

В. Ю. Осипов, д-р техн. наук, проф., вед. науч. сотр., osipov_vasily@mail.ru,
Санкт-Петербургский институт информатики и автоматизации РАН

Нейросетевое прогнозирование событий для интеллектуальных роботов

С точки зрения интеллектуализации автономных роботов рассматривается задача прогнозирования событий без предварительного задания моделей их развития в целях расширения функциональных возможностей по прогнозированию событий с изменяющимися законами их развития. Для решения этой задачи предлагается новый метод нейросетевого прогнозирования, предусматривающий использование рекуррентной нейронной сети с управляемыми синапсами, со структурой слоев в виде двойной спирали. В основе метода лежит ассоциативное запоминание текущих и задержанных входных сигналов и извлечение из памяти нейронной сети последовательности будущих событий. Приведены результаты моделирования, отражающие особенности такого прогнозирования. Сформулированы рекомендации по применению предложенного метода.

Ключевые слова: прогнозирование, нейронная сеть, сигналы, события, автономный робот

Введение

Успешное решение многих прикладных задач интеллектуальными роботами различного назначения непосредственно связано с прогнозированием событий [1]. В качестве таких событий могут выступать привязанные к пространству и времени совокупности фактов. Чем жестче требования реального масштаба времени к действиям робота, тем больше внимания должно уделяться такому прогнозированию. Способности робота осознанно и своевременно реагировать на внешние события позволяют судить об уровне его интеллектуальности. Особенности прогнозирования событий для интеллектуальных роботов состоят в том, что во многих случаях законы развития событий непрерывно меняются. Например, при рассмотрении коллективного поведения автономных интеллектуальных роботов в условиях наличия случайных факторов трудно предсказать возможные ситуации и заблаговременно предусмотреть реакции на них. Это относится и к отдельно взятым автономным роботам, действующим в быстро меняющейся обстановке. Несвоевременная, неадекватная реакция роботов может приводить не только к низкой их эффективности, но и катастрофическим последствиям.

В каждом конкретном случае внешние события могут развиваться по своим законам. Число возможных вариантов развития событий окружающего мира в общем случае стремится к бесконечности. Для того чтобы адекватно предсказывать события, нужны соответствующие модели, обойтись набором типовых моделей не всегда удается. Желательно,

чтобы адекватная модель развития событий оперативно формировалась самим роботом.

Одним из возможных подходов к решению этой задачи выступает нейросетевое прогнозирование. Известен ряд методов такого прогнозирования [2–6], которые в той или иной мере применимы к интеллектуальным роботам. Прежде всего, это методы прогнозирования временных рядов. Однако лежащие в их основе искусственные нейронные сети во многом несовершенны. Среди их недостатков — ориентация на одношаговое прогнозирование с малой глубиной, невысокие возможности по ассоциативному запоминанию связей между обрабатываемыми сигналами и извлечению их из памяти, существенные ограничения на структуру нейронной сети, невозможность использования одной и той же сети для решения других задач. Как правило, нейронную сеть обучают на решение конкретной задачи прогнозирования. После фиксации весов синапсов такие сети реализуют жесткие функции и не способны перестраиваться на другие задачи. Известен также метод прогнозирования событий с использованием рекуррентной нейронной сети (РНС) [7]. Этот метод позволяет гибко настраиваться на различные модели развития событий и прогнозировать их. Для реализации прогнозов в такой РНС предусматривается изменение задержек сигналов внутри сети в зависимости от текущих состояний ее слоев. За счет этого РНС может функционировать в прошедшем, настоящем и будущем времени. В некоторой мере этот метод лишен отмеченных выше недостатков, однако его применение связано с техническими трудностями, связанными с управлением задержками сигналов. Кроме

этого, не совершенна структура слоев РНС и не проработаны условия ее перехода от настоящего к будущему или прошедшему времени. Необходимо совершенствование известных и поиск новых методов нейросетевого прогнозирования событий.

В статье развиваются идеи нейросетевого прогнозирования событий для интеллектуальных роботов. Предлагается метод такого прогнозирования с расширенными функциональными возможностями.

