

APPLICATION OF ARTIFICIAL INTELLIGENCE FOR REAL-TIME STEAMING IN WEB SITES

Nenad Kojić¹⁶
Natalija Vugdelija¹⁷

<https://doi.org/10.31410/itema.2018.37>

Abstract: *When users of websites have the need to access some real-time audio or video signal into a stream, this can cause delays and interruptions in the transmission. This is especially pronounced when a large number of users want to access the server at the same time and look at the same content, thus increasing the linkwave's load. In this paper, one solution of the algorithm, based on artificial intelligence, is presented. The proposed algorithm finds Steiner tree and Pareto an optimal path for transmitting multimedia signals in topology and conditions of the network that is dynamically changing.*

Keywords: *Artificial intelligence, neural network, multimedia streaming, web sites*

1. INTRODUCTION

The intense development of hardware and software has contributed to the fact that a large percentage of people today have different types of electronic devices and direct access to the Internet. These opportunities have increased the need for the most prominent types of services and services offered through websites [1]. Thus, web sites and their versions on various types of devices have become a need of a modern society in everyday life. Today, most users start the day with reading the latest news, watching weather forecasts, tracking traffic conditions, and so on.

The first sites only had a textual content, and that was enough for the information to be transmitted. Over time, the demands of the users, and thus the sites that companies have created, have become increasingly complex and technologically more complex in the desire to be better, more beautiful and innovative. The tastes and demands of the users are constantly increasing and the market constantly searches for something new and better. All these requirements are followed by other areas in the IT industry [2]: development of telecommunication network, development of protocols, development of hardware components, devices, ...

It has long been known that color, and therefore images, have a big impact on the user and the perception of the entire site [1]. Very quickly the sites began to be flooded with a lot of different images. However, the market demanded even more intensive content, and so quickly the need to display video and audio files on the sites. In the first versions of sites this was not possible and such multimedia files could only be downloaded to the personal account and then viewed. As user and market demands became more and more intense, communication networks, web servers and websites had to provide real-time transmission, ie, streaming audio and video signals and within web sites [3]. Today's young users mostly use these types of content and consider it to be the default web site offer.

¹⁶ ICT College for vocational studies, Belgrade, Serbia

¹⁷ ICT College for vocational studies, Belgrade, Serbia

On the other hand, the real-time transmission of audio and video signals is the most complex thing from the aspect of the communication network [4]. This relates to network resource bandwidth, transmission quality, delayed transmission and download speeds, resulting in interruption in user playback, high data transfer rates, and slow loading.

For this reason, a large number of scientific papers are devoted to solving the most optimal way of distributing audio and video signals in order to achieve better quality, faster and cheaper transmission and less engagement of network resources [5], [6], [7].

Special attention is drawn to services that simultaneously access a large number of users, such as video on demand, ip television or realtime broadcasts of radio stations.

Sites offering such content are one of the most complex ways of transmitting and displaying data. Users expect to see audio or video content, in real time, without delay, with more users wanting to view the same content at the same time. This content is displayed in the browser, but it encounters a complex logic of routing through the global network. When more users want to view the same content at the same time, then such a routing is called multicast routing [8].

Routing is one of the most important aspects of communication networks. The basic idea of multicast routing is to transfer copies of packages to their destinations [8], [9].

Optimization of parameters involved in creating routing criteria can be divided into two segments. The first is finding the shortest path tree, in terms of the lowest price between the source and the destination, and the other, optimizing the total cost of routing in the multicast tree, in case some of the tree nodes are not members of the destination group [8], [9].

If the conditions in the network change over time and depending on several parameters, we say that it is dynamic multicast routing.

In this paper we will analyze the possibilities of implementing additional intelligence in the form of neural networks in order to improve the quality of transmission of multimedia content, via the Internet, for viewing in web sites, which is based on dynamic multicast routing. As this routing can be realized in a number of ways, and thus improved, in this paper will be used artificial intelligence, and the possibility of applying Hopfield's neural network [10] for the improvement of the routing process.

This work is organized in four chapters: after the introduction, the basic principles of the work of Hopfield's neural network are explained. After that, it is possible to modify it and apply it in the dynamic multicast routing process of the multimedia content. Finally, conclusions and guidelines for future work are given.

