Enhancing Path Selection in Multihomed Nodes

Bruno Sousa[1], Kostas Pentikousis[2], and Marilia Curado[1]

[1] CISUC, University of Coimbra
Polo II, Pinhal de Marrocos 3030-290, Coimbra, Portugal
{bmsousa,marilia}@dei.uc.pt
[2] Huawei Technologies
Carnotstrasse 4, 10587 Berlin, Germany
k.pentikousis@huawei.com

Abstract. Path selection in multihomed nodes can be enhanced by optimization techniques that consider multiple criteria. With NP-Hard problems, MADM techniques have the flexibility of including any number of benefits or costs criteria and are open regarding the functions that can be employed to normalize data or to determine distances. TOPSIS uses the Euclidean distance (straight line) while DiA employs the Manhattan distance (grid-based) to determine the distance of each path to ideal values. MADM techniques have been employed in distinct areas, as well. Such openness and flexibility may lead to sub-optimal path selection, as their optimality is associated with functions that determine distance as a straight line or as grid path, and not inside an ideal range determined by the type of criteria. In this paper we propose the MeTH distance which considers the type of criteria, whether benefits or costs. In addition, we establish a MADM evaluation methodology based on statistical analysis that enables an objective comparison between MADM mechanisms and respective functions for path selection. With the proposed MADM evaluation methodology, we demonstrate that our MeTH distance is more efficient for the path selection problem than Euclidean and Manhattan distances.

Keywords: MADM, DoE, TOPSIS, path selection, multihoming, evaluation.

1 Introduction

Through the diversity of interfaces, modern device multihoming is characterized by the availability of multiple traffic paths for diverse flows, with different features. However, when considering multiple criteria, optimal path selection becomes a NP-Hard problem [1]. Different approaches exist to solve such kind of problems, namely Linear Programming (LP) or Multiple Attribute Decision Mechanism (MADM). LP techniques are able to provide optimal solutions but with the price of being tied to the problem being optimized [2, 3]. Thus, linear programming cannot be employed, without any adaptation to other problems

D. Pesch et al. (Eds.): MONAMI 2013, LNICST 125, pp. 69–82, 2013.

and scenarios or even to include additional criteria. For instance, overlay multicasting solutions relying on linear programming [4] cannot be applied to path selection optimization.

MADM techniques cope with the limitations of linear programming solutions by supporting optimization without being tied to a particular problem. The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) [5] a MADM technique, is employed in distinct areas, ranging from social sciences to path selection optimization problems [6,7]. The flexibility of incorporating criteria, the possibility of weighting the diverse attributes and the simplicity of use, make MADM preferable in comparison to LP. MADM formulates a score for each path, which is based on the distance that each path has regarding ideal values. To determine such distance, several functions are applied, such as normalization of data and maximum and minimum procedures, to determine ideal values. While determination of the ideal values is similar between techniques, normalization and distance functions are distinct, as the examples of Euclidean employed by TOPSIS [5], Manhattan used by Distance to Ideal Alternative (DiA) [8], Mahalanobis in Novel Method based on Mahalanobis Distance (NMMD) [9], or distance in a geometric plane [10]. But the effect that such function has on the path selection cannot simply rely on handover performance (e.g., ping-pong effects [9]), as such kind of evaluation does not consider the effects that weights of different criteria have on the score. Other kinds of evaluations only consider specific functions in MADM. For instance, normalization techniques [11] are compared, but such an approach is rather incomplete, as distance or scoring functions MADM are ignored.

The contributions of this paper are twofold: First, we specify a distance function that considers the type of criteria. Second, an evaluation methodology is specified to assess the performance of MADM regarding their rankings according to data from networks with multihomed nodes and respective criteria (i.e. bandwidth, round trip time, jitter, loss). The evaluation methodology, publicly available in [12], relies on statistical analysis from the Design of Experiments (DoE) [13]. With DoE diverse experiments are executed to assess the sensitivity on ranking that techniques have with different criteria weights. With Analysis of Variance (ANOVA), the ranking of MADM techniques is compared regarding the model fitness in terms of completeness, coefficient of determination and variance between experiments or inside experiments. We conducted a comparative evaluation of TOPSIS, DiA and MeTH with our proposed methodology, using data from multihomed nodes collected in testbeds. Achieved results demonstrate that MeTH is the only technique that is able to detect interaction between criteria (i.e. if bandwidth increases, round trip time and jitter may decrease).

The remainder of this paper is organized as follows: Section 2 overviews related work. Section 3 introduces MeTH in a comparative approach to TOPSIS and DiA techniques. The evaluation methodology is introduced in Section 4 and evaluations details are described in Section 5. Results are presented and discussed in Section 6. Finally, Section 7 concludes the paper.

