International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume6(2014) pp.474–483

A Methodology for Traffic Shaping Optimization in Next Generation Networks

Ulisses Cavalca[1], Caio Mesquita[1], Adriano C. M. Pereira[2] and Eduardo G. Carrano[3]

[1]Centro Federal de Educação Tecnológica de Minas Gerais, Department of Computing,
Av Amazonas, 7675, Belo Horizonte, Brazil
ulisses@ccc.cefetmg.br; *caioboninho@gmail.com*

[2]Universidade Federal de Minas Gerais, Department of Computer Science,
Av Antônio Carlos, 6627, Belo Horizonte, Brazil
adrianoc@dcc.ufmg.br

[3]Universidade Federal de Minas Gerais, Department of Electrical Engineering,
Av Antônio Carlos, 6627, Belo Horizonte, Brazil
egcarrano@ufmg.br

***Abstract*: In recent years, computer networks have been characterized by heterogeneous traffic and dynamic management of different kinds of services. The network requirements have increased within time and, since bandwidth is limited, it becomes necessary to employ optimization procedures in order to make the network able to operate in its full capacity. Traffic shaping mechanisms implement Quality of Service (QoS) concepts to ensure acceptable service levels. This paper shows preliminary results of an empirical approach dedicated to perform traffic shaping optimization. Such an approach is based on throughput and packet loss optimization through multiobjective genetic algorithms.**

Keywords: QoS optimization, traffic shaping, genetic algorithm, multiobjective optimization,; Next Generation Networks.

I. Introduction

In 2004, International Telecommunication Union (ITU-T) adopted the Next Generation Network (NGN) concepts, which define capabilities to support, to expand and to manage network services and demands [1]. It provides acceptable services levels for all kinds of services, as multimedia demands, unicast, broadcast and multicast messages, real-time and non-real-time services, simple data transfer services [2], and others. Quality of Service (QoS) is one of the NGN principles, and it aims to ensure robustness, scalability, reliability, and availability in network resources. Some efforts like IntServ [3], DiffServ [4], and architectural framework QoS [5, 6] provide practical implementation of QoS in network environments.

Bandwidth is a finite resource, although some corporate users may believe it is not. Not rarely, bandwidth contracts are oversized in comparison with the real workload, what leads to unnecessary expenses. In many practical cases, it is possible to comply with the increase of users and services demands with the simple optimization of the existing network resources, without additional investments on bandwidth and infrastructure.

Theoretically, networks that do not reach full usage do not require traffic shaping. However, traffic shaping solutions must be robust enough to handle with anomalous behaviors and unforeseen network demands. This view proposes a balance between the real traffic load and the required bandwidth. In other words, the network environment must be configured with a ceil rate that assures operation without congestions, and traffic shaping can optimize throughput and packet loss for all kind of services.

In the same view, congested networks represent a challenge on management network and traffic shaping performance. In some cases, Quality of Service solutions do not solve transmission delay and packet loss. Although efforts and actions have been taken in the context of network management, there are scenarios in which it is necessary to expand bandwidth. This occurs when the network demands are much higher than the available resources.

Optimization methods, in mono or multi-objective approaches, have been widely employed to provide efficient resource allocation. Previous works suggest that computational intelligence techniques, such as Neural Network and Evolutionary Computation, can be used to handle with network resource management.

This work is intended to provide optimal management of network resources. In order to deal with this challenge, it is proposed an approach for performing fine tuning of a traffic shaping mechanism. This approach is based on multiobjective evolutionary optimization. The traffic shaping mechanism implements Quality of Service (QoS) concepts and it uses offered throughput maximization (to_i) and packet loss minimization (pl_i) as objectives, under traffic demand td_i.

Such a methodology is based on 4 steps, as follows: *Characterization step*, whose identify network profiles and demands; *Emulation step*, to reproduce major technical features of network operational environment and to generate datasets that describe the operation of the traffic shaping mechanism;

Identification step, that builds neural networks that mimic the traffic shaping mechanism interaction with the network; *Optimization step*, in which a multiobjective genetic algorithm is used to obtain optimal configurations of assured rate (ar_i) and limit rate (c_i); and *Analysis step*, which consists on the validation of a solution on the emulation environment, before real setup implementation.

Furthermore, this methodology is general, since it can be applied to any kind of traffic shaping mechanism that considers assured rate, limit rate and priority as adjustable parameters. As a case study, this work considers a GNU/Linux server environment, and the Hierarchical Token Bucket (HTB) [7] packet scheduler as traffic shaping mechanism. Real data from Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG) is used.

This paper is an extension of a work previously presented at 8th International Conference on Next Generation Web Service Practices [8]. It describes the proposed methodology for traffic shaping optimization, some partial results are presented.

The remainder of this paper is organized as follows. In Section II, some works about Next Generation Network, QoS attributes optimization problem, admission control optimization, and TCP behavior dynamic model are exposed. In Section III, it is discussed the proposed methodology, with focus in traffic characterization and emulation, model identification, optimization and analysis. Finally, some partial results are presented in section IV.

II. Related work

Next Generation Network (NGN) includes the concept of *triple play* that considers voice, audio, and data in the same transmission environment. Packet-based transfer, converged services, and Quality of Service (QoS) capabilities characterize NGN in the context of new demands on IP networks [9]. Furthermore, future practices in network management should take into account self management devices, dynamic performance, and solutions based on bio-inspired systems [10].

International Telecommunication Union (ITU-T) series recommendations define a set of technical and non-technical terms related to Quality of Service. E.800 recommendation lists, in the context of QoS characteristics, speed, accuracy, dependability, availability, and reliability as attributes of network performance [11, 12]. Therefore, these QoS parameters must be considered and, in some cases, guaranteed for new demands in IP networks as Web streaming, videoconferencing, and voice over IP (VoIP).

Additionally, QoS solutions are classified in IntServ (Integrated Services) and DiffServ (Differentiated Services) implementation models. IntServ QoS mechanism, described by RFC 1633 [3], ensures individual resource reservation in network devices for particular requests. DiffServ (RFC 2475 [4]) provides QoS for traffic classes, which consists on a scalable and robust network architecture. Additionally, ITU-T Y.1291 [6] recommendation discusses a framework for support QoS in packet-based network.

