

УДК: 517.93

С.Н. Верзунов, Н.М. Лыченко

Институт автоматизации и информационных технологий

Национальной академии наук Кыргызской Республики, г.Бишкек

verzunov@hotmail.com, nlychenko@mail.ru

ПОЛИМОРФНАЯ ВЕЙВЛЕТ-СЕТЬ С ОБРАТНЫМИ СВЯЗЯМИ ДЛЯ ИДЕНТИФИКАЦИИ НЕСТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ

Рассматривается структура и алгоритм обучения полиморфной вейвлет-сети с обратными связями. Полученные результаты применяются для решения задачи идентификации нестационарных временных рядов.

Ключевые слова: идентификация, нестационарный временной ряд, вейвлет-сеть.

Введение. В классическом смысле под идентификацией понимается выяснение структуры и параметров математической модели, дающей наилучшее соответствие выходных сигналов модели и временного ряда. При выбранной структуре модели задача идентификации представляет собой типичную экстремальную задачу и для ее решения с успехом используется мощный аппарат теории задач такого рода [1]. В последние три десятилетия активно развивались методы адаптивной идентификации с динамической подстройкой оценок неизвестных параметров линейных и нелинейных систем. Появились работы, в которых для идентификации нелинейных динамических систем используются рекуррентные нейронные сети [2]. С точки зрения теории систем, такие сети представляют собой нелинейные динамические системы с обратными связями. В этом случае задача идентификации становится неклассической, поскольку математическая модель объекта в явном виде не определяется и сводится к определению параметров нейронной сети с помощью адаптивного алгоритма, использующего пары наблюдаемых данных вход-выход.

Постановка задачи. Пусть имеется динамическая система, продуцирующая временной ряд, при этом исходные данные о параметрах и структуре модели системы отсутствуют. Требуется по наблюдаемым (измеряемым) входным и выходным переменным системы построить идентификационную модель системы при наличии разного рода ошибок измерений. В общем виде задача идентификации динамической системы поясняется на рис. 1. На систему влияет входное воздействие $u(t_i)$ и ненаблюдаемое возмущение $\varepsilon_s(t_i)$. Вход системы $u(t_i)$ и выход системы $y(t_i)$ измеряются в дискретные моменты времени при наличии некоторых случайных факторов, характеризующихся ошибками измерения, соответственно $\varepsilon_u(t_i)$ и $\varepsilon_y(t_i)$, при этом $M\{\varepsilon\}=0$, $D\{\varepsilon\}<\infty$.

Необходимо, используя выборки зашумленных измерений (наблюдений) реакции системы $d(t_i) = y(t_i) + \varepsilon_y(t_i)$ на входное воздействие $\hat{u}(t_i) = u(t_i) + \varepsilon_u(t_i)$, построить *Идентификатор системы* (идентификационную модель), который при любом входе, подаваемом на объект, продуцирует выход модели, аппроксимирующий временной ряд в смысле малости ошибки идентификации $e(t_i)$.

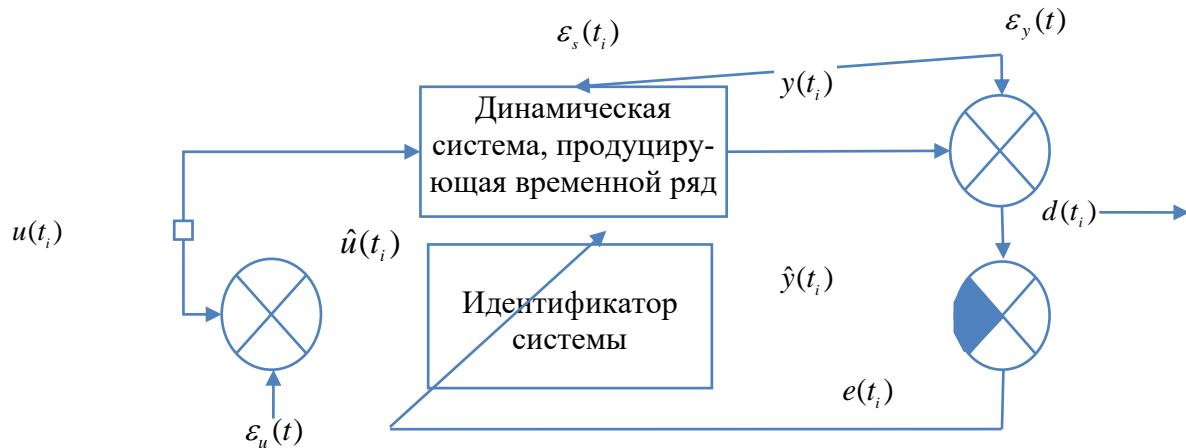


Рисунок 1. Структурная схема идентификации временного ряда

Полиморфная вейвлет-сеть с обратными связями. Для аппроксимации временного ряда предлагается использовать полиморфную вейвлет-сеть с обратными связями (рис. 2), выход которой определяется уравнением:

$$\hat{y}(t_i) = c + t_i \sum_{k=1}^K w_k h_{a_k, b_k, p_k}(t_i) + \sum_{j=1}^R r_j \hat{y}(t_i - j),$$

где c – постоянная составляющая (параметр смещения); K – количество вейвлетов; w_k – весовые коэффициенты прямых связей; R – количество обратных связей; r_j – весовые коэффициенты обратных связей.

