Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 62 (2015) 419 – 426

Procedia Computer Science

The 2015 International Conference on Soft Computing and Software Engineering (SCSE 2015)

MRAI optimization for BGP convergence time reduction without increasing the number of advertisement messages

E.A. Alabdulkreem*, H.S. Al-Raweshidy, M.F. Abbod

College of Engineering & Design and physical sciences, Brunel University

London, UK, UB8 3PH

Abstract

The primary cause for the slowness of the Border Gateway Protocol (BGP) convergence delay is the Minimum Route Advertisement Interval (MRAI). The MRAI is a timer with a default value of 30 seconds, which forces the BGP routers to wait for at least that amount of time before sending an advertisement for the same prefix. This process can delay important BGP advertisements. To date, there has been no specific value used by all the networks around the Internet. This paper aims to find the optimum value for the MRAI timer that maximally reduces the convergence time without increasing the number of advertisement messages. The optimal MRAI value founded by this paper reduced the convergence time by minimum 45%.

Peer-review under responsibility of organizing committee of The 2015 International Conference on Soft Computing and Software Engineering (SCSE 2015)

Keywords: BGP; Convergence Time; MRAI; Neuro Fuzzy Modelling

* Corresponding author. Tel. +44(0)1895 76152.
E-mail address: Eatedal.Alabdulkreem@brunel.ac.uk

Peer-review under responsibility of organizing committee of The 2015 International Conference on Soft Computing and Software Engineering (SCSE 2015)
doi:10.1016/j.procs.2015.08.436

1. Introduction

The Internet consists of numerous heterogeneous computer networks from all over the world [1]. Networks are typically referred to as domains or Autonomous Systems (AS), which are each controlled by single administrative entities. The Border Gateway Protocol (BGP) is the only routing protocol between different networks on the Internet [2]. However, more than one protocol exists within the boundary of each network or AS.

Convergence time is an important metric for routing protocols, such as the BGP. Convergence time is the period required by the BGP to reroute the packet after a routing change. Significant research on BGP convergence time using Internet measurements has revealed the problematic slowness BGP convergence [3]. One of the causes of the delay is a timer called the Minimum Route Advertisement Interval (MRAI). This timer was set to 30 seconds, forcing the BGP routers to wait at least 30 seconds before sending advertisement for the same prefixes [2].

In [4], the MRAI optimal value was dependent on a specific topology and could not be used as a general mechanism to reduce convergence time. However, [5] revealed that, while routing systems with homogeneous MRAI timer values have linear convergence times, diverse MRAI values can cause significant increases in both the convergence time and number of advertisement messages. In [6], the researchers proposed the ghost flushing technique, which reduces convergence times but does not eliminate all cases of unnecessary MRAI timer messages.

Given that the current value for this timer causes significant convergence time delays, we propose an optimum value for the MRAI timer, which improves the convergence process and does not harm the scalability. We examined a number of networks with different topologies to identify the optimum value for the MRAI timer.

The rest of the paper is organized as follows. In Section II, an overview of the BGP is given. The impact of the MRAI on the BGP is explained in Section III. The steps for finding the optimum value for the MRAI are presented in Section IV. Section V discusses the results, and Section VI presents the conclusion and outlines future work.

2. BGP Overview

BGP is based on the path vector routing mechanism. It involves exchanges of network reachability information between the BGP systems to find the most efficient path [7]. Before any information can be exchanged, a BGP session must be established between two BGP routers. The connection is reliable, as the session is supported by the TCP connection. The BGP routers through this connection exchange four different messages [2]:

- OPEN to start the session between any two peers
- UPDATE or advertisement to withdraw an unfeasible route or advertise a new feasible route
- NOTIFICATION to shut down the session whenever an error condition is detected
- KEEPALIVE to verify that the BGP peer is still available; this message must be periodically exchanged

First, a BGP router establishes connection with its neighbours (the other BGP routers) that directly communicate with it. It then downloads the entire routing table of each neighbouring router, finds the best path for each destination and saves that path in the routing table. If there was a routing change for any reason, the BGP router detects it. As a result, the BGP router sends updated messages, which either announces a new path or signifies the withdrawal of a path that no longer exists. As mentioned, convergence time is the time required to reroute packets after a routing change. The current convergence delay could stretch into more than hundreds of seconds, which could lead to high packet drop rates [8].