RFC 2212 [5] discusses about Specification of Guaranteed Quality of Service. This document describes approaches to ensure bandwidth and delay connections end-to-end and it defines requisites for network elements.

Canfora *et al* [13] introduce a QoS attribute optimization applicable to a large number of network scenarios. The proposed fitness function combines the following QoS attributes: cost, response time, availability and reliability. Integer programming (IP) and genetic algorithms are employed to optimize the problem, that is NP-hard. The genetic algorithms has shown good performance when compared to IP.

Yao and Chen [14] deal with QoS optimization considering price, response time, availability, throughput, and reputation as attributes. A multiobjective approach is employed. The main contribution is the application of the *Nondominated Sorting Genetic Algorithm-II* (NSGA-II) [15] to perform optimization.

Rosemberg *et al* [16] also model the QoS attribute optimization as an optimization problem, which is NP-hard. Three metaheuristics that can be applied to the QoS problem are proposed: genetic algorithms, tabu search, and simulated annealing. Moreover, the QoS parameters in this work are distributed in two categories: operational and business-related attributes.

Pop *et al* [17] optimize some QoS attributes, such as response time, reliability, availability, and cost. A differential evolution algorithm is proposed for dealing with the problem. Suciu *et al* [18] compare *Strength Pareto Evolutionary Algorithm* (SPEA) and NSGA-II in multiobjective QoS optimization. This work includes response time, rating, availability, and cost as QoS attributes.

In practice, traffic shaping intends to ensure Quality of Service and performance in hybrid traffic network or congested environment. Jha and Hassan [19] suggest average-rate, burst-size and peak-rate shaping in order to accomplish traffic shaping. Moreover, IP networks consider throughput, delays, packet loss, latency and jitter as traffic shaping metrics. Chu and Lea [20] propose a heuristic search algorithm to solve the Minimum Congestion Ratio problem, in the context of hop-by-hop routing. The work concerns about DiffServ QoS optimization and it focuses on finding optimal link weight settings. The algorithm uses iterative search and two-stage search: at first stage, link weights are given to find the worst case traffic matrix; at second stage traffic matrix is used to look for a new set of weights.

Mikoshi *et a*l [21] optimize input limitation ratio in each edge node, maximizing network throughput. The admission control method proposed in this paper considers some hop-by-hop topologies. Moreover, Linear Programming optimizes the model traffic developed in this work. The observed results improve the ones achieved in [20].

Qaraawy et al [22] implement Particle Swarm Optimization (PSO) to control the network congestion problem. This paper proposes the design of Proportional Integral (PI) controller, an AQM algorithm, to avoid congestion in computer networks. The PI controller designed with PSO provides performance improvement in different congested network scenarios.

Hasegawa *et al.* [23] proposed a distributed radio resource usage optimization algorithm to satisfy QoS requirements in heterogeneous wireless networks. Their optimization algorithm was based on a mutually connected neural network dynamics, which can run distributed and autonomously. The results obtained by computer simulation have shown that the approach is efficient even without centralized processing.

[24] describes experimental tests about adopting improved PSO algorithm to solve the deficiency of QoS optimization in Web service composition. After modeling the Web service composition problem as a multiobjective optimization problem, then an improved PSO algorithm is applied. The results showed that Web service composition model could improve speed and performance for finding the optimal solution. Gueye *et al.* [25] propose a Quality of Service framework for Web Service Composition, which is based on estimated response time of web services. Quality of services constrains are verified in order to provide quality level and performance. Marcon *et al.* [26] discuss about peak issue traffic performance on an Internet Service Provider (ISP). This work analyzes different kinds of traffic shaping techniques, whose objective is to minimize traffic peak, number of shaped flows, and delay. Kanuparthy and Dovrolis [27] propose a methodology for active end-to-end detection of traffic shaping mechanisms, eventually applied on ISPs. The main contribution of this work includes application and analysis of results from a large-scale deployment, which found traffic shaping mechanism in several ISPs. This approach opens discussion about Web neutrality, a topic heavily debated in Quality of Service context, where network connections should not be limited by source or destination requests, type of service, access technology,etc.

Another approach of traffic shaping was developed by La Piana *et al.* [28], who perform the study service control in order to ensure delay, jitter, throughput and loss rates. They used statistical analysis to optimize the adaptation of the scheduling algorithm to the various traffic sources. New recent works about traffic shaping were done by Daian and Giura [29], where the authors developed a study of traffic shaping in a set of dynamic traffic situations, especially stochastic properties of the traffic time series. They proposed a series of tests and compared the results obtained from OPNET Modeler simulation program to simulation with synthetic traffic generated by chaotic dynamic systems.

A brief discussion of these works puts together optimization techniques, particularly computational intelligence and evolutionary computation, and traffic shaping implementations on network computers. Up to the authors knowledge, it was not possible to find in the literature works that address traffic shaping parameters optimization using traffic emulation and artificial intelligence models. The main contribution of this work includes traffic shaping parameter optimization, and the combined implementation of artificial neural networks and genetic algorithms.

III. Methodology

The classification of traffic classes, the models of network workload, and the traffic shaping limitations are discussed along this section. In addition, the steps of the proposed methodology are described.

A. Discussion

The proposed methodology requires that the input and output traffic demands are classified by service or location, such as web services, DNS, SMTP, source or destination IP, and others. This classification generates traffic classes with three associated attributes each: assured rate ar_i, ceil rate c_i, and priority p_i.

Active Queue Managements (AQM), especially Random Early Detection (RED), are the foundation of several studies about TCP analysis. Hollot *et al* [30] propose, in a theoretical approach, a non-linear dynamic model of multiple flow TCP behavior. This work considers network parameters such as link capacity, load and round-trip time. In addition, Li *et al* [31] linearize the Hollot's dynamic model and include UDP traffic support. Using RED as the AQM algorithm, the authors performed and evaluated simulations using Network Simulator 2 (NS2) [32].