Из рис. 2 видно, что предложенная сеть состоит из двух частей: полиморфной вейвлет-сети [3] и блока с обратными связями.

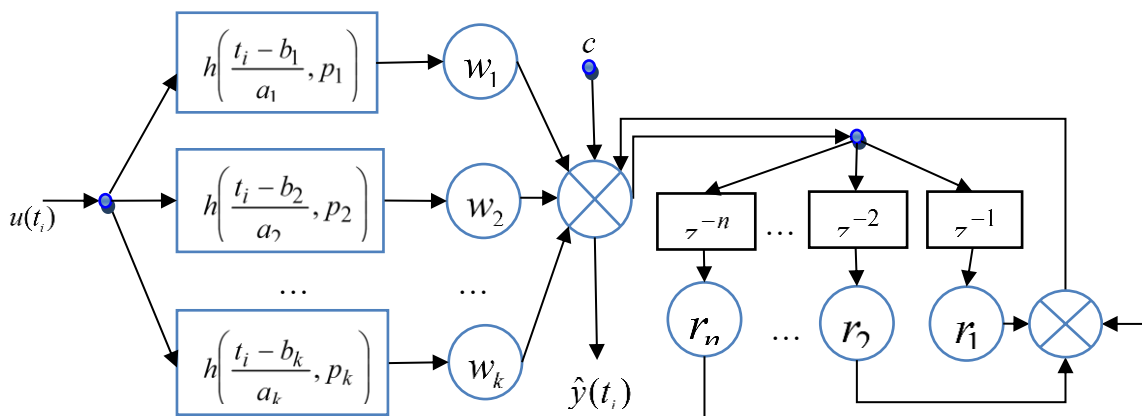


Рисунок 2. Полиморфная вейвлет-сеть с обратными связями

Таким образом, обозначая ошибку в момент t_i как: $e(t_i) = d(t_i) - \hat{y}(t_i)$, функцию

$$E = \frac{1}{2} \sum_{i=1}^N e^2(t_i)$$

энергии ошибки можно записать:

Параметры вейвлет-сети w_k , a_k , b_k , p_k , r_j и c могут быть настроены путем минимизации функции энергии ошибки E на всем множестве отсчетов времени t_j :

$$\sum_{i=1}^N \left(d(t_i) - \left[c + t_i \sum_{k=1}^K w_k h_{a_k, b_k, p_k}(t_i) + \sum_{j=1}^R r_j \hat{y}(t_{j-i}) \right] \right)^2 \rightarrow \min_{c, \bar{w}, \bar{a}, \bar{b}, \bar{p}, \bar{r}}$$

Для минимизации E можно использовать различные численные методы оптимизации, требующие градиенты $\frac{\partial E}{\partial w_k}$, $\frac{\partial E}{\partial a_k}$, $\frac{\partial E}{\partial b_k}$, $\frac{\partial E}{\partial p_k}$, $\frac{\partial E}{\partial r_i}$ и $\frac{\partial E}{\partial c}$ для последовательного измене-

ния каждого конкретного параметра w_k , a_k , b_k , p_k , r_i и c :

$$\begin{aligned} \frac{\partial E}{\partial w_k} &= -\sum_{i=1}^N e(t_i) h(\tau_i, p_k) \hat{u}(t_i), & \frac{\partial E}{\partial b_k} &= -\sum_{i=1}^N e(t_i) \hat{u}(t_i) w_k \frac{\partial h(\tau_i, p_k)}{\partial b_k}, \\ \frac{\partial E}{\partial a_k} &= -\sum_{i=1}^N e(t_i) \hat{u}(t_i) w_k \tau \frac{\partial h(\tau_i, p_k)}{\partial b_k} = \tau_i \frac{\partial E}{\partial b_k}, & \frac{\partial E}{\partial p_k} &= -\sum_{i=1}^N e(t_i) \hat{u}(t_i) w_k \frac{\partial h(\tau_i, p_k)}{\partial p_k}, \\ & & \frac{\partial E}{\partial r_i} &= -\sum_{j=1}^R e(t_i) \hat{y}(t_j - j), & \frac{\partial E}{\partial c} &= -\sum_{i=1}^N e(t_i). \end{aligned}$$

В этих выражениях $h(\tau_i)$ и ее производные определяются конкретным типом вейвлета. Например, для вейвлета RASP, используемого в описанном ниже вычислительном эксперименте, они определяются как:

$$h(\tau_i) = \frac{p_k \tau_i}{(\tau_i^2 + 1)^2}, \quad \frac{\partial h(\tau_i, p_k)}{\partial p_k} = \frac{\tau_i}{(\tau_i^2 + 1)^2}, \quad \frac{\partial h(\tau_i, p_k)}{\partial b_k} = \frac{p_k (3\tau_i^2 - 1)}{a_k (\tau_i^2 + 1)^3}, \quad \text{где } \tau_i = \frac{t_i - b_k}{a_k}.$$

Аппроксимация временных рядов полиморфными вейвлет-сетями с обратными связями, содержащими базовые вейвлеты различных типов. Рассмотрена задача аппроксимации (см. рис. 2), в которой $u(t_i) = t_i$, временной ряд $y(t_i)$ моделируются численным решением обыкновенного дифференциального уравнения на интервале $[t_0, N\Delta t]$ с шагом Δt . Требуется аппроксимировать $d(t_i)$.