3. MRAI impact on the BGP

One primary cause of delayed BGP convergence is the MRAI, which is a timer with a default value of 30 seconds [9]. This timer forces the BGP router to wait for at least that amount of time before sending an advertisement for the same prefix. Though this avoids the storm of BGP advertisements, according to [10], it may unnecessarily delay important BGP advertisements.

Thirty seconds is not the optimal value for every network topology, as this time varies from network to network. Previous studies [11-13] have agreed on the significant influence of the MRAI on BGP convergence time, observing that reduced MRAI can limit convergence time. Since finding the optimal value for each network is extremely difficult [14], the optimal value for this timer is remains unknown. This has resulted in individual network operators decreasing MRAI timers with different values, thereby leading to greater timer setting diversity across the Internet. Routing convergence is substantially worse due to timer heterogeneity between networks [5].

4. Finding the optimum MRAI value

The steps to finding the MRAI optimum value are shown in Figure 1. First, to ensure the timer value would be suitable for any topology, different network graphs were generated (such networks should be similar to the topologies existing on the Internet). Three different topologies were generated with the help of the Boston University Representative Internet Topology gEnerator (BRITE) tool [15], and three more topologies were extracted from the tables at Route Views [17]. A scenario was then configured using the OPNET simulation tool for each topology. Next, different values of the MRAI for each scenario were used. The values for each MRAI began with 30 seconds and then decreased by 0.5 seconds. As a result, 60 different values for MRAI (from 0–30) were generated for use. The 60 MRAI values were used for each scenario. Six scenarios, each with 60 MRAI values, produced 360 different convergence time results and 360 update message results. The results were collected and used as a dataset to train the Neuro Fuzzy system. As a result, the system can provide two modules: one for convergence time and one for the number of update messages. This rendered the optimum value suitable for any topology on the Internet. To find the optimum MRAI value, the Particle Swarm Optimisation (PSO) algorithm was used [19]. This algorithm is designed to find the best value in any module. To find the optimum value for the MRAI timer, the PSO was applied to the two modules from the tanning system. A detailed explanation for each step is given next.

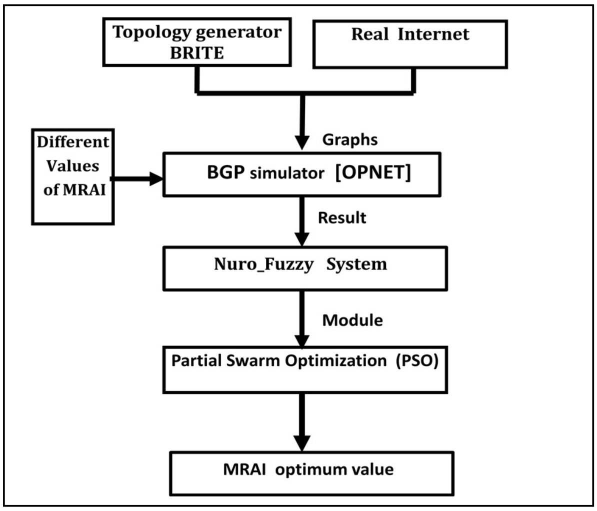

Fig.1 Steps to find the optimum value of MRAI

4.1 Network topology

The topology of the Internet as a group of networks is a series bearing on many management and performance issues. For the development and analysis of internetworking technologies, it is important to have a good example of the network's topological structure [15]. Three different graphs were generated using the BRITE tool, and three more were selected from the AS graphs extracted from the tables at Route Views [17].

The result was six different graphs with different topologies or parameters (Table 1). The parameters were number of nodes and node degree and diameter (the network topology key parameters are the network diameter and node degree). The number of nodes is the number AS in each graph. The average degree is the average number of adjacent nodes in the network. There are different paths between each two nodes. The diameter of the network is considered the maximum shortest path among them.

Table1. The parameters of the six graphs were used in the simulation

	BRITE			REAL INTERNET		
Number of Nodes	5	15	20	40	71	104
Average Degree	2	4.8	6	4.4	4.3	4.3
Diameter	2	3	4	5	6	5

4.2 BGP simulator.

A BGP scenario was simulated for each network topology or graph using OPNET Modeller, a powerful simulation tool with a wide variety of capabilities. Next, the MRAI timer values were changed for each scenario, starting with a default value of 30 seconds and then decreasing in 0.5-second intervals until reaching zero. The convergence time and number of update messages were saved each time a scenario's MRAI timer value was changed.

