

Tomas Macha – Radko Krkos – Vit Novotny \*

## PROPOSAL OF LOAD AWARE ROUTING FOR OSPF ROUTING PROTOCOL

*This paper deals with route cost assignment using one of the most commonly used routing protocols - Open Shortest Path First (OSPF). The limitation of this protocol is that the algorithm does not take actual link load into consideration. This paper proposes using an alternate costing strategy to avoid this problem. The alternate strategy uses an integer value representing the load of a link as the additive component of metric. Since the link load is a dynamic variable, Exponential Weighted Moving Average is used to counter quick changes. This paper proposes a novel approach to calculate OSPF metrics, aimed to better distribute the load in network.*

**Keywords:** OSPF, routing, metric, load.

### 1. Introduction

The Internet or internetworking communication systems are based on universal network-level interconnections. These interconnections enable delivery of data units from their source to their desired destination. Internetworking also deals with the complexity of various underlying communication technologies that today's heterogeneous networks comprise of. The IP (Internet Protocol) plays a central role in the Internet architecture, as it is responsible for appropriate routing of packets in the Internet [1].

Routing is one of the main properties of the Internet and enables establishment of robust and efficient networks. The data unit is sent directly to the destination, if the destination is on the same subnetwork as the source is. Otherwise, the data unit is sequentially being sent to the routing devices in the network that make decisions about directions so the data unit reaches desired destination [1]. Such devices, which interconnect subnet works and pass packets from one to the other, are called routers. The behaviour of routers depends on a routing protocol. A routing protocol defines a set of rules used by a router for communication with neighbouring routers. Routers between a source node and a destination node must choose and maintain paths to be used. The router must create routing tables and compute the best route to use [2].

Since OSPF (Open Shortest Path First) currently represents the most widely used routing protocol, any valuable improvement to keep pace with the rapidly changing Internet environment would be greatly appreciated. The limitation of this protocol is that its link cost calculation algorithm does not take actual link load into consideration. If there are two links with the same bandwidth and the link utilization of the first is very low and for the second link

very high, OSPF assigns both links the same metric [2]. An explanatory example is in Fig. 1. This network consists of six routers; most of the routers are connected with 100 Mbps Ethernet links and their cost corresponds to a metric of 10, see Table 1. Technology of 10 Mbps Ethernet is used to connect router E with routers B and I. Metric assigned to links between these routers is 10. In this network, User 1 and User 3 communicate with the server. Based on path metrics calculated as a sum of metrics of all links on the path from the source to the destination the entire data traffic passes the link between routers D and E. With increasing users' demands on the server this link may become congested. Situation becomes even worse when User 2 also starts to communicate with the server. Congestions cause service performance degradation and in extreme cases even service disruption.

This paper brings a novel mechanism that also includes link load into link cost computation to improve routing in OSPF networks. This mechanism considers link load as an additive parameter for a final metric, which solves the problem of absence of traffic awareness and is aimed to balance the load among network links and routers, which in general decreases latency for transported data and overall improves network performance. This paper is a summary of on-going research presented in [3] and [4].

New Cost Adaptive OSPF (CA-OSPF) is proposed in [5]. However, this solution is not applicable for heavily loaded networks where the CA-OSPF cannot improve the network performance [5]. The recent works related to routing algorithms improvements [6], [7], [8] and [9] can be used to choose a path with specific bandwidth requirements. An implementation of QoS (Quality of Service) routing extensions for OSPF is proposed in [10]. Also many balance heuristic techniques were proposed to avoid con-

\* Tomas Macha, Radko Krkos, Vit Novotny

Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic,  
E-mail: tomas.macha@phd.feec.vutbr.cz