Постановка задачи

Для пояснения предлагаемого метода рассмотрим постановку задачи на примере. Планируется создание интеллектуального автономного робота для выполнения широкого комплекса полезных для человека работ в труднодоступной и опасной среде. Этот робот по замыслу разработчиков должен обладать возможностями наблюдения за окружающими событиями с использованием широкого набора датчиков — оптических, акустических, тепловых, химических, давления, скорости и др. Робот должен обрабатывать воспринимаемую информацию, своевременно вырабатывать и реализовывать целесообразные решения, связанные с его деятельностью в окружающей динамически изменяющейся среде. Такими активными целесообразными действиями интеллектуального робота могут быть различные движения в пространстве, выполнение сложных операций над окружающими динамическими объектами, взаимодействие с другими роботами. Для достижения этого интеллектуальный робот должен обладать способностью прогнозировать возможные события, причем должна обеспечиваться возможность одновременного прогнозирования на различную глубину различных событий.

Необходимо разработать метод нейросетевого прогнозирования для таких интеллектуальных роботов, позволяющий успешно предсказывать возможные события, опираясь на накопленный опыт.

Метод нейросетевого прогнозирования

Согласно предлагаемому методу каждый воспринимаемый роботом сигнал, несущий информацию о событиях, предварительно должен быть разложен на пространственно-частотные составляющие. Каждую из этих составляющих необходимо преобразовать в последовательность единичных образов (ЕО) с возможностью обратного однозначного преобразования. Затем из полученных последовательностей ЕО путем вводимых временных задержек, равных глубине прогнозирования, формируют вторую совокупность последовательностей ЕО. В дискретном времени сформированные таким образом сигналы можно рассматривать как две последовательности совокупностей единичных образов (СЕО). Эти последовательности СЕО подаются на пространственно разнесенные входы двухслой-

ной рекуррентной нейронной сети (РНС) с управляемыми синапсами [8]. Особенности этой РНС в том, что при передаче СЕО от слоя к слою их сдвигают вдоль слоев. За счет этих сдвигов СЕО продвигают вдоль слоев по двойной спирали и скручивают текущую и задержанную последовательности этих совокупностей. В качестве результатов обработки сигналов используют последовательные СЕО на выходах нейронной сети после преобразования их в исходные сигналы. За счет такого продвижения СЕО в РНС расширяются ее возможности по ассоциативному запоминанию и извлечению сигналов из памяти. Для получения прогнозов осуществляют переключение входов РНС. В результате такого переключения текущую последовательность СЕО направляют на другой вход РНС вместо задержанной последовательности. После получения прогноза реализуют обратное переключение для продолжения накопления опыта. В случаях, когда требуется извлечь из памяти РНС предшествующие события, отключают подачу в сеть задержанной последовательности СЕО. Получив интересующий результат, возвращаются к обычному процессу.

Несомненным достоинством предлагаемого метода является возможность параллельного прогнозирования событий на различную глубину с постепенным улучшением результатов. Это позволяет получить скользкие многошаговые прогнозы, на основе которых можно заранее предусматривать реакции интеллектуальных роботов. Результаты скользкого многошагового прогноза на каждый текущий i -й момент времени t_i представляются последовательностью ожидаемых событий для моментов $t_i + 1, t_i + 2, t_i + 3, \dots, t_i + n$ в виде соответствующих последовательностей совокупностей (S) единичных образов:

$$\{S_1(t_i), S_2(t_i), \dots, S_3(t_i), \dots, S_n(t_i)\}$$

$$\{S_1(t_i + 1), S_2(t_i + 1), S_3(t_i + 1), \dots, S_n(t_i + 1)\}$$

$$\{S_1(t_i + 2), S_2(t_i + 2), S_3(t_i + 2), \dots, S_n(t_i + 2)\}$$

$$\dots$$

$$\{S_1(t_i + m), S_2(t_i + m), S_3(t_i + m), \dots, S_n(t_i + m)\},$$

где n — максимальная глубина прогнозирования (число временных тактов, на которые осуществляется прогноз); m — число тактов, в течение которых осуществляется прогнозирование событий.

Получение интеллектуальным роботом информации о будущих событиях в той же самой форме, что и о текущих событиях, позволяет ему адекватно реагировать на них.