4.3 Neuro Fuzzy system.

In the field of artificial intelligence (AI), the Neuro Fuzzy system is the combination of fuzzy logic artificial and neural networks. There are two areas in the Neuro Fuzzy field: the Mamdani model and the Takagi-Sugeno-Kang (TSK) model. The Mamdani model is related to linguistic fuzzy modelling, which focuses on interpretability. The TSK model focuses on accuracy, and it is the model used in the present paper. TSK is a static, nonlinear system modelling tool [18] capable of generating accurate models but requiring good quality historical data to do so. To construct the mapping function, data-driven models should be analysed by historically obtained input-output pairs, which is why the resulting mapping function depends on historical data.

The Adaptive Network-based Fuzzy Inference System (ANFIS) is fuzzy architecture inspired by the TSK model, which performs identification of input-output mapping in the form of N input-output example sets. There are four inputs for each output. Input1 is the number of AS or nodes in the network. Input2 is the node degree or average degree. Input3 is the diameter, and Input4 comprises the different MRAI values between 30 and zero seconds that have been applied to each case. The output is the convergence time in seconds and the number of update messages for each case was recorded. The results for this step include two models: one for the convergence time and the other for the number of update messages Figure 2.

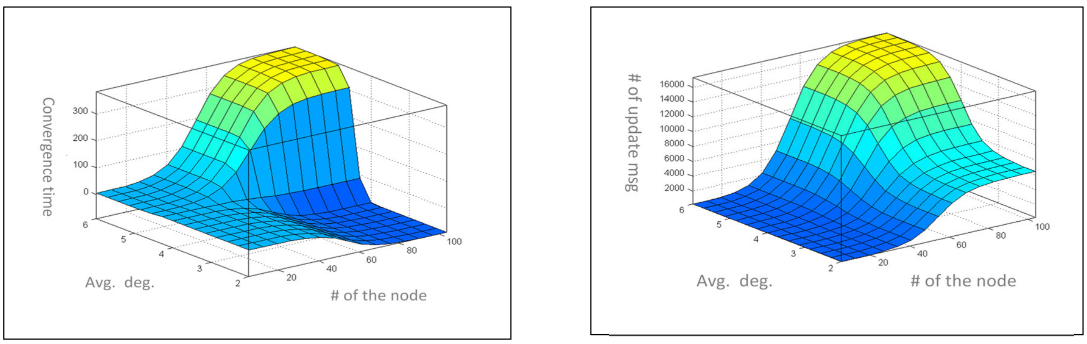

Fig. 2 (a) Module for the convergence time (b) module for the number of update messages

4.4 Particle Swarm Optimization (PSO) algorithm.

The PSO algorithm mimics the behaviour of flocking birds. Consider a scenario where every bird in a group is searching for one piece of food; the birds do not know where the food is, but they do know how far away the food is at any iteration. In this situation, the best strategy would be to follow the nearest bird to the food [19]. The same scenario is used to solve the optimisation problems in a PSO algorithm, where each solution, or particle, represents a bird within the search space. There are fitness values for all the particles to be optimised by the fitness function, and have velocity that direct the particles passage. The particles follow the current optimum particles to pass through the problem space.

First, when a group's random particle initialization is completed, the searches are updated for the optimal particles. With each iteration, every particle is updated based on the two best values. The first is the partial best value, which is the best fitness (solution) achieved so far. The other is the global best value, which is the best value obtained by any particle in the population. The best value is the particle that takes part of the population as its topological neighbours.

For each iteration, the best values for the timer are founded in each module. Then, a global value is found between all the founded values in all the iterations. In this work, the algorithm identified three seconds as the best value for the MRAI, as it maximally reduced the convergence time without increasing the number of advertisement messages.

5. Results and discussion

For each simulated experiment, a single node was disconnected by the failure event. As this caused a routing change in the network, the BGP router sent the updated message. In the original BGP scenario, the MRAI value was the default value; however, in the optimised BGP scenario, the MRAI value was the value recommended by this work. The convergence time and number of bits sent were recorded. The whole process was repeated, but the simulated experiment became a network with a different topology.

Figure 3 shows the result for the network with 20 nodes. The x-axis displays the time (20 minutes) starting at 15:32 and stopping at 15:52. The y-axis displays the number of seconds it took the BGP protocol to converge (convergence time). At 15:42 (after 10 minutes), one of the links failed and the BGP protocol needed to converge. As shown in the figure, the original BGP with the default MRAI value (the blue curve) reached 32 seconds before converging. However, the optimised BGP with the proposed MRAI value (red curve) reached 19 seconds; therefore, the convergence time was reduced in the optimised scenario where the value of the MRAI was three seconds.