2 Related Work

This section reviews related work regarding path optimization employing MADM techniques. Associated evaluation mechanisms are also included. Other kind of optimization techniques like Linear Programming are outside the scope of this paper.

In the path selection problem, TOPSIS is employed as a mechanism to select the best path to enable flow distribution [7]. Nonetheless, the article only addresses implementations issues, as authors aim to demonstrate that MADM techniques can be employed on Linux hosts. The DiA [8] is a MADM mechanism that aims to cope with the ranking abnormality of TOPSIS. An issue in score occurs when one less performant alternative is removed from selection. The NMMD [9], based on the Mahalanobis distance, enables correlation between criteria to overcome ranking abnormality of TOPSIS. The M-TOPSIS [10] uses a modified distance based on geometric planes with the argument of solving the ranking abnormality of TOPSIS. The evaluation of the previous techniques compares the performance of the respective techniques with TOPSIS in ranking abnormality situations. Nonetheless, we argue that such type of evaluation is sub-representative to enable an efficient comparison of MADM techniques in the path selection problem. Furthermore, in scenarios without failures, multihomed nodes may have all the interfaces available without any instability associated.

As stated, MADM techniques comprise several steps. Normalization, one initial step, is evaluated in TOPSIS by considering different normalization functions [11]. Vector normalization is presented as the one providing better support for different problem sizes. For instance, to choose between 2 or 10 paths. Considering the configuration of multihomed nodes, nowadays the path selection problem may include gigabit and wireless interfaces, not exceeding a few paths. Moreover, the evaluation performed is based on synthetic data.

DoE [13] has been employed to assess TOPSIS efficiency in computer-integrated manufacturing technologies [14,15]. Despite, employing DoE there are no comparisons between different MADM techniques, as TOPSIS is assumed to have a better performance than other related techniques. In the path selection problem we do not have such assumption, instead we employ DoE to determine objectively the most performant MADM technique, regarding its statistical results.

3 MADM for Path Selection

This section introduces MeTH, a MADM technique that enhances path selection by introducing correlation between criteria. The correlation support is based on simple functions, such as average and variance. We follow a comparative approach to introduce MeTH. Namely, we perform a comparison with TOPSIS and DiA, highlighting the main differences between the techniques, as described in the next paragraphs.

Table 1. Distance Functions

Step	TOPSIS^{a,c}	DiA^{a,c}	Meth^{a,b,c}
Distance	$D_i = \sqrt{Id_j - v_{i,j}}$	$D_i = \lvert Id_j - v_{i,j} \rvert$	$D_i = \frac{(Id_j - v_{i,j})^2}{\lvert Id_j - Sd_j \rvert + 0.001}$
Score	$S_i = \frac{D_i^-}{D_i^- + D_i^*}$	$S_i = \sqrt{(D_i^*)^2 + (D_i^-)^2}$	$S_i = \sqrt{D_i^* + D_i^-}$
Rank	Best=descend(S_i)	Best=ascend(S_i)	Best=ascend(S_i)

[a] Id_j is the Ideal solution.
[b] Benefits: $Sd_j = \overline{X_j} + Var(X_j)$; Costs: $Sd_j = \overline{X_j} - Var(X_j)$
[c] Benefits: D_i^*; Costs: D_i^-

Step 1 - Decision Matrix. Gather the decision matrix with nb-benefits criteria and nc-costs criteria for the m paths (i.e. alternatives in MADM nomenclature).

Step 2 - Normalization. The decision matrix is normalized using the vector normalization, as it is agnostic to the problem size [11]. Normalized scores r_{ij} are obtained by employing the following relation $r_{ij} = \frac{x_{ij}}{\sqrt{\sum x_{ij}^2}}$ $for\ i = 1, \cdots, m; j = 1, \cdots, n$. x_{ij} correspond to the original values in the decision matrix.

Step 3 - Weighting. The normalized decision matrix is weighted by multiplying the weights w_j of criterion j with the respective normalized score r_{ij}, as follows: $v_{ij} = w_j \cdot r_{ij}$.

Step 4 - Ideal Solutions. Positive-ideal and negative-ideal solutions are determined by A^* and A^- terms, respectively:

$$A^* = \{v_1^*, v_2^*, \cdots, v_{nb}^*\} \tag{1}$$
$$A^- = \{v_1^-, v_2^-, \cdots, v_{nb}^-\} \tag{2}$$
$$\text{Where:}\quad v_j^* = max(v_{i,j})\ \forall i = 1, \cdots, m\ j = 1, \cdots, nb$$
$$v_j^- = min(v_{i,j})\ \forall i = 1, \cdots, m\ j = 1, \cdots, nc$$

Step 5 - Distance. This step computes the separation that each path has to the ideal solution. TOPSIS uses the Euclidean distance, DiA employs the Manhattan distance. For path selection, we introduce the MeTH distance that has the advantage of introducing correlation between criteria, through the arithmetic average and variance functions, as summarized in Table 1. It has been demonstrated that correlation avoids ranking abnormalities of TOPSIS and DiA [9]. MeTH also considers the type of criteria type in the formulation of distance.