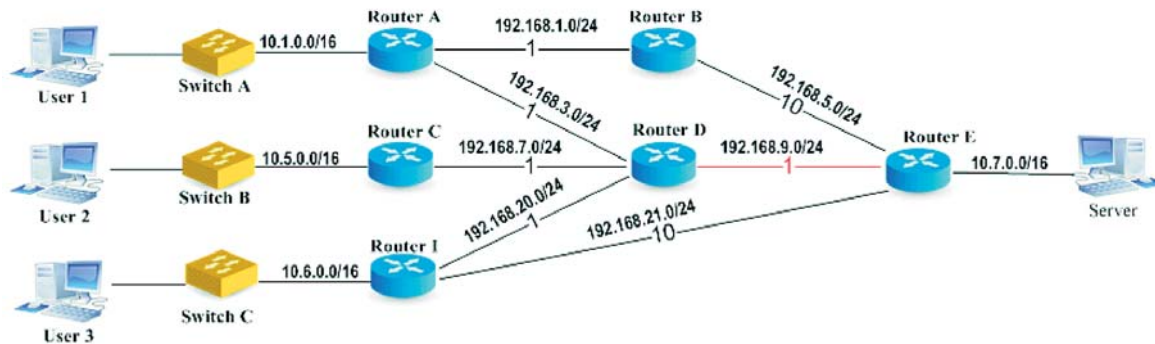


Fig. 1 OSPF network example

gestion and utilize links with low load [11], [12]. However all of these solutions focus only on the case of full link congestion and there is no rerouting of traffic until then. This load balancing gets to a stabilized state after a number of reached link congestions and so the convergence time is very long. The equilibrium is again broken when the volume of traffic changes and convergence process has to start afresh.

The main goal of this paper is to propose a novel and more efficient mechanism to improve routing in data networks using OSPF protocol. To achieve this purpose, a couple of extensions of current routing methods will be examined.

A) OSPF metric

Dijkstra’s algorithm is used by OSPF for calculating the best path through complex networks. The shortest-path calculation is executed with a full knowledge of the topology.

RFC document [13] does not specify what the link metric should be. Most of the routers calculate the metric based on the configured bandwidth. The metric is then inversely proportional to the bandwidth; higher bandwidth means lower metric.

Metric of a path is the sum of the individual link metrics on the path to the destination. The formula used to calculate the metric of one interface is:

$$metric = \frac{reference\ bandwidth}{bandwidth} \tag{1}$$

Different manufacturers use different reference bandwidth, however, common reference bandwidth is 100 Mbps (100,000,000 bps). The reference bandwidth has to be equal for all routers in the same routing area. Only one metric can be assigned per interface. According to formula (1), the metric for Ethernet technology is 10 (10<sup>8</sup>/10 000 000 bps). OSPF metrics for frequently used link types and based on formula (1) are summarized in Table 1. The interface metric is an unsigned 16-bit number; the total metric of a path is a 24-bit unsigned integer, as can be seen from the format of LSA-3, LSA-4, LSA-5, and LSA-7. Metrics are rounded down to the nearest integer. If its total metric overflows the 24-bit value, the path becomes unusable.

If the link bandwidth is 100 Mbps or higher, then based on formula (1) and commonly used reference bandwidth, the link metric will be 1, which is not optimal as there is no distinction between 100 Mbps and faster links. To solve this problem, it is possible to set new reference bandwidth for all routers in the entire network, using a router configuration command.

Default OSPF metrics Table 1

Technology	10 <sup>8</sup> /bps	Metric
Gigabit Ethernet	10 <sup>8</sup> /1 000 000 000 bps	1
Fast Ethernet	10 <sup>8</sup> /100 000 000 bps	1
Token Ring (16 Mbps)	10 <sup>8</sup> /16 000 000 bps	6
Ethernet (100 Mbps)	10 <sup>8</sup> /10 000 000 bps	10
E1	10 <sup>8</sup> /2 048 000 bps	48
T1	10 <sup>8</sup> /1 544 000 bps	64
64 kbps link	10 <sup>8</sup> /64 000 bps	1562
56 kbps link	10 <sup>8</sup> /56 000 bps	1785
9.6 kbps link	10 <sup>8</sup> /9 600 bps	10 416

2. Proposed method

Dijkstra’s algorithm solves the shortest path problem by finding a path of minimal total cost between the nodes based on link metrics. Link metrics do not include the link load. Solution is that an integer value is added to the load of a link as an additive component of bandwidth based metric. This method builds on common metric calculations used by OSPF protocol. The computation of a novel metric for the novel method derives from formula (1):

$$novel\ metric = \frac{reference\ bandwidth}{bandwidth} + load - (metric_{next\ technology} - metric) \tag{2}$$

The component metric is a link metric computed according to the formula (1), metric<sub>next technology</sub> is calculated using the same

formula, but bandwidth of the next slower technology according to Table 1 is used. This means that if the link load approaches full utilization, the next slower link is used as an alternative path because the metric of the faster link is degraded linearly with load to match metric of 0% utilized link of next slower technology when the link is fully loaded. The situation is described in Fig. 2, where the metric for 100 Mbps link grows with link load to finally match the metric of Token Ring link.