Figure 4 shows the result with the x-axis displaying the time (20 minutes), starting at 15:32 and stopping at 15:52. The y-axis displayed the number of bits sent by the BGP protocol. When one of the links failed after 10 minutes (at 15:42) the BGP protocol needed to converge. The original BGP with the default MRAI value (blue curve) and the optimised BGP with the proposed MRAI value (red curve) sent the same amounts of bits.

Figure 5 shows the result for the 40-node network with the x-axis displaying the time. The scenario took 20 minutes, starting at 13:32 and stopping at 13:52. The y-axis displayed the number of seconds it took the BGP protocol to converge (convergence time). When one of the links failed after 10 minutes (at 13:42) the BGP protocol needed to converge. As shown in the figure, the original BGP with the default MRAI value (blue curve) reached 75 seconds before converging. The curve in red is the convergence time for the optimised BGP with the proposed MRAI value (26 seconds).

Figure 6 shows the result with the x-axis displaying the time (20 minutes), starting at 13:32 and stopping at 13:52. The y-axis displays the number of bits sent by the BGP protocol. When one of the links failed after 10 minutes (at 13:42) the BGP protocol needed to converge. The original BGP with the default MRAI value (blue curve) and the optimised BGP with the proposed MRAI value (red curve) sent the same amounts of bits.

Figure 7 shows the trend line for the convergence time. The x-axis displays the time (20 minutes), starting at 15:20 and stopping at 15:40. The y-axis displays the number of seconds it took the BGP protocol to converge. When one of the links failed after 10 minutes (at 15:30) the BGP protocol needed to converge. The figure shows that the original BGP with the default MRAI value (blue curve) reached 75 seconds before converging.

The convergence time for the optimized BGP with the proposed MRAI reached 19 seconds. As shown, the trend line for the original BGP (blue curve) increases faster than the trend line for the optimized BGP (red curve), which means that the convergence was faster with the optimized than the original; that is, the delay was reduced.

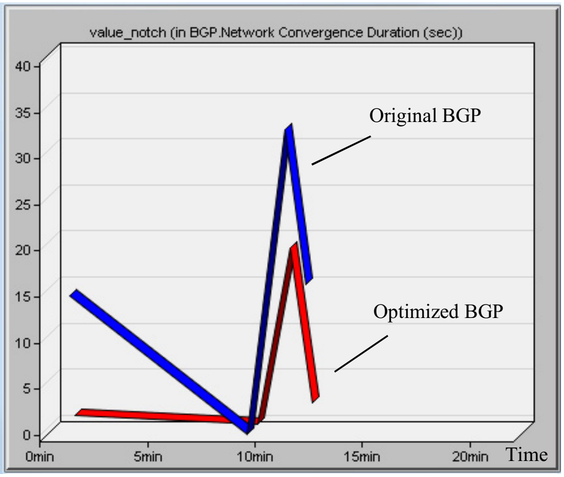

Fig.3 Comparison between the convergence duration with the default value and the optimum value recommended by this paper for the 20- node network

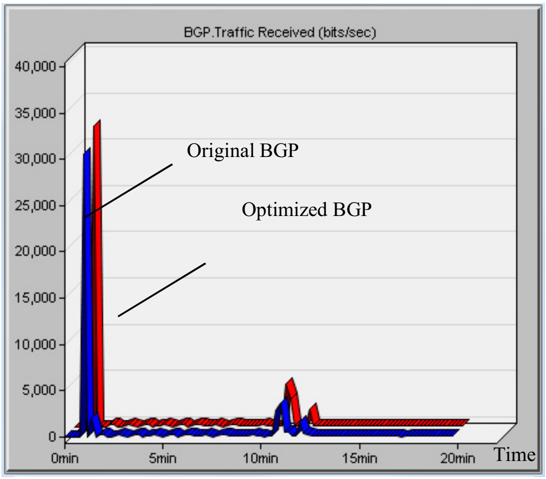

Fig. 4 Comparison between the numbers of update messages with the default BGP and the optimised BGP for the 20- node network