В качестве аппроксиматора предлагается использовать полиморфную вейвлет-сеть с обратными связями (см. рис. 2). Был проведен вычислительный эксперимент, в котором рассматривались два временных ряда, моделируемых дифференциальными уравнениями:

$$\frac{d^2 y}{dt^2} + 5 \frac{dy}{dt} + (6 + 3 \sin(0.5t)) y = \varepsilon_s(t), \quad \frac{d^2 y}{dt^2} + (2 + \sin(0.5t)) \frac{dy}{dt} + 5y = \varepsilon_s(t)$$

для 30 различных реализаций ошибок измерений ε . При этом $\varepsilon_s(t_i) = 0.1 \sin(7t) + \cos(9t)$, $\varepsilon(t) = \xi |y(t)|$, где ξ – некоторая нормально распределенная случайная величина $M\{\xi\} = 0$, $D\{\xi\} = 0.5$.

Для сравнения в качестве аппроксиматора была также применена полиморфная вейвлет-сеть без обратных связей. Кроме того, в качестве базовых использовались вейвлеты различных типов: Морле, POLYWOG и RASP [3]. Результаты статистически обработаны и приведены на рисунке 3.

Как видно из рис.3, сравнение возможностей полиморфной вейвлет-сети с обратными

связями и без них в плане использования их в качестве аппроксиматоров нестационарных временных рядов с инерционностью показало, что полиморфная вейвлет-сеть с обратными связями дает меньшую энергию ошибки E в сравнении с полиморфной вейвлет-сетью без обратных связей и, кроме того, имеет преимущества в скорости сходимости (характеризующейся скоростью сходимости \hat{n} и временем, затрачиваемым на обучение T) [4].

Идентификация нелинейной динамической системы полиморфной вейвлет-сетью с обратными связями.

Для решения задачи идентификации параметры вейвлет-сети нужно настроить так, чтобы минимизировать меру рассогласования между значениями выходных переменных динамической системы, генерирующей временной ряд, и сети. Структура сети должна быть идентичной структуре идентифицируемого динамического объекта, то есть она должна включать обратные связи и временные задержки. Итак, рассмотрим структуру полиморфной вейвлет-сети с обратными связями, на вход которой подается зашумленное измерение (наблюдение) входного воздействия $\hat{u}(t_i)$, поступающего на динамическую систему.

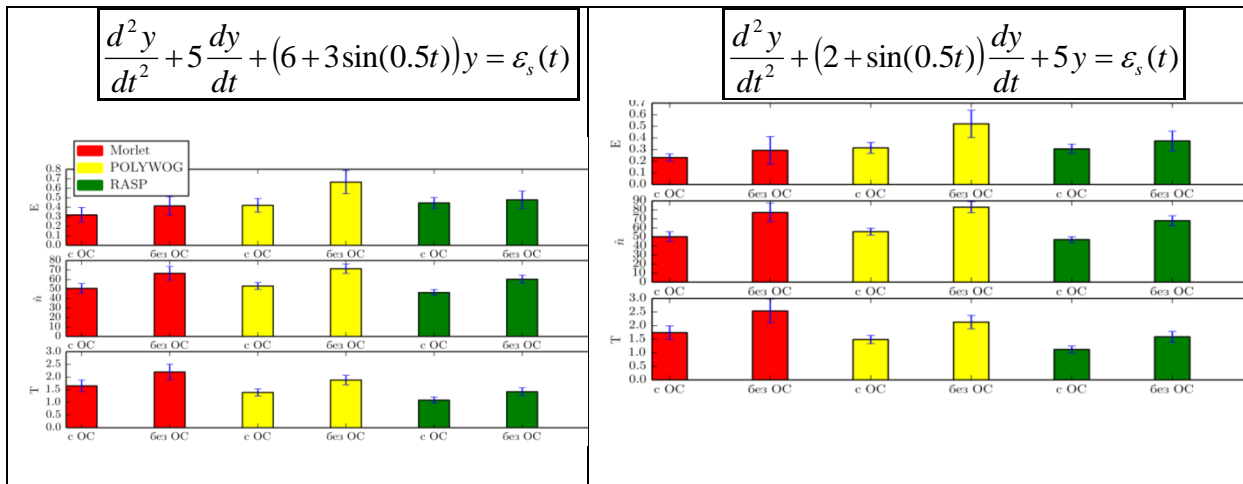


Рисунок 3. Результаты численных исследований аппроксимации нестационарной динамической системы полиморфными вейвлет-сетями с обратными связями и без обратных связей с различными базовыми вейвлетами