Along with the method, we propose several improvements. The first one includes a moving average because of its averaging characteristics and the ability to smooth fluctuations. The second one considers real maximum link load utilization. The third one calculates link metric using highest link bandwidth present in OSPF area as reference bandwidth. And the fourth improvement consists of using next slower technology present in current topology instead of next slower technology according to Table 1, in link load impact calculation.

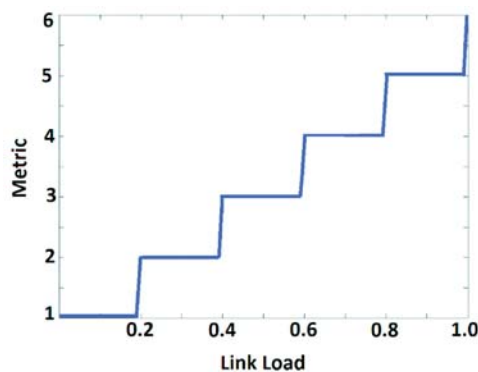


Fig. 2 Metric degradation of 100 Mbps link based on load

A change in the network topology is represented as a change in one or more of the LSAs. LSAs flooding process can noticeably degrade performance. The change causes new LSA flooding and calculation of new routing tables. However, even this amount of necessary traffic is undesirable.

*A) Exponential weighted moving average application*

Data collected in the time order can be averaged over several samples. Moving averages are often used in time series data analysis. Moving averages can be useful for measuring the changes in a trend, smoothing fluctuations and as a forecast for the next period. In this paper, the used raw data are link load observations of a router. Link load values represent data from which the moving average will be computed.

Exponential Weighted Moving Average (EWMA) [13] has become popular process-monitoring tool in a process-control field. Due to EWMA's robustness and ability to monitor a dynamic process with memory and drift, this approach is adopted for proposed method.

EWMA, sometimes also called Exponential Moving Average (EMA), provides a higher level of accuracy when compared with simple or weighted moving average. EWMA uses weight factors that decrease exponentially, which brings much more importance to the recent observation, but all the previous observations are included. Let  $c_n$  be an observation at time  $n$ , the explicit formulation of EWMA is:

$$A(n) = (1 - \lambda)^k A(n - k) + \lambda \sum_{i=0}^{k-1} (1 - \lambda)^i c_{n-i}, \quad (3)$$

where  $A(n)$  is the current EWMA average,  $c_i$  is the value of link load at time  $i$ ,  $k$  is the number of previous values used in calculation or the averaging depth and  $A(n - k)$  is the EWMA average of  $k$  samples before. The starting value  $c_0$  equals zero or is generally being set to the mean of former observations. The effect of the starting constant  $c_0$  decreases over time. The formula relies on an effective period for the exponential moving average called smoothing factor  $\lambda$ ,  $0 < \lambda \leq 1$ ,  $\lambda = 2/(n+1)$ , where  $n$  is the period of the moving average. The parameter  $\lambda$  determines the rate at which previous observations influence the calculation of EWMA. A value of  $\lambda = 1$  implies that only the most recent observation influences the EWMA, thus higher value of  $\lambda$  gives more weight to recent data. The weight of each older observation is then decreased by a factor of  $(1 - \lambda)$  [13]. EWMA can be computed recursively as [14]:

$$A(n) = \lambda c_n + (1 - \lambda)A(n-1). \quad (4)$$

The recursive form of the EWMA calculation simplifies the formula (3) and decreases computing complexity. The recursive fashion (4) requires only two pieces of information to be processed.

Application of the EWMA technique is demonstrated using an example depicted in Fig. 3. The blue line stands for raw link load data, representative of current bandwidth usage. The red line stands for EWMA of the raw data. Formula (4) was used for EWMA computation. It can be seen that the series is smoothed,

