Engineering Applications of Artificial Intelligence 70 (2018) 184–198

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Inter-domain routing for communication networks using Hierarchical Hopfield Neural Networks

Hitalo O. Silva *, Carmelo J.A. Bastos-Filho

Polytechnic School of Pernambuco, University of Pernambuco, Recife, Brazil

ARTICLE INFO

Keywords:
Hopfield Neural Networks
Routing algorithms
Communication networks
Artificial neural networks

ABSTRACT

This paper presents the Hierarchical Hopfield Neural Networks (HHNN). HHNN is a novel Hopfield Neural Network (HNN) approach. HHNN is composed of a hierarchy of self-sufficient HNNs, aiming to reduce the neural network structure and mitigate convergence problems. The HNNN composition depends on the applied problem. In this paper, the problem approached is the inter-domain routing for communication networks. Thus, the hierarchy of HNNs mimics the structure of communication networks (domains, nodes, and links). The proof of concept and the comparison between HNNN with the state-of-art HNN occurs using an implementation of them in the Java programming language. Besides, the performance analysis of the HHNN runs on a parallel hardware platform, using VHDL to develop it. The results have demonstrated a reduction of 93.75% and 99.98% in the number of neurons and connections to build the neural network, respectively. Furthermore, the mean time to achieve convergence of HHNN is rough 1.52% of the total time needed by the current state-of-art HNN approach. It is also less susceptible to early convergence problems when used in communications networks with a large number of nodes. Last, but not least, the VHDL implementation shows that convergence time of HHNN is comparable to routing algorithms used in practical applications.

1. Introduction

The dissemination of personal computers and smartphones has increased the demand for high data transmission networks since these devices are handling several media and applications, such as social networks. New data transmission technologies, efficient routing protocols, and more robust network elements have been proposed to support these claims. Service Providers (SPs) have incorporated some techniques to provide high-quality services, to minimize operational costs and to generate a competitive advantage within the market. These adaptations may demand significant changes in SP's network infrastructure. By contrast, reliability is essential to achieve high throughput with a reasonable Quality of Service (QoS). Several aspects are relevant to do it for communication networks. For example, redundant network elements and links; and adaptive routing protocols. Routing protocols must present some desired characteristics, such as fast response and scalability with the number of network nodes. In general, those features lead to a better network performance.

Routing is a process of establishing paths whereby data between different nodes of the communication network should pass through. Routing Algorithms are the algorithms that calculate those routes. A routing algorithm should select a path in real-time, according to network topology changes. It also needs to respect the necessary QoS demand. The routing process has a significant impact on a communication network performance. In general, heuristics based on pre-defined metrics are used to define the route between two nodes. Such as Shortest Path (SP), Minimum Hops (MH), Least Resistance Weight (LRW) (Wen et al., 2005), Optical Signal to Noise Ratio Routing (OSNR-R) (Martins-Filho et al., 2003), Physical Impairments Aware Adaptive Weight Function (PI-AWF) (Chaves et al., 2007) and Power Series Routing (PSR) (Martins-Filho et al., 2008). Among the routing algorithms, which are applied in network communication, employing these metrics, it is possible to highlight the Dijkstra algorithm (Link State) (Dijkstra, 1959) and the Bellman–Ford algorithm (Distance Vector) (Bellman, 1958; Bonabeau et al., 1998).

The Internet is a conglomerate of Autonomous Systems (AS), and Different administrative authorities control them. Routing policies drive how the ASs interconnect with each other. Routers, equipment that performs routing, run Interior Gateway Protocols (IGPs) within their domains. The IGPs most adopted are Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS–IS). Besides, Exterior Gateway Protocol (EGP) establishes routes between different ASs. They were introduced because IGPs do not support network with

* Corresponding author.
E-mail addresses: hos@ecomp.poli.br (H.O. Silva), carmelofilho@ieee.org (C.J.A. Bastos-Filho).

https://doi.org/10.1016/j.engappai.2018.02.002
Received 19 May 2017; Received in revised form 29 December 2017; Accepted 1 February 2018

thousands of nodes and hundreds of thousands of routes. Even though, IGPs were not developed for this purpose. The establishment of routes between different domains is one of the significant challenges in the area of communication networks. Several algorithms and protocols have been proposed in the literature and standards. Among these algorithms, the most used is the Border Gateway Protocol Version 4 (BGPv4). Being defined in the Request For Comment (RFC) 1771 (Halabi and McPherson, 2000; Warnock and Nathoo, 2011). BGP establishes routes between different domains on the Internet. The number of hops between ASs and some pre-defined network policies are the insights employed to define the best routes.

Despite the BGP qualities, there are some well-known problems about its operation. Among them, for example, the routing tables are large; it may produce different routes for the same source–destination pair, generating instability; sub-optimal paths can arise during the network operation; the time needed to achieve convergence can be extremely high when the number of nodes increases (Pelsser, 0000; Potaroo, 0000).

Computational Intelligence (Engelbrecht, 2007) has been used for routing in communication networks. Some examples of algorithms that have been deployed for this purpose are Ant Colony Optimization (ACO) (Navarro-Varela and Sinclair, 1999; Garlick and Barr, 2002; Bastos-Filho et al., 2009), Genetic Algorithms (GA) (Banerjee et al., 2004; Le et al., 0000) and Hopfield Neural Networks (HNN) (Hopfield, 1982). HNN is a type of recurrent neural network cited as a promise solution for routing because of its adaptation capability (Rauch and Winarske, 1988; Kojic et al., 2007, 2009; Bastos-Filho et al., 2010). Adaptability is a highly desirable feature for routing since communication networks environment can change during the operation. Besides, HNNs do not need a prior training process and the HNN related matrices fully define the convergence conditions. However, the response time of HNN to provide a route between a pair of edge nodes in large networks is still higher than other widely applied routing algorithms, such as the Dijkstra's algorithm (Dijkstra, 1959).