Although these works suggest a dynamic model for TCP and UDP flows, the proposed methodology adopts traffic shaping optimization as static inside an interval time δ_i. Requests made in δ_{i-1} reflects on network performance in δ_i. The traffic shaping optimization X_i^* considers optimal settings only at time δ_i, and optimal solutions X_j^* for δ_j should be obtained separately.

TCP traffic is expressive in most of network environment and, consequently, its transmission mechanism represents a dynamic model. The proposed methodology optimizes traffic shaping parameters in a static interval time δ_i. Therefore, at an interval δ_i, it is not necessary to differentiate traffic transmission protocols, such as TCP, UDP, and ICMP. It only requires the volume of traffic for each class.

This work deals with QoS parameters optimization as an empirical problem, considering the traffic shaping mechanism as a black box. It makes the proposed approach robust enough for dealing with any kind of traffic control mechanism that uses assured rate, ceil rate and priority as QoS parameters. Such as discussed in Section IV, this work adopts HTB as traffic shaping solution in GNU/Linux server environment.

QoS in computer network has limitations. Although DiffServ implementations perform QoS on all route nodes, it is not possible to ensure acceptable services levels. For instance, traffic shaping will not reach acceptable throughput and packet loss when network demand is much larger than the available resources.

B. Problem models

Suppose a traffic volume T, in an environment with bandwidth L, classified in traffic classes $ct_1, ct_2, \ldots, ct_i, \ldots, ct_n$. Each class ct_i is defined by and assured rate ar_i, a ceil rate c_i, and a priority p_i. At time interval δ, consider the throughput td_i as the load demand for each traffic class ct_i, in which the sum $\sum td_i = T$ represents the total load traffic. Traffic shaping mechanism would offer throughput to_i and, in congested networks, it would result in packet loss pli. Figure 2 illustrates the proposed model for traffic shaping optimization.

In order to provide better use of the network, the optimization of traffic shaping should take into account the following constraints:

$$c_i = L \tag{1}$$

$$0 < ar_i \leq L \tag{2}$$

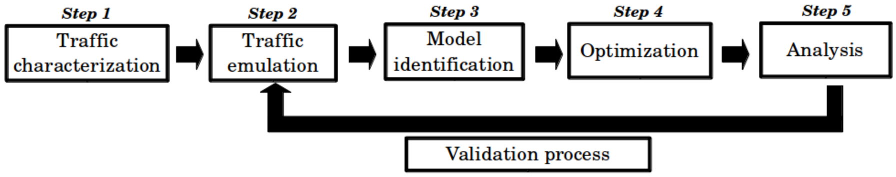

Figure. 1: Methodology steps

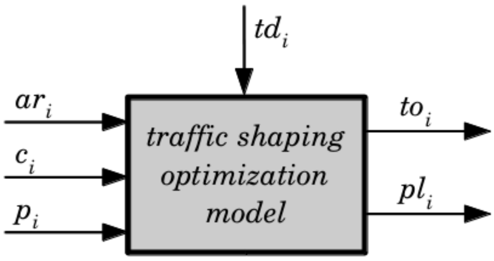

Figure. 2: Proposed model for traffic shaping optimization

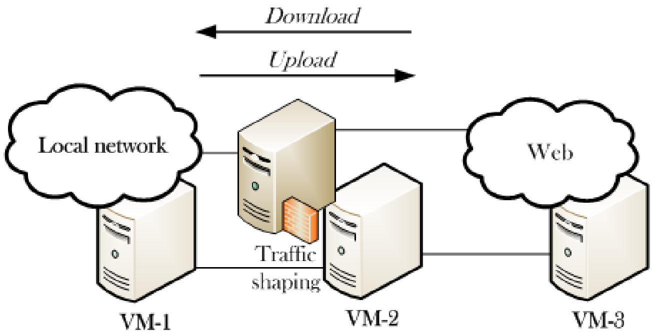

Figure. 3: Testbed environment

$$\sum_{i=1}^{n} ar_i = L \tag{3}$$

Ceil rates c_i for all traffic classes have the same bandwidth L value in order to reduce computational cost and simplify methodology implementation. Although traffic shaping mechanism allows $c_i < L$ on configuration, $c_i = L$ implies on best effort QoS policy. Moreover, all assured rates ar_i should be contained in the interval $(0, L]$, and the assured rate sum should correspond to the total bandwidth L. These constraints represent correct configuration of Hierarchical Token Bucket (HTB), the packet scheduler used as traffic shaping mechanism on this work.

C. Steps

The proposed methodology is illustrated in Figure 1. Its steps are described bellow:

1) Traffic Characterization

It consists on employing usual tools (Cacti[1], Zabbix[2], etc), to collect data from the network, such as throughput, packets sent, packets with errors, latency, jitter, etc. For more details on traffic characterization, as transmission protocol use, type of service index, and flows index, NetFlows [33] may be adopted in this step. Flowtools[3] and fprobe-ulog[4] implement these functions in GNU/Linux environment network.
In the current stage of the proposed methodology, only throughput demand (td_i) for each traffic class is monitored.
As discussed in subsection III-A and modeled in subsection III-B, the proposed methodology includes operation parameters optimization of traffic shaping at time interval δ. It is left to the network administrator to choose the representative interval time δ and granularity of traffic characterization (5, 10, 30 or 60 minutes), before implementing the proposed methodology.

2) Traffic Emulation

In this step, the environment features are reproduced in order to make it possible to perform tests with the traffic shaping mechanism. Figure 3 illustrates the virtual test bed proposed to generate dataset for model identification. $VM-1$ and $VM-3$ represent, respectively, local network demands and *World Wide Web*. $VM-2$ implements traffic shaping mechanism (HTB), through *traffic control*[5] (tc) tools.
$VM-1$ and $VM-3$ generate incoming and outgoing traffic through iperf[6]. Volume traffic demand (td_i) will be implemented as a constant bit rate model (CBR), by UDP transmission. In $VM-2$, qos-tools (developed in this work) generate assured rates (ar_i) and priority values (p_i) using Monte Carlo Simulations. Each execution takes about 2 minutes, being 1 minute for setup. One emulation scenario is composed of 2,000 samples which are employed for model identification.