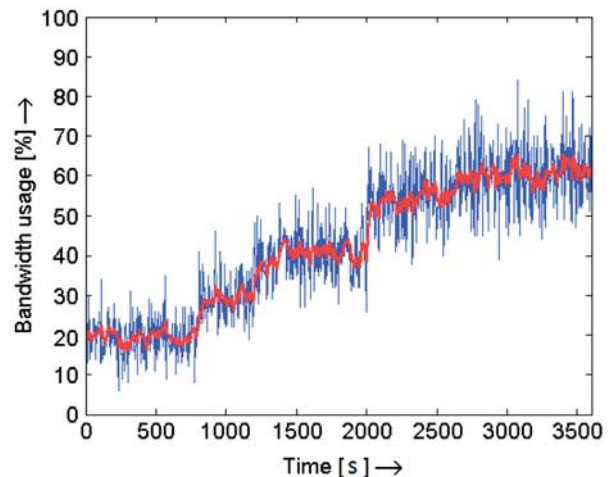


Fig. 3 Exponential weighted moving average example

with much lower variance. The decrease of the weights is an exponential function of the weighting factor  $\lambda$ . If the value of  $\lambda$  is close to zero, the moving average at any time  $n$  is significantly influenced by older values.

As mentioned before, the first improvement of the novel method is the application of EWMA technique onto the novel metric (3). The resulting formula can be expressed as:

$$novel\ metric_1 = metric + EWMA(load) \cdot (metric_{next\ technology} - metric) \tag{5}$$

Figure 4 describes the impact of sampling frequency on EWMA convergence time. There are four similar graphs where the blue line stands for link load and the red line stands for EWMA of load. If the sampling is done every 30 minutes, it can be seen from the first graph that the convergence time is about 2 hours, which is not feasible. On the other hand, if the sampling is done every minute, the EWMA curve copies the load curve. So in this case, parameter  $\lambda$  should be increased.

*B) Maximum link load utilization*

This improvement considers the real maximum load of the links. This method is advantageous when used in heterogeneous

networks or in networks supported by routers of different performance where it provides better scalability and network performance.

The best way how to explain this method is to show an example. For assembling the topology in Fig. 1, Linksys routers WRT54GL were used. The manufacturer claims 100 Mbps link bandwidth. However, according to the real measurement, the maximum bandwidth is only 32 Mbps. If 32 Mbps is the real maximum utilization instead 100 Mbps, the routers are not able to reach link utilization of 100% and the network performance is correspondingly lower. This problem prevails especially with less powerful routers. Therefore, as a solution, a maximum relative load parameter is applied. The resulting formula can be expressed as:

$$novel\ metric_2 = metric + load \cdot \left( \frac{metric_{next\ technology} - metric}{max\ relative\ load} \right) \tag{6}$$

When the network infrastructure is homogenous, the maximum relative load is approximately equal for all the links. But this approach still enables better scaling of metrics with link load. Maximum relative load is most useful when it is quite different from 100%, so the load effect on metric is spread evenly in the range of real bandwidth, which the link is capable of.