2. HOPFIELD NEURAL NETWORK

Hopfield neural network is one type of artificial neural network [10]. This network belongs to a group of recurrent neural networks, because the output signal, with modification, returns to the input, and thus corrects the state of the network. The processing elements (neurons) are full-connected: output, v_i , of each i th cell is connected to inputs of all other neurons via synaptic weights, T_{ij} , producing the modification of cell inputs, u_i , and thus changing the network state, Fig. 1. If the system is stable successive iterations lead to smaller and smaller output changes and network reaches some minima of the system energy.

In hardware implementation, processing cells are realized as summing amplifiers, with nonlinear transfer function $g(u_i)=v_i/u_i$, called also an *activation function*. In Hopfield's early work [10] the activation function was a simple threshold (*hard limiter*), and the model was a discrete one. Later on, Hopfield introduced continuous activation function [11]. A common choice is the sigmoidal or logistic function:

$$g_i(u_i) = \frac{1}{1+e^{-a_i \cdot u_i}}. \quad (1)$$

State equations of the Hopfield NN is described by:

$$C_i \frac{du_i}{dt} = \sum_{j=1, j \neq i}^N T_{ij} v_j - \frac{u_i}{R_i} + I_i, \quad i = 1, 2, \dots, N \quad (2)$$

where R_i is an equivalent resistance connected to the cell's capacitor C_i . Usually (but not necessarily) all neurons are identical, except synaptic conductance.

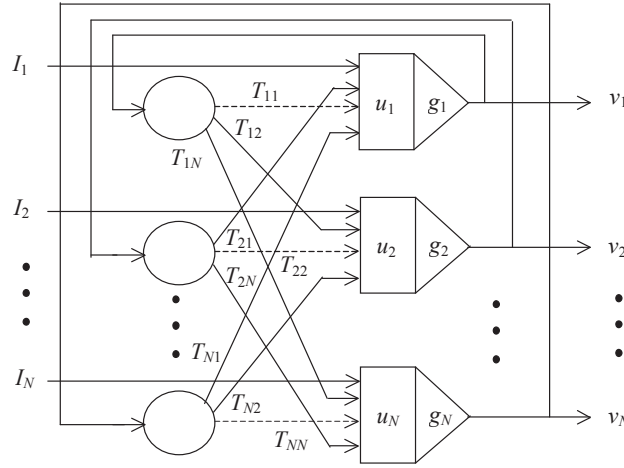


Figure 1: Hopfield network structure

Under these assumptions, and if constants a_i are sufficiently large (for instance, $a_i > 100$), the stability of the network, in Liapunov sense, may be verified by observing the energy function, E , describing the state of the network:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N T_{ij} v_j v_i - \sum_{i=1}^N I_i v_i. \quad (3)$$

Significant improvements in neural network algorithm are derived in paper offered by Ali and Kamoun [12]. For n nodes problem their computational network uses $n(n-1)$ neurons – the diagonal elements in the connection matrix $n \times n$ are removed – and find the shortest path from final stable neuron states. A suitable energy function is of the form:

$$\begin{aligned} E = & \frac{\mu_1}{2} \sum_X \sum_{i \neq X} C_{Xi} v_{Xi} + \frac{\mu_2}{2} \sum_X \sum_{i \neq X} \rho_{Xi} v_{Xi} + \frac{\mu_5}{2} (1 - v_{ds}) \\ & + \frac{\mu_3}{2} \sum_X \left(\sum_{i \neq X} v_{Xi} - \sum_{i \neq X} v_{iX} \right)^2 + \frac{\mu_4}{2} \sum_i \sum_{X \neq i} v_{Xi} (1 - v_{Xi}) \end{aligned} \quad (4)$$

Coefficients C_{Xi} are the link costs from node X to node i and the terms ρ_{Xi} describe the connection between nodes: the value is 1 if nodes are not connected, and 0 for connected nodes. The term μ_1 minimizes the total cost; μ_2 prevents nonexistent links from being included in the chosen path; μ_3 is zero for every node in the valid path (the number of incoming links is equal to the number of outgoing links); μ_4 forces the state of the neural network to converge to one of the stable states – corners of the hypercube defined by $v_i \in \{0,1\}$. The state v_i is close to 1 for node belonging to the valid path, otherwise the state is close to 0. The term μ_5 is zero when the output v_{ds} is equal to 1. This term is introduced to ensure the source and the destination nodes belong to the solution (the shortest path).