Рассмотрим для примера нелинейную дискретную динамическую систему 2-го порядка:

$$y_j(n) = \frac{1.1y(n-1)}{1 + y(n-2)} + u_j(n) + \varepsilon_s(n) \quad (1)$$

где $j=1,2,3,4$; $n=0,1,\dots,99$. Случайные ошибки измерения имеют нормальное распределение, при этом стандартные отклонения для ошибок измерения $\varepsilon_u(t_i)$ и $\varepsilon_y(t_i)$ равны соответственно 0.01 и 0.05 . Стандартное отклонение неизмеряемого возмущения $\varepsilon_s(t_i) = 0.01$. Обучающую выборку образуют пары (u_1, y_1) и (u_2, y_2) , при этом $u_1(n) = 1(n)$, $u_2(n) = 0.5\cos(3\gamma_2)e^{-0.5\gamma_2^2} + 1$, $(\gamma_2 = \frac{n-20}{3})$. $u_3(n) = 0.7\cos(2.5\gamma_3)e^{-0.5\gamma_3^2} + 1$,

при $\gamma_3 = \frac{n-30}{3.5}$, $u_4(n) = 0.5\delta(n-25) + 1$, а тестовую – (u_3, y_3) и (u_4, y_4) .

На рис. 4 показана динамика изменения энергии ошибки E для сети, содержащей 40 RASP вейвлетов, при обучении ее по алгоритму Бройдена-Флетчера-Гольдфарба-Шанно на выборке $\{(u_1, y_1), (u_2, y_2)\}$ в течение 5000 эпох.

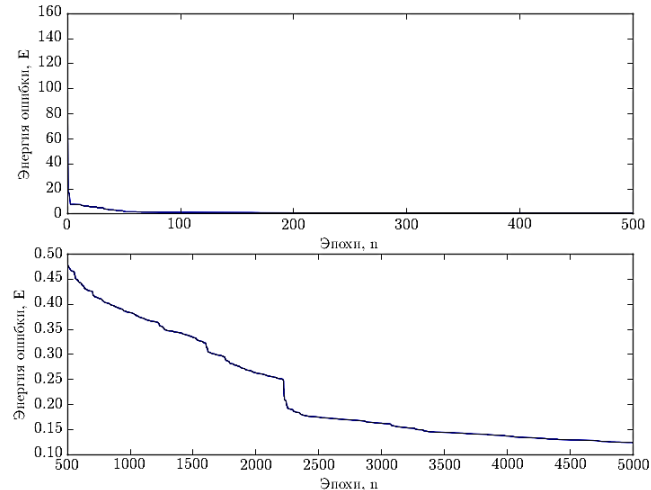


Рисунок 4. Динамика изменения энергии ошибки E в процессе обучения сети

Можно видеть, что в основном обучение сети заканчивается менее чем за 100 эпох, затем, как видно из рис. 5, на протяжении примерно 2200 эпох значения весов сети \bar{a} , \bar{b} , \bar{p} , \bar{w} , c , \bar{r} изменяются довольно плавно. Однако, около 2300 эпох происходят скачкообразные изменения значений параметров сети, которые приводят к значительному уменьшению энергии ошибки. Далее сеть медленно обучается на протяжении более 2500 эпох, после чего параметры сети вновь принимают устойчивые значения.