Структура РНС

В качестве РНС, реализующей предлагаемый метод, может выступать сеть со структурой, представленной на рис. 1.

Результаты прогнозирования

Для подтверждения справедливости сформулированных положений предлагаемого метода нейросетевого прогнозирования событий для интеллектуальных роботов было проведено математическое моделирование. Была разработана небольшая рекуррентная нейронная сеть (РНС) с управляемыми синапсами, со структурой слоев, подобной рис. 1. Каждый слой сети состоял из 1344 нейронов и за счет пространственных сдвигов СЕО разбивался на две строки по 16 полей размером 6×7 нейронов. В качестве базовых элементов такой РНС выступали нейроны с тремя состояниями: ожидания, возбуждения и временной невосприимчивости (рефрактерности) после возбуждения. Время задержки сигналов в образуемых двухслойных контурах РНС задавалось меньшим, чем время рефрактерности нейронов. В этой РНС при протекании токов через синапсы в прямом направлении увеличивались их веса (проводимости). Посредством прохождения токов через синапсы в обратном направлении уменьшались их прямые проводимости и осуществлялось частичное стирание устаревшей информации, аналогично тому, как это делалось в работе [9]. За счет приоритетности коротких связей между нейронами между входами и выходами РНС по входным и выходным языкам обеспечивалось однозначное соответствие. В качестве сигналов в РНС подавались предварительно сформированные текущая и задержанная последовательности СЕО (импульсов).

Результаты вычислительных экспериментов подтвердили возможность успешного прогнозирования событий для интеллектуальных роботов. Сеть успешно запоминала на синапсах связи между текущими и задержанными ЕО и позволяла извлекать из ее памяти результаты прогноза, причем результаты прогноза проявлялись на нескольких полях слоев сети и при продвижении вдоль них к выходу улучшались. Однако чем ближе они к выходу, тем меньше глубина прогнозирования из-за задержек в самой сети.

Установлено, что при накоплении сетью опыта желательно иметь повышенный порог возбуждения нейронов, но не превышающий возможности активизации нейронов, по приоритетным коротким связям. При непосредственном прогнозировании событий целесообразно снижать этот порог, облегчать вызов из памяти сигналов входной последовательностью СЕО.

Простой пример такого многошагового прогнозирования приведен на рис. 2, а, б. Согласно рис. 2, а, б входные СЕО, продвигающиеся вдоль слоев РНС справа налево по второй строке, вызывают из памяти будущие события, отражаемые на первой строке. При этом результаты прогнозирования в виде последовательных СЕО, несущих информацию о пространственном смещении простого объекта в пределах полей $6 \times 7 = 42$ нейронов, про-

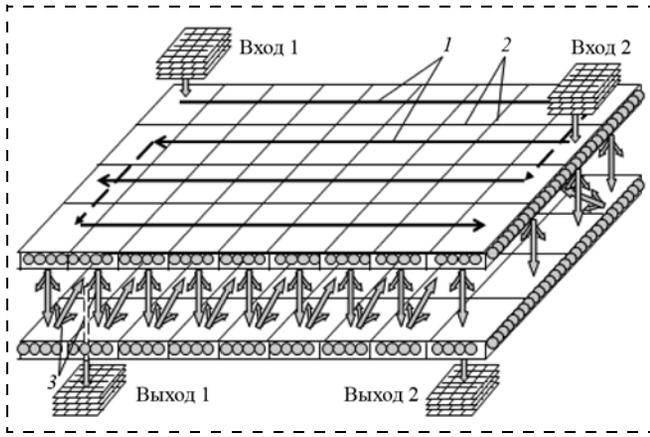


Рис. 1. Рекуррентная нейронная сеть со структурой слоев в виде двойной спирали:

1, 3 — направления продвижения совокупностей единичных образов вдоль слоев и между ними; 2 — линии разбивки слоев на логические поля