3. PROPOSED SOLUTION

Unlike point-to-point routing, now the optimization and definition of matrices in the queue for each of the $S-Dm$ pairs, where Dm is a set of destination nodes. Interactions with each other in terms of minimizing the final solution are given through the first member of the energy function, while the total energy function is obtained after each iteration as:

$$E = \sum_{m=1}^{Dm} E^m \quad (5)$$

This proposal is presented in [13], based on the initial form of the energy function in [14], where the energy function for solving the problem of multicast routing is considered, which considers physical distance and total delay in terms of the maximum allowed delay on the longest path in the tree. By modifying this function and introducing the parameters for capacity and density, but additionally the quality of the Q signal, for the purposes of this paper, the energy function is derived:

$$\begin{aligned} E^m = & \frac{\mu_1}{2} \sum_X \sum_{i \neq X} C_{Xi} v_{Xi}^m \frac{1}{1 + \sum_{\substack{j \neq m \\ j \in D \\ (X,i) \neq (m,s)}} v_{Xi}^j} + \frac{\mu_2}{2} \sum_X \sum_{\substack{i \neq X \\ (X,i) \neq (m,s)}} \rho_{Xi} v_{Xi}^m \\ & + \frac{\mu_3}{2} \sum_X \left(\sum_{i \neq X} v_{Xi}^m - \sum_{i \neq X} v_{Xi}^m \right)^2 + \frac{\mu_4}{2} \sum_i \sum_{X \neq i} v_{Xi}^m (1 - v_{Xi}^m) \\ & + \frac{\mu_5}{2} (1 - v_{ms}^m) + \frac{\mu_6}{2} \sum_X \sum_{\substack{i \neq X \\ (X,i) \neq (m,s)}} [1 - (K_{Xi} - G_{Xi})] v_{Xi}^m \end{aligned} \quad (5)$$

to which the member $\mu_7 Q_{Xi}$ is added and represents the matrix of the signal quality in each of the links. The network consists of routers that are linked by links. The number of routers and links can be arbitrary. One of the routers is connected to the web server on which the multimedia content is located. The other routers in the network connect users to their browsers, and they want to get content from the same web server.

Starting from the network structure given in Fig. 2 and setting the parameters of capacity and density to neutral values 1 ie. 0.1 the same result was obtained as in [13], [15], which shows that the newly introduced matrices do not disturb the optimal solution. The initial price matrix is given in Table 1, while matrix Q is randomly generated, while the obtained path is shown in Fig. 3.

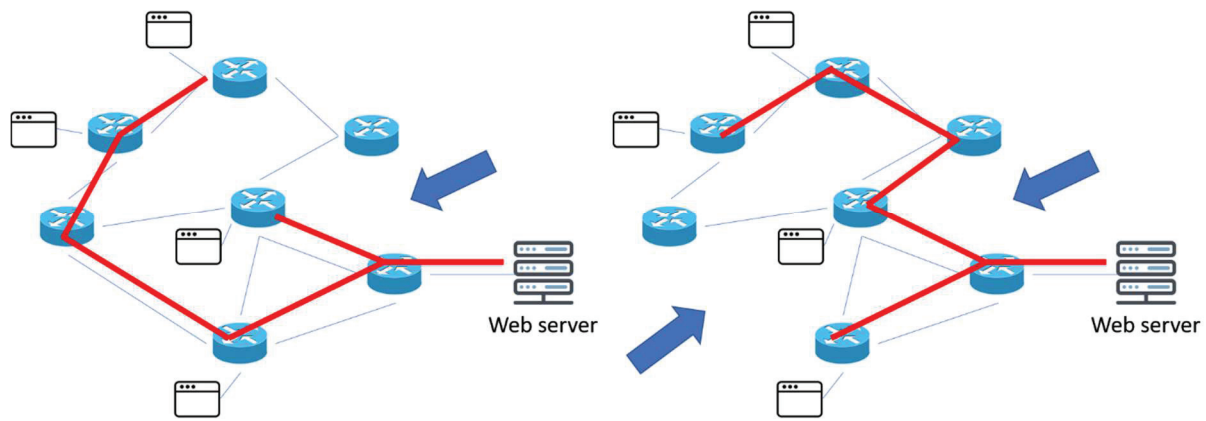


Figure 4: a) Paths after the termination of the link indicated by the right arrow b) Paths after the termination of the link indicated by the right and left arrow.