3) Model Identification

It is used to create a computational model that mimics the traffic shaping mechanism, Hierarchical Token Bucket (HTB) in the specific case of this work. As a result, the model will be used to evaluate solutions in the next step of optimization with genetic algorithms.
This model is built based on Multi Layer Perceptron (MLP) neural networks [34] with one hidden layer, on ANFIS [35, 36], or on Radial Based Function [34]. For MLP implementations, the methodology evaluates models with 3, 5, 10, 12, 15, 20, 25 and 30 neurons on hidden layer. The validation

[1]Cacti: `http://www.cacti.net/`, accessed March 10th, 2013

[2]Zabbix: `http://www.zabbix.com/`, accessed March 10th, 2013

[3]flowtools: `http://freecode.com/projects/flow-tools`, accessed March 10th, 2013

[4]fprobe-ulog: `http://fprobe.sourceforge.net/`, accessed March 10th, 2013

[5]traffic control: `http://tldp.org/HOWTO/Traffic-Control-HOWTO/`, accessed March 10th, 2013

[6]iperf: `http://sourceforge.net/projects/iperf/`, accessed August 20th, 2012

of the models obtained from these architectures is performed using minimal validation error for all architectures.
This model is essential for performing optimization, since the evaluation of a candidate solution using the QoS tool would require an unacceptable large simulation time. Moreover, the difficulty of rebuilding the same features of traffic shaping mechanism from statistics generated by tool itself also justify the use of model identified by neural network on proposed methodology.

4) Optimization

In this step a genetic algorithm is used to optimize the traffic shaping model. Such as discussed in section III-B, this methodology deals with QoS optimization in a multiobjective approach. Therefore, the NSGA-II algorithm evaluates candidate solutions in order to maximize the offered throughput (to_i) and to minimize packet loss (pl_i).
The Equation (4) shows the fitness function employed for representing offered throughput. It is a function of $AR = (ar_1, ar_2, \dots, ar_n)$, $P = (p_1, p_2, \dots, p_n)$, and $TD = (td_1, td_2, \dots, td_n)$. It should be noticed that the constraints related to assured rate (Equations 1, 2, and 3) cannot be ignored.
When the ratio of throughput offered and demanded is close to 1, then the network resources are allocated in a such a way that they meet the traffic demands at time interval δ. As the throughput offered (to_i) becomes smaller than throughput demanded (td_i), it means that network is not capable of complying with demands. Values greater than 1 for throughput offered and demanded ratio are not possible in practice.
For each traffic class i, to_i/td_i is weighted by ω_i, that represents the normalized priority vector. Thus, considering $to_i/td_i = 1$ for all traffic class i, the maximum fitness function value will be $\sum \omega_i$.

$$F^{TO}(AR, P, TD) = \sum_{i=1}^{n} \omega_i(p_i) \frac{to_i(ar_i, p_i, td_i)}{td_i} \quad (4)$$

Equation 5 shows the fitness function that evaluates candidate solutions for packet loss. For each traffic class, the packet loss is between 0% and 100% and it is weighted by ω_i. The maximum value of packet loss fitness function is $n \times 100\%$ (where n represents the total of traffic class), and the minimum value is 0, which is the best case for this work.

$$F^{PL}(AR, P, TD) = \sum_{i=1}^{n} \omega_i(p_i)\ pl_i(ar_i, p_i, td_i) \quad (5)$$

Therefore, throughput offered and packet loss problem optimization can be defined as described on Equation 6.

$$\begin{array}{ll} \text{minimize} & -F^{TO} \\ \text{minimize} & F^{PL} \\ \textit{subject to} & c_i = L, \\ & 0 < ar_i \leq L, \\ & \sum_{i=1}^{n} ar_i = L. \end{array} \quad (6)$$

As a preliminary version of this stage, throughput offered and packet loss will be optimized in a mono objective approach. This procedure is necessary to evaluate selection, crossover, and mutation genetic operators to be used in NSGA-II algorithm on multiobjective optimization. For this instance, Table 1 shows mono objective genetic algorithms proposed for evaluation.

Table 1: Variaes dos algoritmos genticos propostos

Variation	Selection	Crossover	Mutation
AG-1	SUS	Real Interpolation	Gaussian
AG-2	SUS	Real Interpolation	Polinomial
AG-3	SUS	SBX	Gaussiana
AG-4	SUS	SBX	Polinomial
AG-5	Binary Tournament	Real Interpolation	Gaussiana
AG-6	Binary Tournament	Real Interpolation	Polinomial
AG-7	Binary Tournament	SBX	Gaussiana
AG-8	Binary Tournament	SBX	Polinomial

5) Analysis

This step provides genetic algorithm analysis, which consider standard error to indicate the best GA implementation. A multicriteria statistical evaluation tool [37] is employed to evaluate the genetic algorithm implemented in step 4. Moreover, the set of solutions obtained by the GA should be analyzed in this step in order to perform decision making.
For validation of the results, the traffic shaping operational parameters are evaluated in the emulation environment again, such as described in step 2. The validation involves the comparison between the fitness function value obtained in optimization and fitness function observed in emulation. This procedure should be performed before the configuration of the traffic shaping mechanisms under real traffic.
The robustness of the final result of the methodology is evaluated based on the variation of the demand by $\pm 10\%$, $\pm 30\%$ and $\pm 50\%$. Further, the demand is kept fix and the bandwidth is decreased by -40%, -60% and -80%. This analysis allows verifies that the optimal solution is able to offer the same levels of quality for anomalous scenarios
The Figure 4 shows a candidate Pareto-front for such a problem. For policies that require higher throughput rates and high packet loss, network administrator should choose solutions on the region *A* of the front. Policies that consider average values of the objectives should be implemented as solutions of region *B*. Finally, policies that require low packet loss should support the implementation of a solution of region *C*. The diversity of the solutions can be evaluated using the hypervolume performance metric [38].