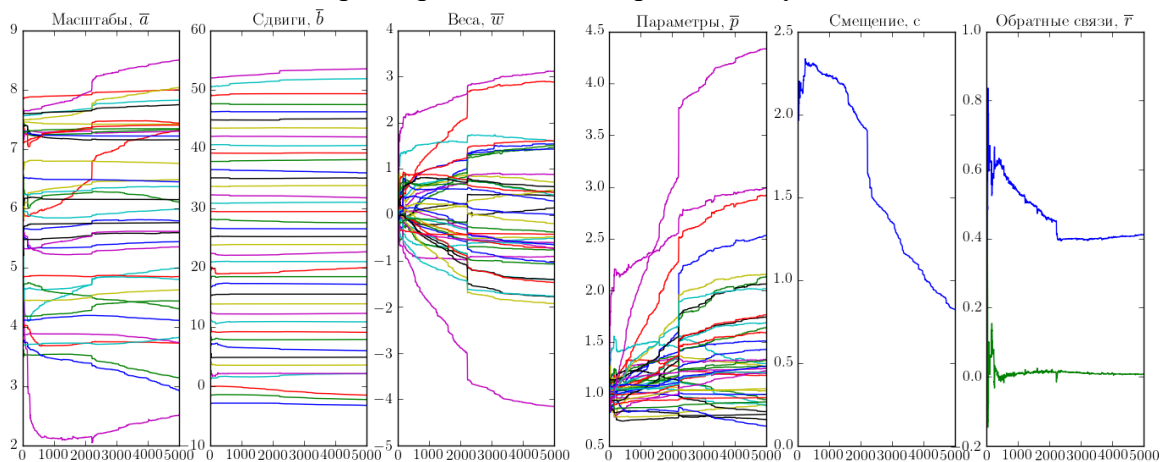


Рисунок 5. Динамика параметров сети в процессе обучения

На рис. 6 показаны выходы системы (1) и идентифицирующей ее вейвлет-сети при входном воздействии $u_1(n) = 1(n)$, воздействии $u_2(n) = 0.5\cos(3\gamma_2)e^{-0.5\gamma_2^2} + 1$, где $\gamma_2 = \frac{n-20}{3}$, использованных при обучении сети. На рис. 7 показаны выходы системы

(1) и вейвлет-сети при входных воздействиях, отличных от использованных при обуче-

нии: $u_3(n) = 0.7 \cos(2.5\gamma_3) e^{-0.5\gamma_3^2} + 1$, где $\gamma_3 = \frac{n-30}{3.5}$ и $u_4(n) = 0.5\delta(n-25) + 1$, где

$$\delta(n-25) = \begin{cases} 1, & \text{при } n = 25, \\ 0, & \text{в противном случае} \end{cases}$$

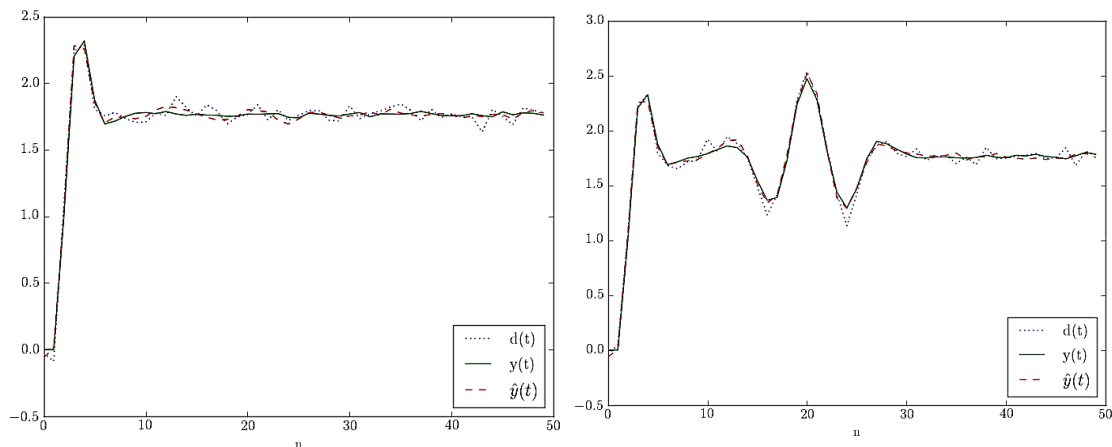


Рисунок 6. Результат идентификации системы полиморфной вейвлет-сетью при входном воздействии u_1 и u_2 из обучающей выборки

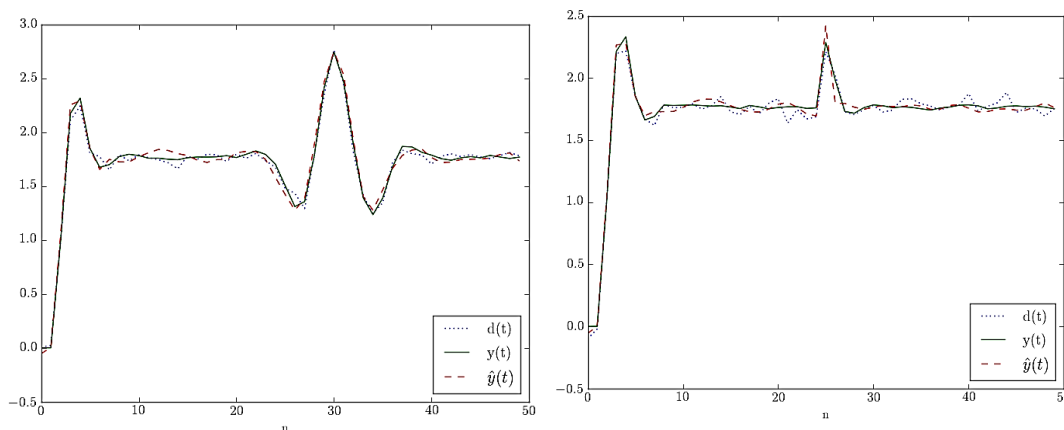


Рисунок 7. Результат идентификации системы полиморфной вейвлет-сетью при входном воздействии u_3 и u_4 из тестовой выборки

Из рис. 6–7 видно, что полиморфная вейвлет-сеть с обратными связями достаточно хорошо идентифицирует нелинейные динамические системы при различных входных воздействиях. На основании модели динамической системы в виде обученной полиморфной вейвлет-сети с обратными связями можно осуществлять прогноз детерминированной составляющей временного ряда, генерируемого этой системой при различных входных воздействиях. Следует отметить, что обучать сеть следует на двух и более парах вход/выход, в противном случае результат идентификации неудовлетворительный [5].