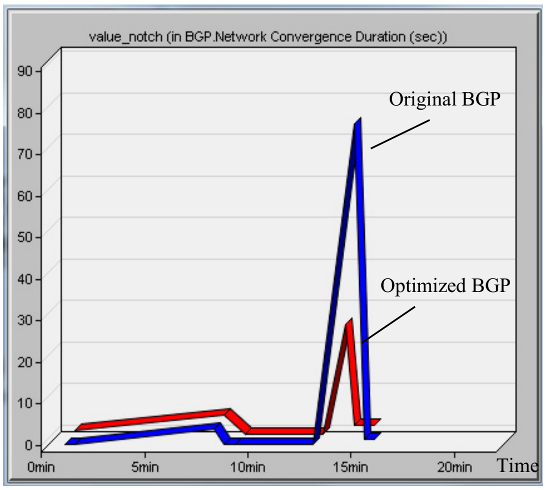

Fig.5 Comparison between the convergence time with the default and the optimum value recommended by this paper for the 40-node network

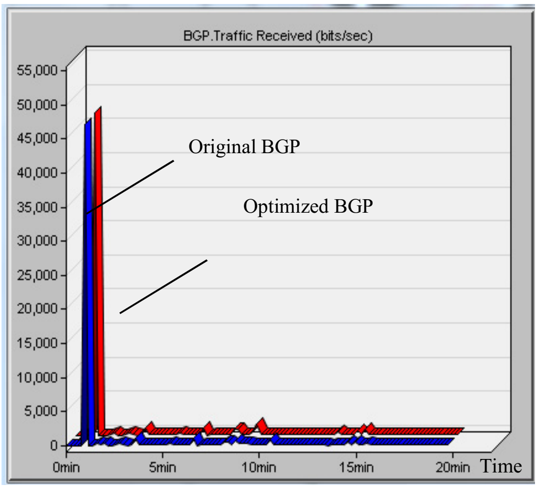

Fig.6 Comparison between the numbers of update messages with the default BGP and the optimised BGP for the 40-node network

Figure 8 shows the network convergence activity for three different scenarios: 1) the original BGP, 2) the ghost flushing (as it is the most cited idea) and 3) the optimised BGP based on this paper. The x-axis displays the time (20 minutes), starting at 15:20 and stopping at 15:40. The y-axis displays the number of seconds it took the BGP protocol to converge. When one of the links failed after 10 minutes (at 15:30) the BGP protocol needed to converge. As shown in the figure, the original BGP with the default MRAI value (blue curve) reached 71 seconds before converging. The BGP based on the ghost flushing idea (red curve) required 47 seconds to converge. The convergence time for the optimised BGP with the proposed MRAI value was 19 seconds.

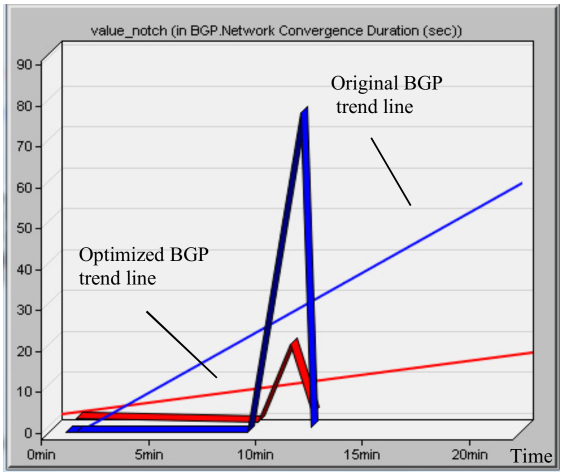

Fig.7 Comparison between the trends lines for the convergence time with the default and the optimum value recommended by this paper

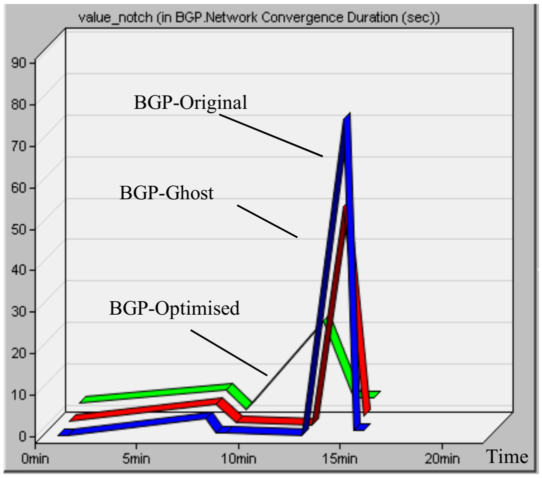

Fig.8 Comparison between the trends lines for the number of the update messages with the default BGP and the optimised BGP

6. Conclusion

This paper aimed to identify the optimum value for the MRAI timer that could maximally reduce convergence time without increasing the number of advertisement messages. We examined different networks with various configurations. For each network, a different scenario was configured using the OPNET tool. For each scenario, different MRAI values were applied, and the results for convergence time and number of update messages were collected. The results were used to train the Neuro Fuzzy system to provide two modules: one for the convergence time and another for the number of update messages. The PSO algorithm was applied to the modules to find the optimum value for the MRAI timer.