IV. Case study

The proposed methodology has been applied to a real case study observed in CEFET-MG Local Area Network. The infrastructure of this network includes Web, DNS and database services. It provides Internet connection to approximately 1,200 hosts. Cacti and Nagios[7] tools are used to perform network management, and *tc* employs *HTB* mechanism for traffic shaping.
Currently, CEFET-MG has an 1Gbps bandwidth. However, it operated with 6Mbps until February 2012, for the same traffic demand. It obviously resulted into a very congested

[7]Nagios: `http://www.nagios.org/`, accessed March 10th, 2012

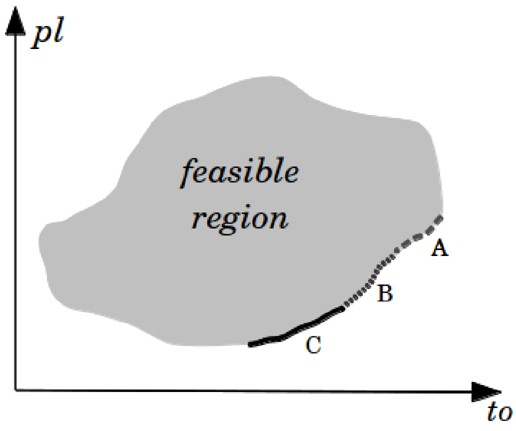

Figure. 4: Pareto for decision making for traffic shaping configuration

environment. In this scenario, the traffic shaping parameters employed were not enough to ensure acceptable service levels, what motivated this work.

The total demand of CEFET-MG was divided in six classes, as follows: class 10 – local web servers; class 20 – HTTP traffic; class 30 – socials networks; class 40 – CEFET-MG web services (email, DNS, HTTP, etc); class 50 - others institutional services; and class 60 - non classified traffic. Figures 5 and 6 show a sample of incoming and outgoing traffic for each traffic class. These samples were taken of on September 20th, 2012, between 12:00 p.m. and 13:00 p.m. In these descriptions, incoming and outgoing traffic have 01 hour of granularity.

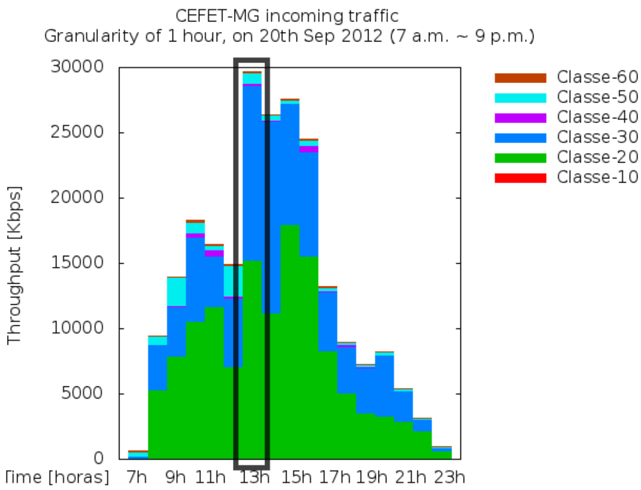

Figure. 5: CEFET-MG incoming traffic managed by traffing shapping

Table 2 illustrates network characterization in terms of traffic volume, such as required by the proposed methodology. Besides, average throughput demands for download and upload traffic are also shown (p_i is the priority of class i). It should be noticed the predominance of classes 20 and 30: download traffic for classes 20 and 30 are bigger than CEFET-MG available bandwidth ($L = 6Mbps$).

Since the demand for upload traffic is considerably lower than the bandwith available (less than 3Mbps vs 6Mbps),

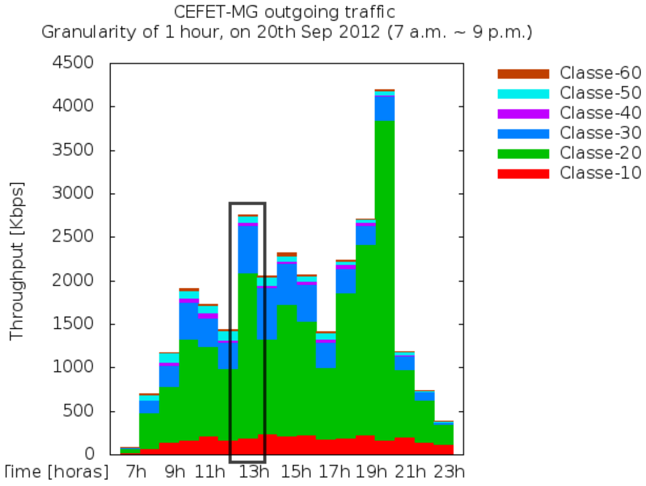

Figure. 6: CEFET-MG outgoing traffic managed by traffing shapping

Table 2: Traffic on Septemper 20th, 2012 - 12h 13h (granularity of 1 hour)

Traffic class	Priority	Download [Kbps]		Upload [Kbps]	
		Kbps	%	Kbps	%
10: Local servers	3	114	0.38	199	7.22
20: Web: HTTP traffic	4	15,160	51.15	1,897	68.83
30: Web: Social networks	6	13,357	45.06	543	19.70
40: CEFET-MG Web	1	159	0.54	36	1.31
50: CEFET-MG Systems	2	801	2.70	71	2.58
60: Nonclassified traffic	5	49	0.17	10	0.36
Total		29,640	100.00	2,756	100.00

then the implementation of traffic shaping to upload would not lead to performance gains. Based on this observation, outgoing traffic has been ignored during optimization.

The traffic emulation (step 2) considers a bandwidth (L) of 6Mbps for download and upload traffic. For these experiments, the priorities p_i have been set as static values, known a priori in the interval $\{1, \ldots, n\}$. The total download and upload throughput demand can also be static (T) or random, following uniform distribution $U(T - \epsilon, T + \epsilon)$, with $\epsilon = 17\%$. Table 3 summarizes the instances considered for traffic emulation.

