Висновок. Методом локального потенціалу знайдено границю стійкості вільного конвективного руху відносно стоячих збурень нестисливої в'язкої рідини з парним профілем швидкості, зумовленим зовнішнім температурним градієнтом, перпендикулярним до напрямку гравітаційного поля.

Результати, отримані з використанням тільки однієї пробної функції для швидкості і температури, достатньо добре узгоджуються з даними, отриманими кінетичним методом з використанням складних пробних функцій **[**1**]**.

Література

1. Gerchuni G.Z. Convective stability of incompressible fluid / G.Z. Gerchuni, E.M. Zhukhovitsky, A.A. Yakimov // J. Heat Mass Transfer. – 1974. – Vol. 17. – 717 p.
2. Гершуни Г.З. Конвективная устойчивость несжимаемой жидкости / Г.З. Гершуни, И.М. Жуховицкий. – М. : Изд-во "Наука", 1972. – 392 с.
3. Гершуни Г.З. Результати, отримані з використанням тільки однієї пробної функції для швидкості і температури / Г.З. Гершуни, Е.Н. Жуховицкий // Вісник вузів : зб. наук. техн. праць. – Сер.: Фізика. – 2014. – № 45. – С. 43-48.
4. Левич В.Г. Физико-химическая гидродинамика / В.Г. Левич. – М. : Изд-во АН СССР, 1952, 132 с.
5. Гленсдорф П. Термодинамическая теория структуры, устойчивости и флуктуацій / П. Гленсдорф, И. Пригожин. – М. : Изд-во "Мир", 1973. – 273 с.
6. Стратонович Р.Л. Нелинейная неравновесная термодинамика / Р.Л. Стратонович. – М. : Изд-во "Наука", 1985. – 480 с.

Надійшла до редакції 23.10.2016 р.

Негрич В.В., Демьянчук Я.М., Процюк В.Г. **Исследование устойчивости конвективного движения с внутренним источником тепла**

Исследована гидродинамическая устойчивость вязкой несжимаемой жидкости, размещенной между вертикальными параллельными поверхностями. Рассмотрена свободная конвекция с четным профилем скорости. Для исследования использован вариационный принцип неравновесной термодинамики – метод локального потенциала. Определено критическое значение критерия Грасгофа в зависимости от числа Прандтля и значение волнового числа, при котором происходит переход от простой до сложной диссипативной структуры. Также учитывался угол наклона слоя по отношению к гравитационному полю. Представленный метод расчетов с использованием только одной пробной парной функции скорости и одной непарной для температуры позволяет получить удовлетворительные результаты.

Ключевые слова: конвективное движение жидкости, локальный потенциал, диссипативные структуры, критерий Грасгофа, критерий Прандтля.

Negrich V.V., Demyanchuk Ya.M., Protsiuk B.G. **The Study of the Stability of Convective Motions with Internal Heat Source**

We investigate the hydrodynamic stability of viscous incompressible fluid located between vertical parallel surfaces. We considered free convection with an even velocity profiles. For research we have used variational principle of non-equilibrium thermodynamics – the method of local potential. We have determined the critical Grashof criterion, depending on the Prandtl number and value of the wave number at which the transition from simple to complex dissipative structures occurs. We have also taken into account the angle of the layer with respect to the gravitational field. The method of calculation using only one test feature provides comparable results with the kinetic method, using complex test functions.

Keywords: convective fluid motion, the local potential, dissipative structures, criteria Grashof, Prandtl number.

УДК 519.63:004.02

РОЗВ'ЯЗУВАННЯ ОДНОВИМІРНОЇ ЗАДАЧІ ТЕПЛОПРОВІДНОСТІ РЕКУРЕНТНОЮ НЕЙРОННОЮ МЕРЕЖЕЮ ДЖОРДАНА

Н.О. Семенишин[1]

Наведено методологію для розв'язання крайових задач, а саме одновимірного рівняння теплопровідності за допомогою штучної нейронної мережі Джордана. Подано результати моделювання. В основі функції вартості мережі лежить метод Кранка-Ніколсона. Архітектура представленої нейронної мережі має класичну рекурентну структуру, але з додатковим прихованим шаром, в якій вузли вихідного шару мають ефект на вузли попереднього шару.