In a similar way, the proposed algorithm can detect all changes in the topology of the network as well as in the state of the links. Changing the value of the parameters on the links directly affects the state of the input parameters in the algorithm, and the algorithm finds a new path. This new path is relatively fast, as it should allow for a short stop to transmit the multimedia signal from the web server to the browser. Additionally, the path should still include all users who are currently connected to a web server, with as few branches as possible to make the transfer speed as fast as possible.

4. CONCLUSION

This paper presents one solution for transferring audio and video signals from the web server to clients, which access the content at the same time. To find the path, a Steiner Tree has been used that has the least number of links to connect the desired users. Hopfield's neural network was used to find the path, as a form of artificial intelligence. The goal of the algorithm is to detect changes in network topology or link parameters as soon as possible, and on the basis of this information find the Pareto optimal path. This path is found in a dynamic environment that is one of the most complex, and which is needed in situations where there is a required guaranteed quality of signal transmission across the network. Further research will focus on the hardware implementation of the proposed solution and testing in the networks of higher topology.

REFERENCES

- [1] Herring, S. C. (2009). Web content analysis: Expanding the paradigm. In *International handbook of Internet research* (pp. 233-249). Springer, Dordrecht.
- [2] Zhang, D. (2007). Web content adaptation for mobile handheld devices. *Communications of the ACM*, 50(2), 75-79.
- [3] Sánchez-Nielsen, E., Chávez-Gutiérrez, F., Lorenzo-Navarro, J., & Castrillón-Santana, M. (2017). A multimedia system to produce and deliver video fragments on demand on parliamentary websites. *Multimedia Tools and Applications*, 76(5), 6281-6307.
- [4] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, 17(4), 2347-2376.

- [5] Lee, H. R., Jung, T. J., Seo, K. D., & Kim, C. K. (2015). Delay constrained ARQ mechanism for MPEG media transport protocol based video streaming over Internet. *AFIN* 2015, 68.
- [6] Xu, J., Andreopoulos, Y., Xiao, Y., & van Der Schaar, M. (2014). Non-stationary resource allocation policies for delay-constrained video streaming: Application to video over Internet-of-Things-enabled networks. *IEEE Journal on Selected Areas in Communications*, 32(4), 782-794.
- [7] Kumar, N., Zeadally, S., & Rodrigues, J. J. (2015). QoS-aware hierarchical web caching scheme for online video streaming applications in internet-based vehicular ad hoc networks. *IEEE Transactions on Industrial Electronics*, 62(12), 7892-7900.
- [8] L. A. Giuliano, B. M. Edwards, B. R. Wright, *Interdomain Multicast Routing*, Addison-Wesley, 2002.
- [9] D. Chakraborty, G. Chakraborty, C. Pornovalai, N. Shiratori, "Cost Minimization for Dynamic Multicast without Rerouting", *Proceeding of the Internet Global Summit 1999*, San Jose, June 22-25, 1999
- [10] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", *Proc. Nat. Acad. Sci.*, Vol. 79, pp. 2554-2558, 1982.
- [11] J. J. Hopfield, D. W Tank, "'Neural' computations of decision in optimization problems", *Biol. Cybern.*, Vol. 52, pp. 141-152, 1985
- [12] M. Ali, F. Kamoun, "Neural networks for shortest path computation and routing in computer networks", *IEEE Trans. on Neural Networks*, Vol. 4, No. 6, pp. 941-953, 1993.
- [13] N. Kojić, I. Reljin, B. Reljin, "Primena neuralnih mreža za dinamičko multicast rutiranje video signala", *TELFOR*, Beograd , Nov. 20-22, 2007.
- [14] C. Pornovalai, G. Chakraborty, N. Shiratori, "A neural network approach to multicast routing in real-time communication networks", *Third International Conference on Network Protocols (ICNP'95)*, pp. 332-339, 1995.
- [15] Kojic, N., Reljin, I., & Reljin, B. (2013). Neural network based dynamic multicast routing. *Elektronika ir Elektrotehnika*, 19(3), 92-97.