Table 3: Instances for traffic emulation

Instances	P_i	**Network workload** (T)	**Demand *Throughput*** (td_i)
Instance 1	Static values predefined	$T = \sum td_i$	td_i
Instance 2	Arrangement $\{1, \ldots, n\}$ with repetition	$T = \sum td_i$	td_i
Instance 3	Arrangement $\{1, \ldots, n\}$ without repetition	$T = \sum td_i$	td_i
Instance 4	Static values predefined	$T = \sum td_i$ $T' \sim U(T-\epsilon, T+\epsilon)$	$\frac{td_i}{\sum td_i} T'$
Instance 5	Arrangement $\{1, \ldots, n\}$ with repetition	$T = \sum td_i$ $T' \sim U(T-\epsilon, T+\epsilon)$	$\frac{td_i}{\sum td_i} T'$
Instance 6	Arrangement $\{1, \ldots, n\}$ without repetition	$T = \sum td_i$ $T' \sim U(T-\epsilon, T+\epsilon)$	$\frac{td_i}{\sum td_i} T'$

In a first test, the instance 1 (*download*) has been used to

evaluate the proposed methodology. At this moment, only instance 1 has been implemented in order to proceed the work on Identification Model step. Ten MLPs of format (6,k,1) (6 inputs, k neurons in the hidden layer, 1 output), ANFIS (with 2 membership functions and 5 epochs) and RBF (spread factor equal 100) neural networks have been built for generating models for throughput and packet loss for the six classes. Values $k = 3, 5, 10, 12, 15, 20, 25, 30$ have been tested. The MLP networks have been trained using Levenberg-Marquardt training algorithm. The stop criteria were global error or norm of the gradient vector (both should be higher than 10^{-8}), with 1,000 epochs ceil.
Three scenarios are considered on optimization: *Scenario 1* ($L = 30Mbps$), *Scenario 2* ($L = 24Mbps$), and *Scenario 3* ($L = 6Mbps$). The third scenario was the real situation of CEFET-MG before the implantation of the 1Gbps network. The other ones are theoretical scenarios of bandwidth expansion.

Table 4: Scenario 1 ($L = 30Mbps$): Neural network error validation for throughput offered and packet loss

Variable	Best Model	Error validation	Epochs
td_1	MLP(6,3,1)	0.0035472	10
td_2	MLP(6,15,1)	0.0038104	10
td_3	MLP(6,10,1)	0.0011807	13
td_4	MLP(6,5,1)	0.0034722	13
td_5	MLP(6,3,1)	0.0034340	12
td_6	MLP(6,3,1)	0.0016098	9
pl_1	MLP(6,3,1)	0.0000000	0
pl_2	MLP(6,12,1)	0.0007846	17
pl_3	MLP(6,12,1)	0.0009170	14
pl_4	MLP(6,3,1)	0.0000000	0
pl_5	MLP(6,3,1)	0.0000000	0
pl_6	MLP(6,3,1)	0.0000000	0

Table 5: Scenario 2 ($L = 24Mbps$): Neural network error validation for throughput offered and packet loss

Variable	Best Model	Error validation	Epochs
td_1	MLP(6,30,1)	0.0047572	13
td_2	MLP(6,5,1)	0.0031371	12
td_3	MLP(6,3,1)	0.0012160	45
td_4	MLP(6,12,1)	0.0046591	12
td_5	MLP(6,15,1)	0.0046365	16
td_6	MLP(6,3,1)	0.0021880	11
pl_1	MLP(6,3,1)	0.0000000	0
pl_2	MLP(6,15,1)	0.0012241	46
pl_3	MLP(6,10,1)	0.0018991	13
pl_4	MLP(6,3,1)	0.0000000	0
pl_5	MLP(6,3,1)	0.0000000	0
pl_6	MLP(6,3,1)	0.0000000	0

A comparison between the neutral network models built to scenarios 1, 2 and 3 can be performed from Tables 4, 5 and 6. These results refer to throughput td_i and packet loss pl_i observed for all classes. Validation errors and number of epochs required for training, important metrics to evaluate models, are also presented. Classes 1, 4, 5, and 6 have reached validation error 0 for packet loss because the traffic demand for these classes is very low. Besides, it was possible to note that the best model for all the scenarios is always a MLP network, what suggests that this kind of regression mechanism is more suitable for the considered application.
It should be noticed from these results that the proposed approach was able to generate good models for all classes. The validation errors observed were very low, what indicates that it was able to obtain generalized models. It means robustness on the evaluation of solutions inside the optimization algorithm, since it depends on the approximated models in order to evaluate the candidate parameters.

Table 6: Scenario 3 ($L = 6Mbps$): Neural network error validation for throughput offered and packet loss

Variable	Best Model	Error validation	Epochs
td_1	MLP(6,3,1)	0.0025289	8
td_2	MLP(6,3,1)	0.0003766	34
td_3	MLP(6,5,1)	0.0000001	656
td_4	MLP(6,5,1)	0.0025243	10
td_5	MLP(6,3,1)	0.0025042	10
td_6	MLP(6,3,1)	0.0024328	10
pl_1	MLP(6,3,1)	0.0000000	0
pl_2	MLP(6,3,1)	0.0014640	430
pl_3	MLP(6,3,1)	0.0001141	60
pl_4	MLP(6,3,1)	0.0000000	0
pl_5	MLP(6,3,1)	0.0000000	0
pl_6	MLP(6,3,1)	0.0000000	0

At this moment, the authors are working on steps 4 and 5. The genetic algorithm that will be used to optimize the QoS parameters is being implemented. It will use the MLP models in order perform the optimization. The development of the analysis module will start after a preliminary evaluation of the proposed algorithm.

V. Conclusion

This work proposes a new methodology for traffic shaping multiobjective optimization. In a scenario of Next Generation Network, including web services practices, QoS insurance is essential to provide acceptable service levels.
The proposed methodology includes traffic estimation, characterization and emulation, optimization of QoS and further solution analysis. This approach enhances the applicability in environments with high traffic demands, characterized by hybrid traffic. The model identification of traffic shaping mechanism, implemented by Artificial Neural Networks, allows to optimize any traffic shaping tool which is based on assured rate, ceil rate and priority parameter tuning. It has been employed to a real case study, handled by CEFET-MG. Until now, three steps have been completed: traffic characterization, traffic emulation and model identification. The next steps involve the implementation of a multiobjective genetic algorithm (NSGA-II) to optimize traffic shaping parameters. Furthermore, a module which is able to perform solution analysis will be implemented too.
Future work based on the discussion and partial implementation of the methodology can be presented. The first one involves studying methodologies to support different behaviors and changes on network environment over time. As the optimization process occurs only at a single time interval, it is desirable that the methodology presents a set of solutions applicable over a longer period of time. This improvement can be achieved by identifying changes in traffic demand by mechanisms of fuzzy logic. For each new pattern demand network identified, the methodology is reapplied to build a database of optimized solutions for traffic shaping mechanism.