Для навчання рекурентної мережі використано розширення стандартного алгоритму зворотного поширення – "Зворотного поширення в часі". Його отримано шляхом розгортання часових операцій мережі в багатошарову мережу прямого поширення, топологія якої розширюється на один шар на кожному часовому кроці. Мета навчання нейронної мережі є зведення до мінімуму неув'язки вихідного рівняння, яке описує проблему.

Ключові слова: рівняння теплопровідності, рекурентна нейронна мережа Джордана, метод Кранка-Ніколсона, метод "Зворотного поширення в часі".

Вступ. Точні розв'язки крайових задач для рівнянь з частинними похідними, до яких зводиться дослідження багатьох важливих проблем практики, вдається знайти тільки для часткових випадків. Тому, зазвичай, такі задачі розв'язують наближено. Серед відомих числових методів для розв'язування крайових задач особливо поширеними є метод скінченних різниць (МСР) та метод скінченних елементів (МСЕ). Ідея цих методів полягає в редукції вихідної диференціальної задачі до дискретної: область зміни неперервного аргументу заміняють дискретною множиною точок (вузлів), які називають сіткою. Згідно з МСР, замість функцій неперервного аргументу розглядають функції дискретного аргументу, визначені у вузлах сітки. Похідні, що містять диференціальні рівняння і граничні умови, заміняють різницевими співвідношеннями. Теоретичною основою МСЕ є варіаційні методи.

Незважаючи на розвиток сучасної обчислювальної техніки, досягнення у програмному забезпеченні, конструюванні нових алгоритмів, є багато задач, які або не піддаються розв'язуванню наявними числовими методами, або не досягають задовільної точності. Це зумовило пошук нових ідейно методів, зокрема нейромережевих **[**1-5**]**. Набутий досвід застосування штучних нейронних мереж до розв'язування крайових задач математичної фізики і зумовив написання цього дослідження.

Мережа Джордана. В цій роботі розглядаємо одновимірну задачу теплопровідності для ізотропних матеріалів із граничними умовами першого типу

$$\begin{gathered}\frac{\partial u}{\partial t}=\alpha\frac{\partial^2 u}{\partial x^2},\quad 0<x<L,\ 0<t<T\\ \begin{cases}u(0,t)=C_1\\ u(L,t)=C_2\end{cases},\quad 0<t<T \qquad ,\\ u(x,0)=C_0(x),\quad 0<x<L\end{gathered} \tag{1}$$

[1] інж. Н.О. Семенишин – НЛТУ України, м. Львів

де: u – температура стрижня у даній точці x у певний момент часу t (градусів С); x і t – незалежні змінні; α – коефіцієнт температуропровідності дорівнює $\alpha = k/c\rho$ (м²/с); де c – теплоємність матеріалу стрижня (Дж/кг·°С); ρ – густина матеріалу стрижня (кг/м³) і k – коефіцієнт теплопровідності (Вт/(м·К)); $C_0(x)$ – початковий розподіл температури у стрижні; C_1, C_2 температура на лівому та правому кінці стрижня відповідно.

Для розв'язання цієї задачі розглянуто рекурентну штучну нейронну мережу Джордана. Мережа за функціональністю є дещо схожою на клітинну [4, 5] з тією відмінністю, що не використовуємо логіку задачі, що розв'язується. Тобто нейрон наступного шару мережі містить зв'язки всіх нейронів попереднього, а не тільки сусідів, які входять у деякий різницевий шаблон, щоправда використання шаблону істотно зменшує час навчання, за рахунок зменшення спектра вирішуваних задач математичної фізики.

Мережа Джордана – вид рекурентних нейронних мереж, яка отримується з багатошарового перцептрону, якщо на його прихований шар подати, крім вхідного вектора, вихідний із затримкою на один або кілька тактів. Використано мережу цього типу, оскільки задача теплопровідності містить часову залежність, а в рекурентних мережах є зворотний зв'язок, що дає змогу отримати інформацію з попереднього часового кроку. Ще одним важливим чинником вибору мережі Джордана є присутність прихованого шару, а це, водночас, позитивно впливає на апроксимацію розв'язку цієї задачі [1].