Although some relevant approaches based on HNNs were proposed in the literature, they were not developed to be applied in networks with many domains and with a large number of nodes. The central hypothesis is that HNN models that present these two characteristics can run in large-scale networks, such as the Internet. Thus, it is proposed in this paper a hierarchical HNN model to mitigate these current limitations of the already proposed HNNs.

The remainder of this paper is organized as follows: Section 2 presents the background concepts regarding HNN; Section 3 presents our proposal; Section 4 shows the simulation setup; Section 5 presents the results; Finally, Section 6 presents the conclusions.

2. Hopfield neural networks applied to routing in communication networks

Routing in communication networks consists of finding the shortest path from a source node s to a destination node d. A communication network topology can be described by an undirected graph $G = (N, A)$, which N is the number of nodes (the vertices of the topology) and A is a set of links (arcs or edges). For each arc (xi) there is a non-negative value C_{xi} representing the cost of the link. A path P^{nd} is a sequence of ordered nodes connecting s to d: $P^{sd} = (s, i, j, k, \dots, r, d)$. The total cost of a path is defined by $L^{sd} = C_{si} + C_{ij} + C_{jk} + \cdots + C_{rd}$. The problem is to find the path that contains the least total cost L^{sd}.

Rauch and Winarske (1988) created the first deployment of an HNN to solve the problem of the shortest path from a given source–destination pair. Authors created and published others HNN models proposing improvements. They were not adaptive enough to run in dynamic communication networks. In 1993, Ali and Kamoun (1993) introduced a new model in which the synaptic weight matrix $(T_{xi,yj})$ carries only the convergence attributes. The external polarization of neurons (I_{xi}) contains the information about the link cost and the topology.

Ali and Kamoun's (1993) model is organized in a $n{\times}n$ matrix, where n is the number of nodes in the network topology. A neuron represents each element in the matrix. For each link (even if it do not exist in the topology) between two adjacent nodes (connection from node x to node i) in the communication network, there will be a neuron (xi) associated. All diagonal elements in the matrix are equal to zero since one node cannot be self-connected for this type of problem. As a consequence, one needs $n \cdot (n-1)$ neurons to represent a communication network with n communication nodes.

The cost of the link from node x to node i is denoted by C_{xi}, that assuming positive real and normalized values between $(0, 1]$. The cost will be zero for non-existent links. ρ_{xi} is the matrix that defines whether the link xi exists in the network. If the link xi does not exist, then $\rho_{xi} = 1$, otherwise $\rho_{xi} = 0$.

Fig. 1 shows a small communication network represented by a graph and its matrices of cost and topology.

It is necessary to determine an energy function to guide the neural network to the lowest energy state when establishing routes using HNN. This state must correspond to the solution of the shortest path. The energy function should favor states that correspond to valid paths between the source–destination pair specified. Among them, it should also support those who have the shortest length. Ali and Kamoun (1993) proposed an energy function for this purpose:

$$E = E_1 + E_2 + E_3 + E_4 + E_5 \tag{1}$$

where:

$$E_1 = \frac{\mu_1}{2} \sum_{\substack{x=1 \\ (x,i)\neq(d,s)}}^{n} \sum_{\substack{i=1 \\ i\neq x}}^{n} C_{xi} V_{xi}, \tag{1a}$$

$$E_2 = \frac{\mu_2}{2} \sum_{\substack{x=1 \\ (x,i)\neq(d,s)}}^{n} \sum_{\substack{i=1 \\ i\neq x}}^{n} \rho_{xi} V_{xi}, \tag{1b}$$

$$E_3 = \frac{\mu_4}{2} \sum_{x=1}^{n} \left\{ \sum_{\substack{i=1 \\ i\neq x}}^{n} V_{xi} - \sum_{\substack{i=1 \\ i\neq x}}^{n} V_{ix} \right\}^2, \tag{1c}$$

$$E_4 = \frac{\mu_3}{2} \sum_{x=1}^{n} \sum_{\substack{i=1 \\ i\neq x}}^{n} V_{xi}(1 - V_{xi}), \tag{1d}$$

$$E_5 = \frac{\mu_5}{2}(1 - V_{ds}) \tag{1e}$$

where μ_1, μ_2, μ_3, μ_4 and μ_5 are constants. Each term of Eq. (1) has a particular function during the routing process. The term (1a) minimizes the total cost of a path, considering only the cost of existing links. The term (1b) avoids non-existent links being part of the path. The term (1c) is zero if, for every node in the solution, the number of inbound links is equal to the number of outbound links. The term (1d) drives the neural network to converge to one of 2^{n^2-2} stable states. The term (1e) is zero when the output of the neuron (d, s) assumes the value 1. Although the link from d to s is not part of the topology, it is introduced to ensure that the source and destination nodes are part of the solution. The external polarization of neurons (I_{xi}) is used in this model to adjust the excitation level of the entire network and to pass data about the network topology, as well as the source and destination, to the neural network. The expression for I_{xi} is defined by:

$$I_{xi} = \frac{\mu_1}{2} C_{xi}(1 - \delta_{xd}\delta_{is}) - \frac{\mu_2}{2} \rho_{xi}(1 - \delta_{xd}\delta_{is}) - \frac{\mu_4}{2} + \frac{\mu_5}{2} \delta_{xd}\delta_{is}, \tag{2}$$

where δ is the Kronecker delta function, defined by:

$$\delta_{ab} = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{on the other cases.} \end{cases} \tag{3}$$