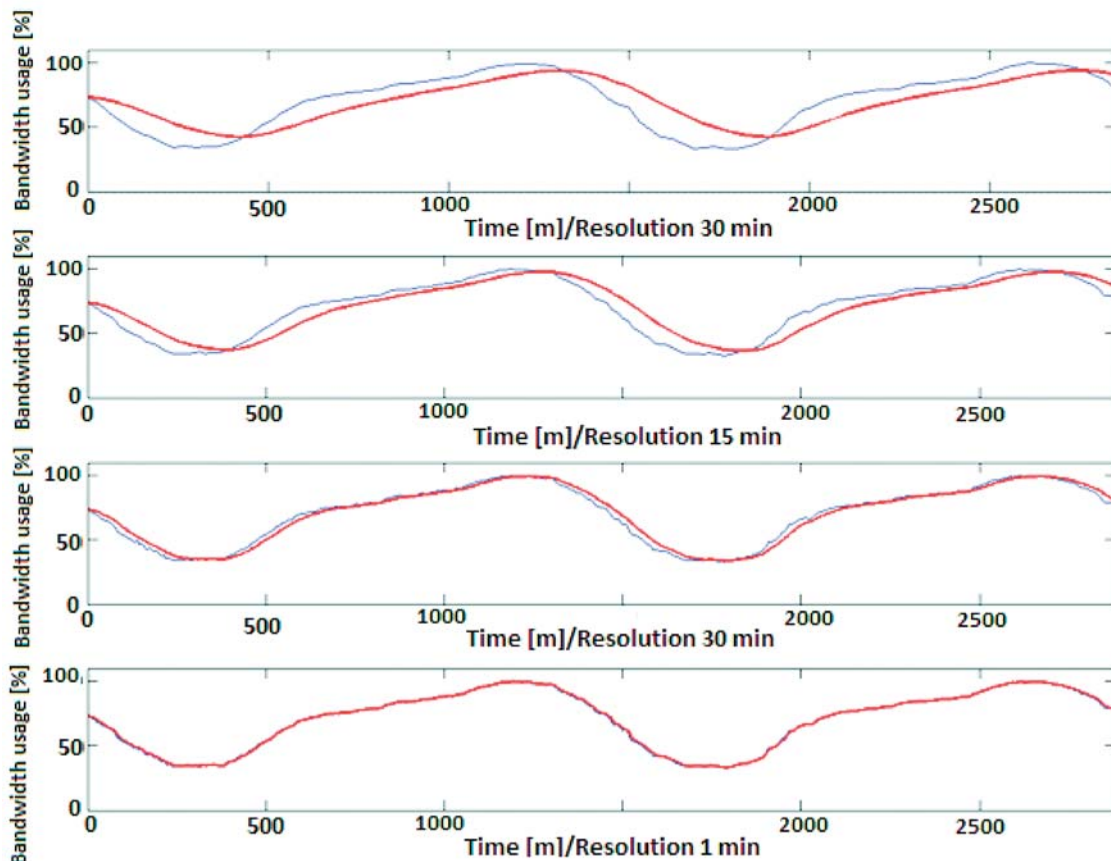


Fig. 4 Impact of sampling frequency on EWMA convergence time

### C) Reference bandwidth value optimization

Original OSPF metric is calculated using reference bandwidth set to 108 [15]. This value is too small for today's link bandwidths and it causes that 100 Mbps, 1 Gbps and faster links have equal metric, see Table 1. Reference bandwidth value can be changed by network administrator, but when altered, it has to be manually set to the same value on all routers in order to shortest path selection to work properly. The solution would be an automatic fastest link seeking. This can be done in two ways. First possibility is to specify and distribute information about fastest link before the first topology creation. This is easier and often should be enough, because link bandwidths rarely change during operation. But it requires new LSA type or alteration of existing one to spread the highest encountered link bandwidth value. Second solution is to resolve it during topology creation. This is more complex, because the metrics have to be calculated before the maximum link bandwidth in network is known. Routers can calculate metrics using largest value already known to them, but also have to transmit used reference bandwidth, so other routers can correct their values if used reference values are mismatched. The advantage is possibly lower convergence time, especially for larger topologies.

Extension to this technique is an automatic detection of all present link bandwidths. This enables, in cooperation with improvement in section D), more accurate spreading of link load effect on metric because non present link bandwidths are left out and only the existing ones are considered when calculating next line speed metric, as shown by equation (6).

### D) Matching technologies present in topology

Based on the formula (3), link load impact is multiplied by metric difference of current technology and the next slower technology according to Table 1. This is often sufficient, as multiple technologies used in one network mostly differ by one level. In some situations, when this is not the case, e.g. 100 Mbps lines are used for headquarters networking, but remote sites are interconnected by a leased line E1 link, this approach would yield no benefit.

The problem is, that in Table 1, there is another technology with metric value between the two mentioned. So in full load of 100 Mbps link, the metric for this link would be degraded to match the next slower technology, 10 Mbps link. The metric of an E1 link would still be much higher and therefore the link would not be used as an alternative. Solution is using the next technology present in current topology when calculating  $metric_{next\ technology}$ . This requires modifications to the OSPF topology discovery procedure to ensure that all routers know all technologies present in topology and therefore can change the link metrics based on load in a consistent and efficient manner. Such procedure is a question for further research beyond this paper.

## References

- [1] CHAO, JONATHAN H., BIN LIU: *High Performance Switches and Routers*. Hoboken, New Jersey: John Wiley & Sons, 2007, 613 p., ISBN 978-0-470-05367-6.