На першому етапі проводимо дискретизацію вихідної задачі і розіб'ємо прямокутну область зміни аргументів x і t на рівномірну сітку $\left\{(t_i, x_j)\right\}_{i=0,\, j=0}^{M_t M_x}$ де $t_i = i\Delta t,\ x_j = j\Delta x,\ \Delta t = T / M_t,\ \Delta x = L / M_x$, тут Δx і Δt – крок розбиття області по змінній x і t відповідно. Побудуємо нейромережу, що складається із M_x нейронів вихідного шару, яка буде давати на виході нейрона j для часового кроку i значення температури $u_{ij} = u(t_i, x_j),\ j = 1,...,M_x - 1$ на підставі вихідних значень нейронів попереднього часового кроку. При цьому $u_{0j} = u(0, x_j) = C_0(x)$; $u_{i0} = u(i,0) = C_1,\ u_{iM_x} = u(i, M_x) = C_2$ – ці значення подаються мережі на першому кроці по часу. Для прихованого шару кількість нейронів N встановлюється рівною кількості нейронів вихідного шару, проте це значення може відрізнятись.

На кожному часовому кроці i отримуємо нове значення температури із попередніх виходів за формулою

$$u_{ij} = \sum_{k=1}^{N} a_{jk} s\left(\sum_{l=1}^{M_x} w_{kl} u_{i-1l}\right), \quad 0 < i < M_t, \tag{2}$$

де: a – матриця ваг, що з'єднує прихований шар з вихідним; w – матриця ваг, що з'єднує вихідний шар попереднього часового кроку з прихованим шаром для даного моменту часу; s – функція активації, зазвичай сигмоїдного типу; N – кількість нейронів прихованого шару. Автор [7] для кращої збіжності радить обирати гіперболічний тангенс з коефіцієнтами

$$s(x) = 1{,}71 \tanh\left(\frac{2}{3} x\right). \tag{3}$$

Але для цієї функції вхідний сигнал має бути змасштабований, тобто потрібно відняти середньоквадратичне значення і за змогою декорелювати.

Цільова функція. Вибір цільової функції – ключовий момент при побудові нейронної мережі. Наша цільова функція (4) є складною квадратичною функцією, яка ставить у залежність виходи нейромережі для кожного моменту часу. Така структура функції дає змогу використовувати методи оптимізації нульового, першого чи, навіть, другого порядку.

$$J(u) = \sum_{i=0}^{M_t} \sum_{j=0}^{M_x} \left(\begin{array}{l} \dfrac{u_{i,j-1} + u_{i,j} + u_{i,j+1} - u_{i-1,j-1} - u_{i-1,j} - u_{i-1,j+1}}{3\Delta t} - \\ -\alpha \dfrac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1} + u_{i-1,j-1} - 2u_{i-1,j} + u_{i-1,j+1}}{2\Delta x^2} \end{array} \right)^2. \tag{4}$$

Як бачимо, функція є різницею лівої і правої частини (1) із заміною диференціального оператора на скінченно-різницевий методом *Кранка-Ніколсона* [10], піднесена до квадрату і просумована по всіх точках області. Ця функція мінімізується по a_{jk} і w_{kl}. Важливою перевагою такого підходу є те, що можемо навчити мережу для деякої множини початкових і граних умов.

Навчання нейронної мережі. Для навчання мережі Джордана обрано алгоритм *Зворотного поширення похибки в часі по епохах* (Epochwise Back-propagation through Time) – алгоритм навчання нейронних рекурентних мереж, який є розширенням стандартного алгоритму зворотного поширення [6]. Він може бути отриманий шляхом розгортання часових операцій мережі в багатошарову мережу прямого поширення, топологія якої розширюється на один шар для кожного кроку часу.

Навчання по епохах означає, що мережа починає свою роботу з деякого вихідного стану і розвивається доки не досягне іншого стану. У цій точці навчання призупиняється і мережа скидається у вихідний стан, після чого починається наступна епоха навчання. Термін "епоха" тут означає один приклад навчання, тобто в нашому випадку початкові умови однієї задачі.

Отже, наша мережа є, по суті, мережею прямого поширення із матрицею ваг, що повторюється кожного моменту часу. Відповідно, і метод навчання є дещо подібним до методу *Зворотного поширення похибки* [6]. Для цього методу шукаємо похідну від цільової функції по кожній синаптичній вазі. Тому для опису методу введемо поняття *Впорядкованої похідної* [8], яке за своєю суттю є звичайним ланцюговим правилом і використовується для того, щоб уникнути обчислення складних формул при виведенні похідної. Достоїнством такого підходу, на думку автора [6], є гомогенне трактування прямих і рекурентних зв'язків, тобто збільшення виразу при прямому поширенні веде до відповідного збільшення виразу при зворотному поширенні.