However, considering the dynamics of the network environment, it is desirable that the methodology support Variable Bit Rate (VBR) model data transmission or replication of traces network. This approach suggests that gains representing traffic shaping are obtained, from the time that the changes are well identified through the evaluation process.
Another important topic is to evaluate the impact of changing the design time intervals δ_k. The granularity required for the application of this methodology represents a trade-off: small time intervals increase the gains in performance with high computational cost, as well as large time intervals reduces the cost of the methodology with no significant gains.
Finally, after the implementation of dynamic characteristic of computer networks in the methodology, the traffic shaping setups can be optimized. Network connections can be identified and classified into classes from traffic patterns with the support of fuzzy logic and artificial neural networks. Thus, the architecture of traffic shaping mechanisms implemented in the network can be optimized in terms of the adequacy of the number and traffic classes allocation, as well as metrics such as throughput offered, packet loss and latency.

Acknowledgments

This research was supported by the Brazilian National Institute of Science and Technology for the Web (CNPq grant no. 573871/2008-6), CAPES, CNPq, Finep, and Fapemig. The authors also would like to thank CEFET-MG for provide all network environment and computational resources needed for this work.

References

[1] *Y.2011: General principles and general reference model for Next Generation Networks*, International Telecommunication Union (ITU-T) Std. Y.2011, 10 2004. [Online]. Available: http://www.itu.int/rec/T-REC-Y.2011-200410-I/en

[2] *Y.2001: General overview of NGN*, International Telecommunication Union (ITU-T) Std. Y.2001, 12 2004. [Online]. Available: http://www.itu.int/rec/T-REC-Y.2001-200412-I/en

[3] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC 1633 (Informational), Internet Engineering Task Force, Jun. 1994. [Online]. Available: http://www.ietf.org/rfc/rfc1633.txt

[4] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," RFC 2475 (Informational), Internet Engineering Task Force, Dec. 1998, updated by RFC 3260. [Online]. Available: http://www.ietf.org/rfc/rfc2475.txt

[5] S. Shenker, C. Partridge, and R. Guerin, "Specification of Guaranteed Quality of Service," RFC 2212 (Proposed Standard), Internet Engineering Task Force, Sep. 1997. [Online]. Available: http://www.ietf.org/rfc/rfc2212.txt

[6] *Y.1291: An architectural framework for support of Quality of Service in packet networks*, International Telecommunication Union (ITU-T) Std. Y.1291, 05 2004. [Online]. Available: http://www.itu.int/rec/T-REC-Y.1291-200405-I/en

[7] M. Devera, "Htb linux queuing discipline manual - user guide," Tech. Rep., 2002. [Online]. Available: http://luxik.cdi.cz/~devik/qos/htb/manual/userg.htm

[8] U. Cavalca, C. Mesquista, A. Pereira, and E. Carrano, "A methodology for traffic shaping multiobjective optimization for next generation network," in *Computational Aspects of Social Networks (CASoN), 2012 Fourth International Conference on*, Nov., pp. 345–350.

[9] Y. Liu and X. Liang, "New regulations to the next generation network," in *Communications and Mobile Computing, 2009. CMC '09. WRI International Conference on*, jan. 2009, vol. 2, pp. 172 –174.

[10] A. Gupta, "Network management: Current trends and future perspectives," *Journal of Network and Systems Management*, vol. 14, pp. 483–491, 2006, 10.1007/s10922-006-9044-7.

[11] *E.800: Definitions of terms related to quality of service*, International Telecommunication Union (ITU-T) Std. E.800, 09 2008. [Online]. Available: http://www.itu.int/rec/T-REC-E.800-200809-I/en

[12] *E.802: Framework and methodologies for the determination and application of QoS parameters*, International Telecommunication Union (ITU-T) Std. E.800, 02 2007. [Online]. Available: http://www.itu.int/rec/T-REC-E.802-200702-I/en

[13] G. Canfora, M. Di Penta, R. Esposito, and M. L. Villani, "An approach for qos-aware service composition based on genetic algorithms," in *Proceedings of the 2005 conference on Genetic and evolutionary computation*, ser. GECCO '05. New York, NY, USA: ACM, 2005, pp. 1069–1075.

[14] Y. Yao and H. Chen, "Qos-aware service composition using nsga-ii," in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, ser. ICIS '09. New York, NY, USA: ACM, 2009, pp. 358–363.

[15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *Evolutionary Computation, IEEE Transactions on*, vol. 6, no. 2, pp. 182 –197, apr 2002.

[16] F. Rosenberg, M. B. Müller, P. Leitner, A. Michlmayr, A. Bouguettaya, and S. Dustdar, "Metaheuristic optimization of large-scale qos-aware service compositions," *Proceedings of the 2010 IEEE International Conference on Services Computing*, pp. 97–104, 2010.

[17] F.-C. Pop, D. Pallez, M. Cremene, A. Tettamanzi, M. Suciu, and M. Vaida, "Qos-based service optimization using differential evolution," in *Proceedings of the*

13th annual conference on Genetic and evolutionary computation, ser. GECCO '11. New York, NY, USA: ACM, 2011, pp. 1891–1898. [Online]. Available: http://doi.acm.org/10.1145/2001576.2001830

[18] M. Suciu, M. Cremene, F. Pop, and D. Dumitrescu, "Equitable solutions in qos-aware service optimization," in *Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference companion*, ser. GECCO Companion '12. New York, NY, USA: ACM, 2012, pp. 1537–1538.

[19] S. K. Jha and M. Hassan, *Engineering Internet Qos*. Norwood, MA, USA: Artech House, Inc., 2002.