Согласно рис. 1 подаваемые в сеть СЕО продвигаются вдоль слоев в заданных направлениях. При таком продвижении единичные образы из этих совокупностей ассоциируются друг с другом. Результаты ассоциативного взаимодействия ЕО запоминаются на синапсах сети. При накоплении опыта сетью на вход 1 подаются не задержанные, а на вход 2 — задержанные СЕО, несущие информацию о наблюдаемых сигналах. При прогнозировании событий вход 1 остается свободным, а не задержанные СЕО подаются на вход 2 вместо задержанных. В результате текущие СЕО, выступающие в качестве задержанных, вызывают ассоциативно из памяти сети связанные с ними будущие совокупности таких образов. Они могут сниматься как с выхода 1, так и с предшествующих ему полей, на которые разбиты слои сети. Другой случай, когда есть некоторая текущая последовательность СЕО, подаваемая на вход 1, но неизвестна предшествующая ей последовательность. Тогда в результате обработки сигналов при свободном входе 2 на выходе 2 получаем интересующие предшествующие результаты.

Заметим, что результаты такого нейросетевого прогнозирования зависят от объема накопленного опыта. Если не подкреплять входы сети результатами обработки сигналов, то любое ассоциативное извлечение сигналов из ее памяти влечет за собой ослабление этого опыта. Необходимо соблюдать баланс между временем накопления опыта (обучения) и временем прогнозирования событий. Время прогнозирования не должно превышать времени накопления опыта. Чтобы получать результаты прогноза на выходе 1, время задержки СЕО перед подачей в сеть должно быть больше задержки сигналов в самой сети. Для получения результатов прогноза на большую глубину достаточно наблюдать за состояниями соответствующих внутренних полей слоев сети.

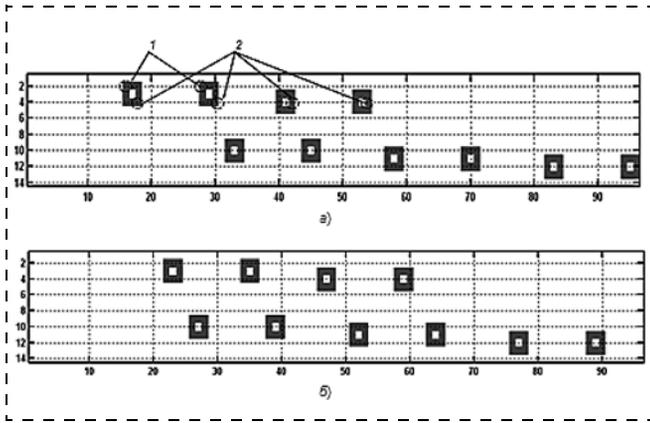


Рис. 2. Примеры состояния первого слоя РНС при прогнозировании событий:

a — результат прогнозирования перед очередными пространственными сдвигами СЕО вдоль слоев (*1* — возбужденные нейроны, привязанные к одной и той же пространственно-частотной составляющей, но к различным временным тактам; *2* — аналогично, но для другой пространственно-частотной составляющей); *б* — после таких сдвигов

двигаются вдоль слоев слева направо. Несомненно, в практическом плане при прогнозировании сложных пространственно-соотнесенных цветных событий (имеется в виду, что в качестве событий могут выступать цветные динамические образы. Результаты прогноза могут отображаться цветными кадрами) размеры таких полей должны исчисляться не в десятках, а в миллионах нейронов.

Если в соответствии с рис. 2, *a* рассматривать результаты прогноза в виде отдельных временных рядов, соотнесенных с каждой пространственно-частотной составляющей, то они могут быть представлены в виде рис. 3.

Согласно рис. 3 верхний временной ряд соответствует цепочке *1* на рис. 2, а средний ряд — цепочке *2* на рис. 2. Всего таких ненулевых рядов в конкретном случае 12. Максимально возможное число временных рядов для этого примера равно 42 ($42 = 6 \times 7$). Заметим, что на рис. 2 результаты прогноза на первой строке сдвигаются при очеред-