У випадку нашої мережі обчислюємо і зберігаємо локальний градієнт, тобто похідну від (4) по виходу для кожного нейрона, щоб передати його при зворотному проходженні іншим нейронам для аналогічних обчислень.

$$\frac{\partial^+ J}{\partial u_{ij}} = \frac{\partial J}{\partial u_{ij}} + \sum_{k=0}^{M_x} \frac{\partial^+ J}{\partial u_{i+1k}} * \frac{\partial u_{i+1k}}{\partial u_{ij}}. \tag{5}$$

У рівнянні (5) лівий член означає явну похідну, а правий неявну. Для кращого розуміння розглянемо такий приклад:

$$z_2 = 4 \cdot z_1;\; z_3 = 3 \cdot z_1 + 5 \cdot z_2 .$$

Частинна похідна z_3 по z_1 дорівнює 3 – це явна похідна (прямий вплив). Проте впорядкована похідна дорівнює 23 завдяки непрямому впливу z_1 через z_2. На підставі (5) та (2) виведемо формули для навчання нашої мережі.

Локальне індуковане поле нейрона [6] прихованого та вихідного шару відповідно, дорівнює

$$net_{ik} = \sum_{l=1}^{M_x} w_{kl} u_{i-1l},\; 1 < k < N;\; x_{ik} = s(net_{ik});\; u_{ij} = \sum_{k=1}^{N} a_{jk} x_{ik},\, 1 < j < M_x . \tag{6}$$

Із формули (6) маємо вихідні сигнали всіх нейронів. Перейдемо до виведення формул у зворотному режимі. Тут і надалі у формулах префікс *J_* означає похідну від функції (4) по змінній, яка йде після цього префікса.

$$J\_\hat{u}_{ij} = \frac{\partial J}{\partial u_{ij}} ; \tag{7}$$

$$\begin{aligned} J\_u_{ij} &= J\_\hat{u}_{ij} + \sum_{k=1}^{N} w_{kj} * J\_x_{i+1k} \quad M_x > j > 1 \\ J\_net_{ik} &= \sum_{j=1}^{M_x} a_{jk} * J\_u_{ij} \quad N > k > 1 \end{aligned} ; \tag{8}$$

$$J\_x_{ik} = s'(net_{ik}) * J\_net_{ik} \quad N > k > 1 . \tag{9}$$

Формули (8) і (9) обчислюють похідну згідно з ланцюговим правилом диференціювання від (4), щоб визначити вплив кожного нейрона на формування вихідного сигналу. Тут $J\_\hat{u}_{ij}$ – явна похідна від (4) по виходу нейрона *j* тільки у вихідному шарі мережі для кроку часу *i*, а J_u_{ij} – впорядкована похідна нейрона у вихідному шарі. Далі обчислюємо похідну від (4) по синаптичних вагах

$$J\_a_{kj} = \sum_{i=1}^{M_t} J\_u_{ik} * x_{ij}; \qquad J\_w_{kj} = \sum_{i=1}^{M_t} J\_x_{i+1k} * u_{ij} . \tag{10}$$

Методи оптимізації. Сформулюємо метод оптимізації для нашої мережі. Отже, задача полягає в тому, що потрібно знайти такі значення вагових коефіцієнтів мережі, щоб мінімізувати (4). Для цього введемо формулу, яка показує, що зміна матриці ваг має відбуватись по антиградієнту цільової функції

$$\Delta w_{i,j} = -\eta \frac{\partial J}{\partial w_{ij}}, \quad 0 < \eta < 1 , \tag{11}$$

де η – множник, що задає швидкість зміни *w*.

Зрозуміло, що $\frac{\partial J}{\partial w_{ij}} = J\_w_{ij}$ із формули (10). Залишається відкритим питання про те, в який спосіб має змінюватись *w*, щоб в (4) було досягнуто певного

"кращого" мінімуму. Найпростішим таким способом є *Метод найшвидшого спуску*

$$New\_w_{i,j} = w_{i,j} - \eta \cdot J\_w_{i,j} , \tag{12}$$

де $New\_w_{i,j}$ – це нове значення синаптичних ваг. Але цей метод має дуже звивистий характер, рух майже перпендикулярний дну заглиблення, тоді як треба рухатись уздовж. Тому рекомендують вибирати ефективніші методи наприклад різновид методу *Спряжених градієнтів – Флетчера-Рівса* [1].

$$w_{k+1} = w_k + \eta p_k, \qquad p_k = g_k + \beta p_{k-1};\; p_0 = g_0, \quad \beta_k = \frac{(g_k, g_k)}{(g_{k-1}, g_{k-1})} , \tag{13}$$

де $g_k = -\frac{\partial J}{\partial w}$. Тут варто зазначити, що k – номер глобальної ітерації навчання нейронної мережі і g_k присвоюються нові значення із (10) для кожної такої ітерації відповідно, доки зміни не стануть достатньо малими або похибка (4) досягне малого рівня.