[20] J. Chu and C.-T. Lea, "Optimal link weights for maximizing qos traffic," in *Communications, 2007. ICC '07. IEEE International Conference on*, june 2007, pp. 610 –615.

[21] T. Mikoshi, T. Takenaka, T. Fujiwara, E. Oki, and K. Shiomoto, "Optimization of input admissible traffic flow guaranteeing qos," *Communications Letters, IEEE*, vol. 13, no. 1, pp. 49 –51, january 2009.

[22] S. Qaraawy, H. Ali, and A. Mahmood, "Particle swarm optimization based robust controller for congestion avoidance in computer networks," in *Future Communication Networks (ICFCN), 2012 International Conference on*, april 2012, pp. 18 –22.

[23] M. Hasegawa, A. Funatsu, and H. Harada, "A distributed radio resource usage optimization algorithm to satisfy qos requirements by neural network dynamics," *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, pp. 2255 –2259, sept. 2009.

[24] L. Huang, X. Zhang, Y. Huang, G. Wang, and R. Wang, "A qos optimization for intelligent and dynamic web service composition based on improved pso algorithm," *Networking and Distributed Computing (ICNDC), 2011 Second International Conference on*, pp. 214 –217, sept. 2011.

[25] B. Gueye, I. Niang, B. Gueye, and M. O. Deye, "Qos4wsc: A framework for web services composition based on qos constraints," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 5, pp. 488–498, 2013. [Online]. Available: http://www.mirlabs.org/ijcisim/regular_papers_2013/Paper125.pdf

[26] M. Marcon, M. Dischinger, K. Gummadi, and A. Vahdat, "The local and global effects of traffic shaping in the internet," in *Communication Systems and Networks (COMSNETS), 2011 Third International Conference on*, jan. 2011, pp. 1 –10.

[27] P. Kanuparthy and C. Dovrolis, "Shaperprobe: end-to-end detection of isp traffic shaping using active methods," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, ser. IMC '11. New York, NY, USA: ACM, 2011, pp. 473–482. [Online]. Available: http://doi.acm.org/10.1145/2068816.2068860

[28] A. La Piana, I. Ferraro, and C. Beoni, "Queuing, scheduling and traffic shaping over ip networks," *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, vol. 4, pp. 2322 – 2326 Vol.4, may 2004.

[29] D.-S. Daian and D.-H. Giura, "Traffic shaping and traffic policing impacts on aggregate traffic behaviour in high speed networks," *Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on*, pp. 465 –467, may 2011.

[30] C. Hollot, V. Misra, D. Towsley, and W.-B. Gong, "A control theoretic analysis of red," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, 2001, pp. 1510–1519.

[31] W. Li, L. Zeng-zhi, and C. Yan-ping, "A control theoretic analysis of mixed tcp and udp traffic under red based on nonlinear dynamic model," in *Information Technology and Applications, 2005. ICITA 2005. Third International Conference on*, vol. 2, july 2005, pp. 747–750.

[32] "The Network Simulator ns-2 (v2.1b8a)," http://www.isi.edu/nsnam/ns/, Oct. 2001.

[33] B. Claise, "Cisco Systems NetFlow Services Export Version 9," RFC 3954 (Informational), Internet Engineering Task Force, Oct. 2004. [Online]. Available: http://www.ietf.org/rfc/rfc3954.txt

[34] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

[35] J.-S. R. Jang, "Fuzzy modeling using generalized neural networks and kalman filter algorithm," in *Proceedings of the ninth National conference on Artificial intelligence - Volume 2*, ser. AAAI'91. AAAI Press, 1991, pp. 762–767. [Online]. Available: http://dl.acm.org/citation.cfm?id=1865756.1865795

[36] J.-S. Jang, "Anfis: adaptive-network-based fuzzy inference system," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 23, no. 3, pp. 665–685, May/Jun 1993.

[37] E. Carrano, E. Wanner, and R. Takahashi, "A multicriteria statistical based comparison methodology for evaluating evolutionary algorithms," *Evolutionary Computation, IEEE Transactions on*, vol. 15, no. 6, pp. 848 –870, dec. 2011.

[38] K. Deb and D. Kalyanmoy, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: John Wiley & Sons, Inc., 2001.

Author Biographies

Ulisses Cavalca received Mathematics degree from Universidade Estadual Paulista (UNESP), Guaratinguetá, Brazil, in 2007. He also work with management of infrastructure and services in computer network, and GNU/Linux system operating administration. Since February 2011, he has been a master student of Mathematical and Computational Modeling at Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte. His research interests include network management, network performance, Quality of Service solutions, evolutionary algorithms, and multiobjective optimization.

Caio Mesquita since 2008 is a student of Computer Engineering at CEFET-MG until now a days. His research interests include network management, network performance and Quality of Service solutions.

Adriano C. Machado Pereira is an Adjunct Professor in Computer Science Department at Federal University of Minas Gerais (DCC / UFMG), Brazil. He received his bachelor degree in Computer Science at UFMG in 2000, his MSc. in 2002, and his Ph.D. in 2007. He also had performed a Post-Doc research in electronic markets in 2008-2009. His research interests include e-Business, e-Commerce, Workload Characterization, Distributed Systems, Web 2.0, Social Networks, Social Computing, Social Commerce, Performance of Computer Systems, Web Technologies, and Business Intelligence. He is also a member of the Brazilian National Institute of Science and Technology for the Web - INWEB (www.inweb.org.br).

Eduardo Gontijo Carrano received the Ph.D. degree in electrical engineering from the Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, in 2007. From May 2006 to May 2007, he was a Visiting Student with the Center for Intelligent Systems, Universidade do Algarve, Faro, Portugal. From September 2007 to January 2008, he was a Post–Doctoral Researcher with the Department of Mathematics, UFMG. From January 2008 to June 2011, he was an Assistant Professor with Centro Federal de Educação Tecnológica de Minas Gerais (CEFET–MG), Belo Horizonte. Since June 2011, he has been an Assistant Professor with Department of Electrical Engineering, UFMG. His current research interests include network design, combinatorial optimization, evolutionary algorithms, multiobjective optimization, optimization theory, algorithm evaluation, and dynamic systems.