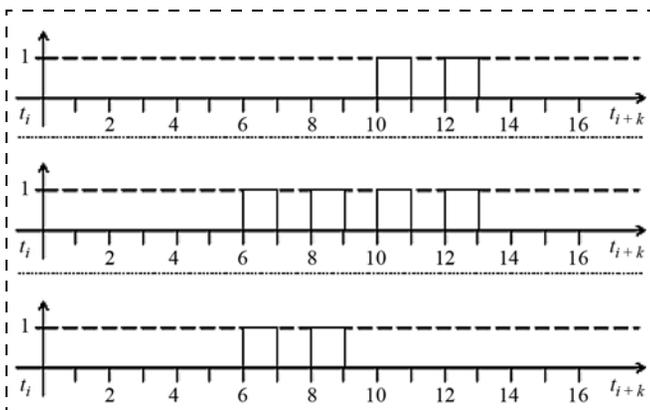


Рис. 3. Временные ряды, соответствующие результатам нейросетевого прогноза

ном шаге слева направо, а на рис. 3 — справа налево.

В отличие от известных решений согласно предлагаемому методу текущие результаты прогноза учитываются при получении следующих прогнозов.

Для практического использования в интеллектуальных роботах такого прогнозирования при больших размерах РНС желательна их не программная, а аппаратная реализация. Программная реализация предлагаемого метода успешно применяется при небольших РНС.

Заключение

Рассмотренный метод нейросетевого прогнозирования событий для интеллектуальных роботов ориентирован на широкий круг возможных условий. Согласно этому методу не требуется заранее знать, по какому закону будут развиваться наблюдаемые события. Модель развития событий динамически формируется самой сетью в процессе накопления ею опыта. Одна и та же РНС может применяться для прогнозирования различных событий. Предлагаемый метод позволяет получать на каждый момент времени прогноз не на один, а на несколько шагов вперед, действовать на один шаг и снова прогнозировать события на несколько шагов вперед. При этом за счет встречного продвижения обрабатываемых СЕО вдоль слоев сети улучшаются ее возможности по запоминанию и извлечению из памяти результатов прогнозирования событий. Отдаленные результаты прогноза динамически по мере приближения к реальному времени улучшаются. Формально за выбор режима работы РНС (накопление опыта, прогнозирование, восстановление прошедших событий) может отвечать как внешний потребитель, так и сама сеть. Результаты прогнозирования могут представляться в том же самом виде, что и исходные сигналы. Это позволяет интеллектуальному роботу видеть мир в будущих событиях и своевременно реагировать на них. С применением этого метода в общем случае осуществимо одновременное прогнозирование различных характеристик, наблюдаемых объектов. В качестве таких характеристик могут выступать параметры движения, изменения формы, цвета, яркости и др.). С использованием этого метода могут прогнозироваться не только внешние события, но и действия самого интеллектуального робота, подлежащие реализации.

Предложенный метод нейросетевого прогнозирования рекомендуется применять в перспективных автономных интеллектуальных роботах, функционирующих в быстро меняющихся внешних условиях, при жестких требованиях реального времени на возможные реакции. Рекуррентные нейронные сети для реализации этого метода могут быть исполнены аппаратно с применением ждущих мультивибраторов в качестве нейронов и мемристоров (резисторов с памятью) как искусственных синапсов.

Список литературы

1. Добрынин Д. А., Карпов В. Э. Моделирование некоторых форм адаптивного поведения интеллектуальных роботов // Информационные технологии и вычислительные системы. 2006. № 2. С. 45–56.
2. Рабочая книга по прогнозированию. М.: Мысль, 1982. 430 с.
3. Haykin S. Neural Networks and Learning Machines. New-York: Prentice Hall, 2008. 936 p.
4. Galushkin A. I. Neural Networks Theory. Springer-Verlag Berlin Heidelberg, 2007. 396 p.
5. Осовский С. Нейронные сети для обработки информации: Пер. с польского И. Д. Рудницкого. М.: Финансы и статистика, 2002. 344 с.
6. Ghiassi M., Saidane H., Zimbra D. K. A dynamic artificial neural network model for forecasting time series events // International Journal of Forecasting. 2005. Vol. 21, Iss. 2. P. 341–362.
7. Осипов В. Ю. Нейронная сеть с прошедшим, настоящим и будущим временем // Информационно-управляющие системы. 2011. № 4. С. 30–33.
8. Осипов В. Ю. Рекуррентная нейронная сеть со структурой слоев в виде двойной спирали // Информационные технологии. 2014. № 7. С. 56–60.
9. Осипов В. Ю. Стирание устаревшей информации в ассоциативных интеллектуальных системах // Мехатроника, автоматизация, управление. 2012. № 3. С. 16–20.