Аналогічні обчислення виконують і для матриці *a* з єдиною відмінністю $g_k = -\frac{\partial J}{\partial a}$.

Результати моделювання. Щоб дослідити властивості апроксимації нашої мережі, розглянемо приклад одновимірної задачі теплопровідності (1).

Дані початкові умови $u(0,x) = 0{,}4\sin(\frac{\pi x}{L})$, 0 <= *x* <= *L*, *L* =2; 0 <= *t* <=*T*, *T*=0,05; Δt = 0,005; Δx = 0,2; $u(t,0) = 0$, $u(t,L) = 0$; коефіцієнт температуропровідності (α) = 2,4094 (м²/с);

Точний розв'язок $u(t,x) = 0{,}2e^{\frac{-\alpha\pi^2 t}{L^2}} 2\sin(\frac{\pi x}{L})$.

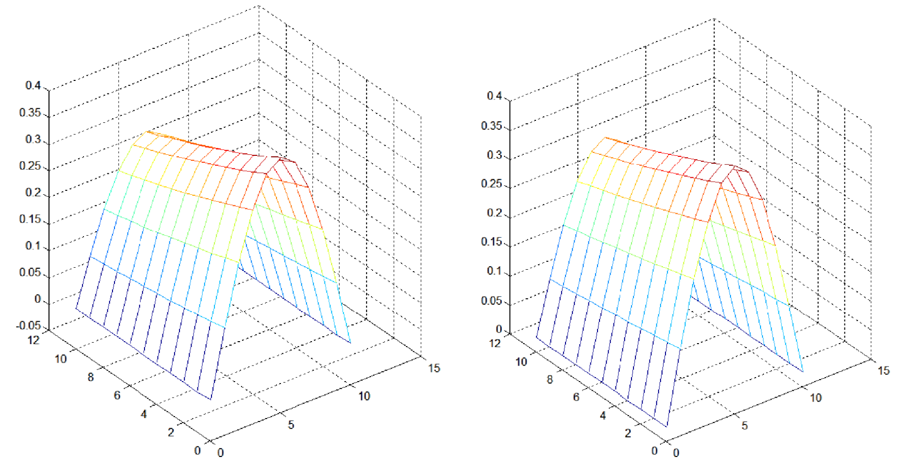

Рис. Температурний розподіл задачі в середовищі Matlab
(зліва результат нейронної мережі, справа точний розв'язок)

На графіках видно зображення точної функції – права поверхня та графік функції апроксимованої нашою мережею – ліва поверхня. Обидва графіки майже ідентичні. Середньоквадратична похибка дорівнює 0,0042 для всієї поверхні (рис.).

Висновки:

1. Проведено аналіз особливостей застосування нейромережевого підходу при побудові наближених розв'язків крайових задач для рівнянь параболічного типу. Розглянуто важливу для практики задачу теплопровідності.
2. Підібрано архітектуру рекурентної нейронної мережі зі зворотним зв'язком (мережа Джордана) для розв'язування задачі теплопровідності.

Результати моделювання дають змогу виявити переваги та недоліки нейромережевого підходу. До недоліків можна віднести те, що метод скінченних різниць дає значення шуканої функції у деякій скінченній множині точок. Це можна виправити, використавши другу мережу (чи метод) для інтерполяції.

До переваг відносять здатність мережі давати результат за різних початкових і граничних умов; використовувати некласичну постановку задачі – додавати експериментальні значення прямо в цільову функцію; також класичні переваги нейронних мереж [6]: нелінійність, завадостійкість, одноманітність аналізу та проектування – мережа може використовуватись в різних предметних областях. За допомогою рекурентних мереж можна моделювати і більш складні системи – з фазовими переходами, наприклад задачу Стефана [1].