Аппроксимация нестационарных временных рядов полиморфными вейвлет-сетями с обратными связями, обученными различными численными методами. Для того, чтобы сравнить различные численные методы оптимизации, используемые при обучении полиморфных вейвлет-сетей, рассмотрим пример аппроксимации поли-

морфной вейвлет-сетью временного ряда, детерминированная составляющая которого $y(t_i)$ моделируется численным решением обыкновенного дифференциального уравнения на интервале $[t_0, N\Delta t]$ с шагом Δt :

$$\frac{d^2 y}{dt^2} + 5 \frac{dy}{dt} + (6 + 20 \sin(3t))y = u(t) + \varepsilon_s(t) \quad (2)$$

где $\varepsilon_s(t_i) = 0.1 \sin(7t) + \cos(9t)$, $u(t) = 1 + 0.2 \sin(2t)$. □ На рис. 8 показана аппроксимация системы (2) полиморфной вейвлет-сетью с 10-ю вейвлетами Морле с суммарной квадратической ошибкой $E=0.025$. Как видно из рисунка, эта система достаточно нестационарна. Обучение полиморфной вейвлет-сети с 10-ю вейвлетами Морле по алгоритму наискорейшего спуска до суммарной квадратической ошибки $E=0.025$ занимает 1440 эпох (рис. 9). На рис. 10 показана динамика изменения параметров сети в процессе обучения. При тех же условиях обучение сети по алгоритму сопряженных градиентов занимает 598 эпох (рис. 11). Обращает на себя внимание больший диапазон изменения дополнительных параметров p по сравнению с предыдущим алгоритмом (рис. 13).

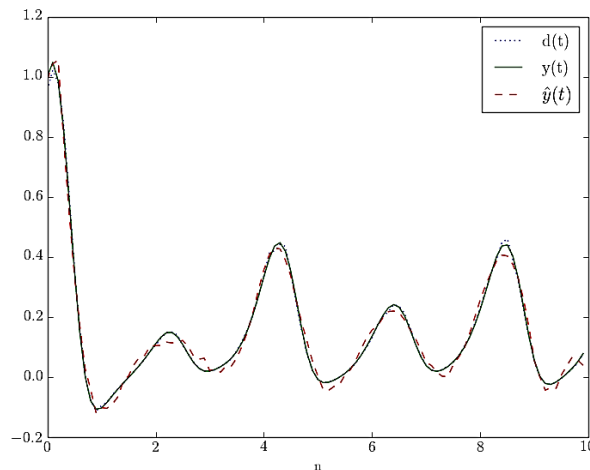


Рисунок 8. Аппроксимация нестационарной системы полиморфной вейвлет-сетью с 10-ю вейвлетами Морле по алгоритму наискорейшего спуска

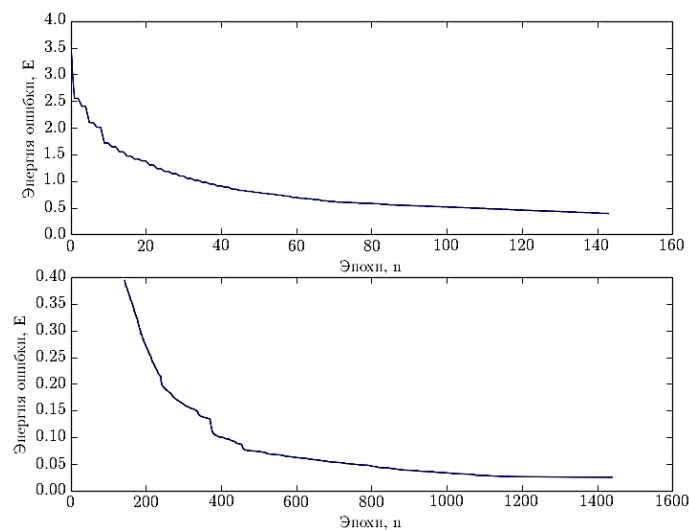


Рисунок 9. Динамика изменения энергии ошибки E в процессе обучения сети по алгоритму наискорейшего спуска

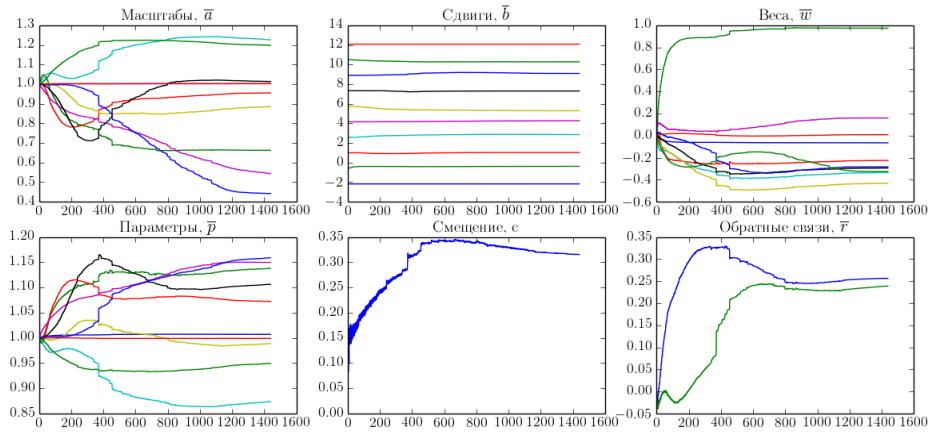


Рисунок 10. Динамика изменения параметров сети в процессе обучения по алгоритму наискорейшего спуска