Step 6 - Score. Scoring is obtained by combining the separation from positive and negative ideal solutions, D_i^* and D_i^-, respectively. Each technique has different forms of combining distances, as depicted in Table. 1.

Step 7 - Ranking. Ranking relies on ordering score vectors Si. Since scoring is different between techniques, ordering is performed in descending for TOPSIS and in ascending order for DiA and MeTH.

Table 2. Decision matrix for 3 criteria with 2^k factorial design

Id	x_1	x_2	x_3	**Effect**
1	-	-	-	(1)
2	+	-	-	x_1
3	-	+	-	x_2
4	+	+	-	x_1x_2
5	-	-	+	x_3
6	+	-	+	x_1x_3
7	-	+	+	x_2x_3
8	+	+	+	$x_1x_2x_3$

$$
\begin{aligned}
Y =& \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\
& \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \\
& \beta_6 x_2 x_3 + \beta_7 x_1 x_2 x_3 + \epsilon
\end{aligned} \tag{3}
$$

As demonstrated, MADM share functions in some steps. Functions performing the same goal (e.g., distance determination) but with different formulations, justify the performance difference between these techniques. The following section presents a methodology to assess such difference in performance.

4 An Evaluation Methodology for MADM Techniques

This section specifies an evaluation methodology for MADM techniques, which can be used in the context of multihoming nodes for path selection. The aim is to compare MADM techniques more efficiently and without relying on sub-representative evaluation metrics, such as handover ratios, in the path selection problem.

The DoE or experimental design [13] allows to plan experiments, in such a way that facilitates analyses and conclusions. DoE has different techniques to promote analyses, specially the 2^k factorial design that allows to assess the effect of several variables over a response. In the path selection problem, the criteria may include benefits, such as security, coverage, bandwidth and costs like round tip time, jitter and packet loss. The 2^k factorial design specifies full factory experiments for the k main effects, $(\frac{k}{2})$ two-factor interactions, $(\frac{k}{3})$ three-factor interactions, and so on, in a total of $2^k - 1$ effects. By applying full factorial a decision matrix is obtained for the k effects, considering two levels: (-) representing the minimum values and (+) representing maximum values. Table 2 exemplifies the decision matrix for 3 factors (x_1, x_2, x_3), considering a 2^k factorial design. The n^k factorial design considers n levels of the criteria. In the path selection problem with 3 paths, the n levels can correspond to the maximum values of the diverse criteria, $max_{p1}(x1)$, $max_{p2}(x1)$, $max_{p3}(x1)$ and so on.

With the results of several experiments, Y (score), the response variable, can be estimated through a regression model, as depicted in Eq. 3, where x_1, x_2 and, x_3 represent effects/criteria, β_0 is the intercept coefficient, β_1, β_2 are effect coefficients and σ is the error estimate. Experiments include the same data for the diverse criteria but with different weight sets.

ANOVA applies regression to formulate a linear model in the form of the Eq. 3 and has associated statistical values that determine the efficiency of the model. The goodness of fit can be assessed by the coefficient of determination R^2, which corresponds to the total variance in response variable (Y) by effects/criteria. Higher values of R^2, close to one, indicate that the model explains almost 100% of the variation in Y due to the effects/criteria and their possible interactions. The F-statistic is also important to assess the variation between groups and within groups. Such groups represent the different experiments. For instance, higher values of F-statistic indicate that mean variation between experiments is greater than variation within experiments. If variation is between experiments, it highlights that the score varies due to the different configured weights.

The proposed methodology to compare MADM techniques includes several steps, as detailed bellow:

Step 1 - Gather data of the different paths for each criteria. Such step can be performed in a controlled way or relying on data collected by others, outside control. In this step n decision matrices $dM_n[m,k]$ are obtained, with m measurements for the n paths with k criteria.