Література

1. Васильев А.Н. Нейросетевое моделирование. Принципы. Алгоритмы. Приложения / А.Н. Васильев, Д.А. Тархов. – СПб. : Изд-во Государственный Политехнический Университет, 2009. – С. 34-42.
2. Горбаченко В.И. Распараллеливание нейросетевого алгоритма решения краевых задач математической физики на многоядерных процессорах NVIDIA / В.И. Горбаченко, Н.О. Матвеева, Е.И. Гурин, 2012. – М. : Изд-во "Наука". – 342 с.
3. Горбаченко В.И. Нейрокомпьютеры в решении краевых задач теории поля / В.И. Горбаченко. – К. : Изд-во "Лыбидь", 2003. – 336 с.
4. Нестеренко Б.Б. Асинхронні паралельні алгоритми нейронних мереж / Б.Б. Нестеренко, М.А. Новаторський. – М. : Изд-во "Вершина". – 238 с.
5. Новотарський М.А. Штучні нейронні мережі: обчислення / М.А. Новотарський, Б.Б. Нестеренко // Праці Ін-ту математики НАН України. – К. : Вид-во Ін-ту математики НАН України. – 2004. – Т. 50. – 408 с.
6. Саймон Хайкин. Нейронные сети: полный курс : пер. с англ. / Хайкин Саймон. – Изд. 2-ое, [перераб. и доп.]. – М. : Изд. дом "Вильямса", 2006. – 1104 с.
7. LeCun Y. Efficient BackProp / Y. LeCun, L. Bottou, Orr, G. and K. Muller (Eds), Neural Networks: Tricks of the trade, Springer, 1998. – 322 p.
8. Werbos P.J. Backpropagation through time: what it does and how to do it / P.J. Werbos // Proceedings of the IEEE, No. 10, October. – 1990. – Vol. 78. – Pp. 1550-1560.
9. [Electronic resource]. – Mode of access http://en.wikipedia.org/wiki/Crank%E2 %80%93 Nicolson_method

Надійшла до редакції 22.10.2016 р.

Семенишин Н.О. **Решение одномерного уравнения теплопроводности с помощью рекуррентной нейронной сети Джордана**

Приведена методология для решения краевых задач, а именно одномерного уравнения теплопроводности с помощью искусственной нейронной сети Джордана. Представлены результаты моделирования. В основе функции стоимости сети лежит метод Кранка-Николсона. Архитектура представленной нейронной сети имеет классическую рекуррентную структуру, но с дополнительным скрытым слоем, в котором узлы выходного слоя имеют эффект на узлы предыдущего слоя.

Для обучения рекуррентной сети использовано расширение стандартного алгоритма обратного распространения – "Back-propagation –through-time". Он получен путем развертывания временных операций сети в многослойную сеть прямого распространения, топология которой расширяется на один слой на каждом временном шаге. Фактически этот алгоритм подразумевает цепное правило дифференцирования, где производная от функции стоимости по выходам нейронов и их весам идет в обратном направлении через всю сеть. Целью рекуррентного обучения нейронной сети является сведение к минимуму невязки исходного уравнения, которое описывает проблему.

Ключевые слова: уравнение теплопроводности, рекуррентная нейронная сеть Джордана, метод Кранка-Николсона, метод "Обратного распространения во времени".

Semenyshyn N.O. **Solving Heat Equation in One Dimension Using Jordan Recurrent Neural Network**

The author has presented a methodology to solve boundary value problems, namely heat equation in one dimension using Jordan artificial neural networks. The results of simulation are presented. A recurrent neural network has been constructed for the implementation of the Crank – Nicolson method, which is the basis for constructed networks cost functions. The architecture of the presented Jordan neural network has almost classical recurrent structure where nodes of output layer have effect to nodes of former layer. The difference is that such a network contain additional hidden layer. This architecture was used to get better rate of convergence of neural network training method and also to have possibility to obtain output value from previous time step. For training a recurrent network an extension of the standart back-propagation algorithm was used referred to as "Back-propagation –through-time". It is derived by unfolding the temporal operation of the network into a layered feedforward network, the topology of which grows by one layer at every time step. In fact this algorithm imply simple chain rule of differentiation, where derivative of cost function with respect to neurons outputs and weights going backwards through the whole network. The purpose of the recurrent neural network training is to minimize the discrepancy of the original equation, which describe the problem.

Keywords: heat equation, Jordan recurrent neural network, Crank – Nicolson method, "back-propagation –through-time" method.