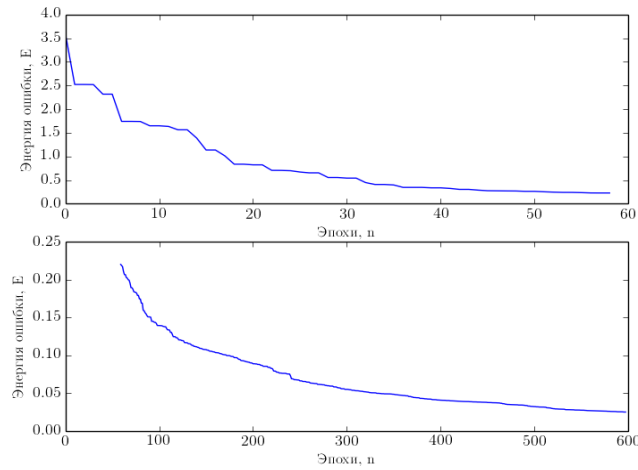


Рисунок 11. Динамика изменения энергии ошибки E в процессе обучения сети по алгоритму сопряженных градиентов

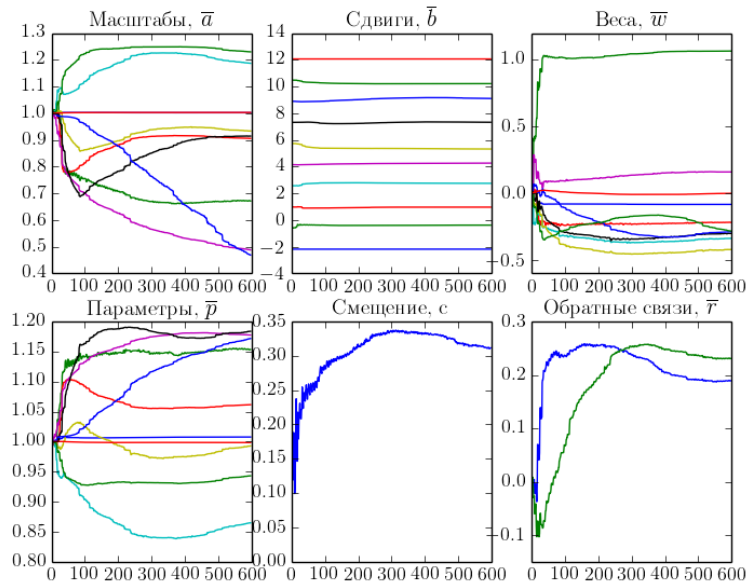


Рисунок 12. Динамика изменения параметров сети в процессе обучения по алгоритму сопряженных градиентов

Обучение сети по алгоритму Бройдена-Флетчера-Гольдфарба-Шанно занимает 293 эпохи (рис. 12).

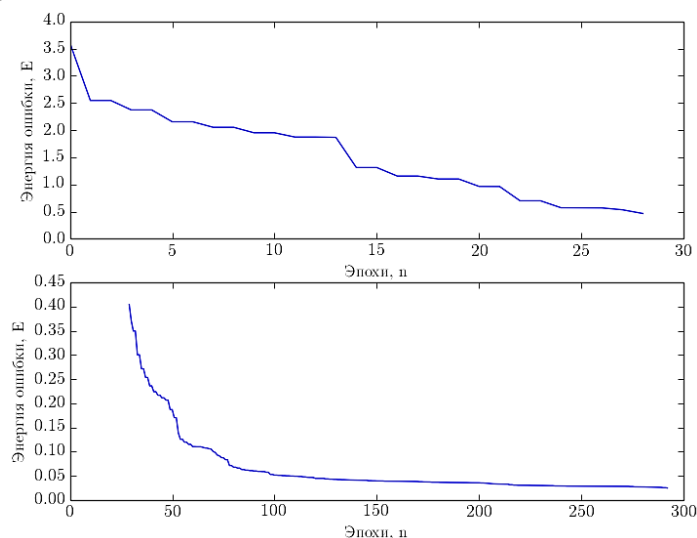


Рисунок 13. Динамика изменения энергии ошибки E в процессе обучения сети по алгоритму Бройдена-Флетчера-Гольдфарба-Шанно

При использовании этого алгоритма обучения веса сети изменяются в более широких диапазонах, причем сильнее всего изменяются именно масштабы вейвлетов a и дополнительные параметры p (рис. 14).

