не оказывает существенного влияния на качество оценки, что может рассматриваться как подтверждение адекватности используемого критерия выбора I_A^* .

ЛИТЕРАТУРА:

1. Добрецов С.В. Сравнительный анализ нелинейных алгоритмов планирования действий интеллектуальных агентов // Труды VII международной конференции "Информационные сети, системы и технологии". Том 2. Мн.: БГЭУ, 2001. С. 109-114.
2. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей // Сибирский журнал вычислительной математики. 1998. Т. 1, N 1. С. 12-24.
3. Swingler, K. Applying Neural Networks: A Practical Guide. London: Academic Press. 1996. p.303.