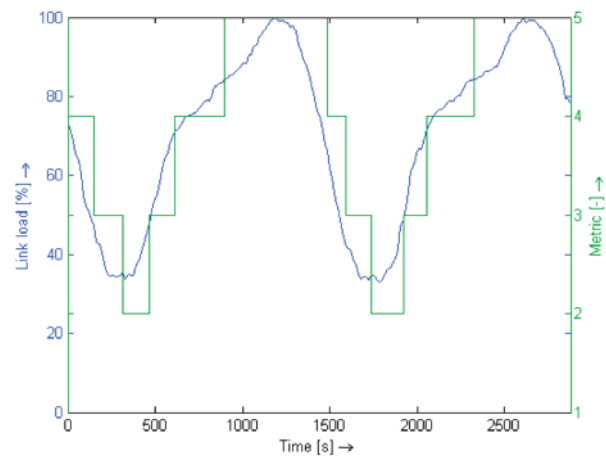


Fig. 5 Metric evolution based on link load for 100 Mbps link

## 3. Conclusion and future work

Open Shortest Path First is the most commonly used Internet routing protocol. The limitation of this protocol is that the algorithm does not take link load into consideration. This paper suggests using an alternate costing strategy to counter this problem. A novel method of OSPF metric calculation was proposed. An additive metric that considers current load of the link is added to the default metric. Also four methods to improve the OSPF metric calculation were proposed. All the proposals are summarized in the paper. Fig. 5 shows the metric evolution based on link load for 100 Mbps link. It can be seen that the novel method is functional and adapts the metric according to the link load. Proposed method may cause unwished characteristic, as frequent routing table changes can degrade the service and create downtimes. Also as IP protocol routing is dynamic, two-stage oscillation effect can occur. In this case, the traffic switches between two paths back and forth. Further work will be focused on solutions to these problems. As discussed in [3], further simulations will be done in ns-3 with direct code execution support, running altered Quagga OSPF daemon on the simulated nodes. Next, this altered OSPF implementation will be deployed on real devices to confirm the simulated results and carry out performance measurements.

### Acknowledgment

The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

- [2] SPORTACK, Mark A.: *IP Routing Fundamentals*. Indianapolis: Cisco Systems, 1999, xvii, 510 p. ISBN 15-787-0071-X.
- [3] MACHA, T., KRKOS, R., NOVOTNY, V.: OSPF Alternate Costing Strategy. In 2012 Intern. Conference on Telecommunication Systems, Modeling and Analysis (ICTSM2012), Prague: Czech Technical University in Prague, pp. 136–140, 2012, ISBN: 978-0-9820958-6-7.
- [4] MACHA, T., KRKOS, R.: *A Novel Approach to OSPF Metric Calculation*. Proc. of Research in Telecommunication Technologies, 14th Intern. Conference, pp. 85–91, 2012, ISBN: 978-80-554-0570-4.
- [5] ZHOU, H., PAN, J., SHEN, P.: *Cost Adaptive OSPF*. Proc. of the Fifth Intern. Conference on Computational Intelligence and Multimedia Applications. 2003.
- [6] GUERIN, R., ORDA, A., WILLIAMS, D.: *QoS Routing Mechanisms and OSPF Extensions*. Proc. of IEEE GLOBECOM, Phoenix. AZ. pp. 1903–1908, November 1997.
- [7] WANG, Z., CROWCROFT, J.: Quality-of-service Routing for Supporting Multimedia Applications. *IEEE J. on Selected Areas of Communication*, vol. 14, pp. 1288–1234, September 1996.
- [8] MA, Q., STEENKISTE, P., ZHANG, H.: *Routing High-bandwidth Traffic in Max-min Fair Share Network*. Proc. of ACM SIGCOMM. Stanford. CA. pp. 206–217. August 1996.
- [9] WANG, J., NAHRSTEDT, K.: *Hop-by-hop Routing Algorithms for Premium Class Traffic in DiffServ Networks*. Proc. of IEEE INFOCOM, pp. 705–714, June 2002.
- [10] APOSTOLOPOULOS, G., KAMAT, S. GUERIN, R.: *Implementation and Performance Measurements of QoS Routing Extensions to OSPF*. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 680–688, March 1999.
- [11] FORTZ, B., THORUP, M.: Increasing Internet Capacity Using Local Search. *Computational Optimization and Applications*, pp. 13–48, 2004.
- [12] FORTZ, B., REXFORD, J., THORUP, M.: Traffic Engineering with Traditional IP Routing Protocols. *IEEE Communication Magazine*, vol. 40, pp. 118–124, October 2002.
- [13] HUNTER, J. S.: The Exponential Weighted Moving Average. *J. of Quality Technology*.
- [14] ROBERTS, S. W.: *Control Chart Tests Based on Geometric Moving Averages*. Technometrics.
- [15] MOY, J.: OSPF Version 2. RFC 2328. *Internet Engineering Task Force*. April 1998.