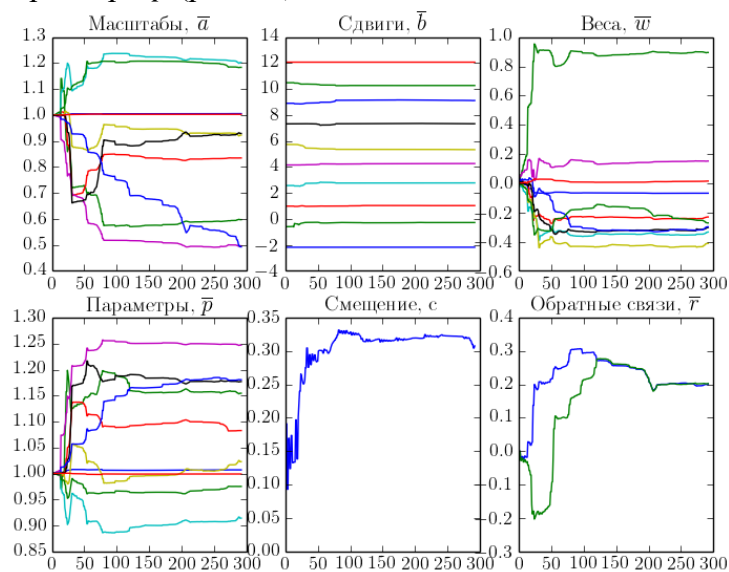


Рисунок 14. Динамика изменения параметров сети в процессе обучения по алгоритму Бройдена-Флетчера-Гольдфарба-Шанно

Таким образом, на основании приведенных выше примеров можно говорить, что наиболее эффективным алгоритмом обучения полиморфной вейвлет-сети является метод Бройдена-Флетчера-Гольдфарба-Шанно. Причем значительная роль в обучении вейвлет-сети принадлежит правильной настройке масштабов вейвлетов a и дополнительных настраиваемых параметров p . Метод Бройдена-Флетчера-Гольдфарба-Шанно – один из наиболее часто используемых квазиньютоновских методов, при котором не

вычисляется непосредственно гессиан оптимизируемой функции. Вместо этого гессиан оценивается приближенно, что позволяет достичь высокой скорости работы алгоритма [6].

Выводы. В настоящей работе для идентификации нестационарных временных рядов предложена новая структура полиморфной вейвлет-сети с обратными связями. Рассмотрены алгоритмы для адаптивной настройки параметров сети и примеры аппроксимации нестационарных временных рядов, детерминированные составляющие которых моделируются численными решениями обыкновенных дифференциальных уравнений на некотором интервале времени. Показаны преимущества использования полиморфных вейвлет-сетей с обратными связями в сравнении с сетями без обратных связей.

Проведено исследование сетей с базовыми вейвлетами различных типов: Морле, POLYWOG и RASP. Рассмотрена задача идентификации нелинейных динамических систем, в которой в качестве идентификатора предложено использовать полиморфную вейвлет-сеть с обратными связями. Результаты моделирования показали, что разработанная вейвлет-сеть с обратными связями достаточно хорошо идентифицирует нелинейную динамическую систему при различных входных воздействиях. Полученные результаты могут быть использованы при создании вейвлет-сетевых систем управления нелинейными процессами.

Проведено сравнение различных численных алгоритмов оптимизации параметров полиморфной вейвлет-сети. Показано, что наилучшим алгоритмом обучения полиморфной вейвлет-сети является метод Бройдена-Флетчера-Гольдфарба-Шанно.

Литература

1. Льюнг Л. Идентификация систем. Теория для пользователя. – М.: Наука, 1991. – 432 с.
2. Narendra, K. Parthasarathy K. Identification and Control of Dynamical Systems Using Neural Networks // IEEE Transaction on Neural Networks. – 1990. – Vol. 1. – P. 4–27.
3. Верзунов С. Н. Синтез полиморфной вейвлет-сети и исследование ее свойств для аппроксимации нестационарных временных рядов // Информатика и системы управления. – Благовещенск. № 2 (2015) – С. 60–69.
4. Верзунов С.Н., Лыченко Н.М. Аппроксимация временных рядов полиморфной вейвлет-сетью с обратными связями // Математические структуры и моделирование. 2016. № 2 (38). С. 16-26.
5. Верзунов С.Н., Лыченко Н.М. Идентификация нелинейных динамических систем полиморфной вейвлет-сетью с обратными связями // Вестник Кыргызско-Российского Славянского университета. 2015. Т. – 15. № 9. – С. 18–22.
6. Верзунов С.Н., Лыченко Н.М. Аппроксимация нестационарных динамических систем полиморфной вейвлет-сетью // Фундаментальные и прикладные проблемы науки. Том 3. Материалы Кыргызской секции X международного симпозиума 8–10 сентября 2015 года. – М.: РАН, 2015 – С. 3–11.