Neural Network Prediction of Events for Intelligent Robots

V. Yu. Osipov, osipov_vasily@mail.ru✉, St. Petersburg Institute for Informatics and Automation of RAS, St. Petersburg, 199178, Russian Federation

Corresponding author: Osipov Vasily Yu., D. Sc., Professor, Leading Researcher, St. Petersburg Institute for Informatics and Automation of RAS, St. Petersburg, Russian Federation, e-mail: osipov_vasily@mail.ru

Received on March 02, 2015

Accepted on August 25, 2015

The traditional forecasting of events is based on a preliminary design of their models. Employment of such models as a part of the intelligent robots cannot take into account the whole variety of the situations, which may arise. It is desirable that an adequate model of events is formed by a robot itself. In the interests of intellectualization of the autonomous robots the task of predicting events without setting models of their development is considered. Objective is to increase the functionalities of forecasting of the events with the changing laws of their development. In order to address this problem a new method for predicting is proposed. The method involves the use of a recurrent neural network with controlled synapses and with a layer structure in the form of a double spiral. The method is based on associative memorizing of the current and delayed signals, and extraction of the future events from the memory of the neural network. According to the method is not necessary to know in advance, in accordance with which law the observed events will develop. The development model of such events is formed by the dynamic neural network itself in the process of accumulation of its experience. The method makes it possible to predict the parallel events at various depths with a gradual improvement of the results. These results are the sliding multistep forecasts. Due to this the intelligent robots can qualitatively predict events and plan their responses. The simulation results, reflecting the peculiarities of such prediction, are presented. The recommendations for the use of the proposed method have been formulated.

Keywords: prediction, neural network, signals, events, autonomous robot

For citation:

Osipov V. Yu. Neural Network Prediction of Events for Intelligent Robots, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2015, vol. 16, no. 12, pp. 836–840.

DOI: 10.17587/mau/16.836-840

References

1. Dobrynin D. A., Karpov V. Je. *Modelirovanie nekotoryh form adaptivnogo povedeniya intellektual'nyh robotov* (Modeling of some forms of adaptive behavior of intelligent robots), *Informacionnye Tehnologii i Vychislitel'nye Sistemy*, 2006, no. 2, pp. 45–56 (in Russian).
2. Bestuzhev-Lada I. V. ed. *Rabochaja kniga po prognozirovaniju* [Workbook Forecasting], Moscow, Mysl', 1982, 430 p. (in Russian).
3. Haykin S. Neural Networks and Learning Machines, Prentice Hall, 2008, 936 p.
4. Galushkin A. I. Neural Networks Theory, Springer-Verlag Berlin Heidelberg, 2007, 396 p.
5. Osovskii S. *Neironnye seti dlya obrabotki informatsii* (Neural networks for processing of information), Moscow, Finansy i Statistika, 2004. 344 p. (in Russian).
6. Ghiassi M., Saidane H., Zimbra D. K. A dynamic artificial neural network model for forecasting time series events, *International Journal of Forecasting*, 2005, vol. 21, iss. 2, pp. 341–362.
7. Osipov V. Yu. *Nejronnaja set' s proshedshim, nastojashhim i budushhim vremenem* (The Neural Network with Past, Present and Future Time), *Informacionno-Upravljajushhie Sistemy*, 2011, no. 4, pp. 30–33 (in Russian).
8. Osipov V. Yu. *Rekurrentnaja nejronnaja set' so strukturoj sloev v vide dvojnoj spirali* (Recurrent Neural Network with Structure of Layers in the Form of the Double Spiral), *Informacionnye Tehnologii*, 2014, no. 7, pp. 56–60 (in Russian).
9. Osipov V. Yu. *Stiranje ustarevshei informatsii v assotsiativnykh intellektual'nykh sistemakh* (Erase Outdated Information in Associative Intelligent Systems), *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2012, no. 3, P. 16–20 (in Russian).