It is concluded that the optimum value for the MRAI is 3 seconds. With this value, the convergence time was reduced by minimum 45% as seen in figures 3 and 5. The number of update messages was the same figures 4 and 6. This result proves the great sensitivity of the BGP to minor modifications of some of its parameters.

Future work will be related to path exploration, as there is a need to improve the speed of this process, which will certainly reduce the time required for the BGP convergence.

References

1. Huaming Guo, Wei Su, Hongke Zh, and Sy-yen Kuo, On the convergence condition and convergence time of BGP. Computer Communications, *Elsevier*, 2011, vol. 34, no. 2, pp. 192-199.
2. Rekhter Y, and Li T, A Border Gateway Protocol 4 (BGP-4), Rfc 1771, SRI Network Information Centre, 1995.
3. Labovitz Cr, Ahuja Ab, Bose Ab, and Jahanian Fa, Delayed Internet Routing Convergence, IEEE/ACM Transactions on Networking, 200, vol. 9, no. 3, pp. 293-306.
4. Griffin Ti, and Presmore Bj, An Experimental Analysis of BGP Convergence, Network Protocol, *IEEE*, 2001, pp. 53-61,.
5. Fabrikant A, Syed U, and Rexford J, There's something about MRAI: timing diversity can exponentially worsen BGP convergence. INFOCOM, *IEEE*, 2011, pp. 2975-2983.
6. Afek Ye, Bremler-Barr An, Schwarz Sh, Improved BGP convergence via ghost flushing, Selected Areas in Communications, *IEEE* Journal, 2004, vol. 22, no.10, pp. 1933 - 1948.
7. Yannuzzi M., and Masip-Bruin X., Open Issues in Interdomain Routing: A Survey, Network, *IEEE*, 2005, vol. 19, no. 6, pp. 49-56.
8. Shaoo Am, Kant Kr, and Mohapatra Pr, Bgp convergence delay after simultaneous router failures: characterization and solutions.Computer Communications, *Elsevier*, 2009, vol. 32, no. 10, pp. 1207-1218.
9. Qio Ji, Hao Ru, and Li Xi, The Optimal Rate_-Limiting Timer of BGP for Routing Convergence. The institute of electronics, IEICE TRANS. COMMUN, 2005, vol. 88, no. 4, pp. 1338-1346.
10. Suchara Ma, Fabrikant Al, and Rexford Je, BGP safety with spurious Updates. INFCOM, *IEEE*, 2011, pp. 2966-2974.
11. W. Wenhua, S. Qinguo, S. Yu, and H. Chunyong, "New MRAI Setup Mechanism for Enhancing BGP Route Convergence". ICAT, pp. 1-7, 2009.
12. Wang Be, The Research of BGP convergence time, IPCCC, IEEE, 2011, ITAIC, pp. 354-357.
13. Deshpande S., and Sikdar B., On the Impact of Route Processing and MRAI Timers on BGP Convergence Times. GLOBCOME, IEEE, 2004 pp. 1147-1151.
14. Laskovic N., and Trajkovic L., BGP with an adaptive Minimum Route Advertisement Interval. IPCCC, IEEE, 2006, pp. 135-142.
15. Calvert K., Georgia L., Doar M., and Zegura B., Modeling Internet topology. Communications Magazine, IEEE, 1997, vol. 35, no. 6, pp. 160-163.
16. Brite: Universal topology generator . Available at: www.cs.bu.edu/brite/, 2001.
17. Route views project pages, available at: www.routeviews.org, 1994.
18. Ashuton Te, Marina Mi, Knowledge-based parameter identification of TSK fuzzy models, Applied Soft Computing, *ELSEVIER*, 2010, vol. 10, no. 2, pp. 481-489.
19. Hongyan Cu, Jian Li, Xiang Li, and Yunlogng Ca, Particle Swarm Optimization for Multi-constrained Routing in Telecommunication Networks. International Journal of Computer Network and Information Security, 2011, vol. 4, pp.10-17.