Step 2 - Determine the levels of each criteria for the diverse paths. Levels correspond to the minimum, min_j, and maximum, max_j, for path i in the n overall paths. *LevelMin* corresponds to the minimum level (-) while *LevelMax* corresponds to the maximum level (+), and are determined according to Eq. 4 and Eq. 5, respectively.

$$LevelMin_j = min(dM_1[,j], dM_2[,j], \cdots, dM_n[,j]) \; with \; j = 1, \cdots, k \quad (4)$$

$$LevelMax_j = max(dM_1[,j], dM_2[,j], \cdots, dM_n[,j]) \; with \; j = 1, \cdots, k \quad (5)$$

This step determines the logic of employing 2^k or n^k factorial design. If there are no zeros in both levels, 2^k factorial design can be followed, otherwise n^k factorial design must be employed. Data with zeros can represent issues in ANOVA, such as outliers. With a 2^k factorial design levels correspond to the vectors $LevelMin_j$ and $LevelMax_j$. In a n^k factorial design, levels for criteria j are based on the maximum (+) values for the n paths, assuming maximum values are different from zero.

$$lMax_j = \left[max(dM_1[,j]), \cdots, max(dM_n[,j])\right] \; with \; j = 1, \cdots, k \quad (6)$$

Step 3 - Specify weights sets for the different z experiments. Each j criterion in the k criteria has associated a weight. $dW_{sets}[z,k]$, the matrix with weight sets is determined for the z experiments. Weights define how important a criterion is over another, tailoring the final ranking determined by MADM techniques.

Step 4 - Determine factorial design matrix $dF[a,k]$, with a relying on the factorial design, $a = 2^k$ or $a = n^k$. For instance, Table 2 depicts the combinations of three criteria under 2^k factorial design, resulting in $a = 8$, $dF[8,3]$.

Step 5 - Run MADM technique for the full set of factors specified in the $dF[a,k]$ matrix with the respective weight sets in the $dW_{sets}[z,k]$ matrix. Experiments lead to scores, which are combined with the full set of factors to form the input matrix $dI[a, k+z]$ as illustrated in Matrix 7, where $level_{a,k}$ holds the minimum or maximum values.

$$dI[a,k+z] = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ a \end{array} \begin{array}{c} \begin{array}{ccccccc} k_1 & \cdots & k_k & z1 & z2 & \cdots & z_z \end{array} \\ \begin{bmatrix} level_{1,1} & \cdots & level_{1,k} & Score_{1,k+1} & Score_{1,k+2} & \cdots & Score_{1,k+z} \\ level_{2,1} & \cdots & level_{2,k} & Score_{2,k+1} & Score_{2,k+2} & \cdots & Score_{2,k+z} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ level_{a,1} & \cdots & level_{a,k} & Score_{a,k+1} & Score_{a,k+2} & \cdots & Score_{a,k+z} \end{bmatrix} \end{array} \tag{7}$$

Step 6 - Perform ANOVA where the response variable is $Y = Score$ determined by MADM techniques depending on the diverse covariates (k criteria). The initial linear model must include all the covariates and their possible interactions, as exemplified in Table 2 and Eq. 3 for 3 covariates (x_1, x_2, x_3). Interactions are important as the values of one criterion might be related with the values of other criteria. For instance, the score, besides being based on bandwidth, round trip time and jitter can be based on a relation between these parameters. We stress on interactions between criteria, as they can be typical in path selection problems. For instance, higher bandwidths have associated lower RTT, as well as lower jitter values.

Step 7 - Reformulate linear model by including only the effects that are significant, those with *p-value* < 0.05. Run ANOVA with the reformulated model and validate if assumptions for ANOVA models are fulfilled. Namely the model must comply with normality, homogeneity and independence assumptions [13]. Normality assumes that under the same conditions, the observations are normally distributed for each value of X (recall Eq. 3). Homogeneity assumes that the variance for all X values is the same. Independence means that Y values of one observation (X_i) should not influence the Y values for other observations. In DoE, with the factorial design, the independence assumption is assured. The normality assumption can be checked via histograms, where bars must follow the trend of the normal curve. Homogeneity can be checked by plotting the residuals versus the fitted models. If the model complies with normality and homogeneity assumptions, statistical analysis of the regression model must be performed as detailed in the next step.

Step 8 - Analyse the model regarding its completeness, if all the criteria is included, as well as interactions. The analysis must also rely on coefficient of determination, R^2 that assesses how the model explains the variance of Y (score) and F-statistic that complements R^2 in the sense that it measures if variance is inside experiments or between experiments. F-statistic assesses how a MADM technique deals with weights. Higher values of R^2 (close to one) and higher values of F-statistic are preferred. In addition, the significance of the effects and interactions must be considered. Significant effects indicate strong contribution to the score.

Next section presents examples of the TOPSIS, DiA and MeTH evaluation using the evaluation methodology herein proposed and publicly available in [12].

5 Evaluation

The proposed methodology has been applied in two distinct scenarios: *Dropbox* and *Heterogenous* scenarios, which are describe bellow. These scenarios use the same criteria for benefits and costs. Benefits include security (Sec), coverage (Cov) in meters, and bandwidth (BW) in Mb/s. Costs include round trip time (RTT) in milliseconds, Jitter in milliseconds and packet loss (Loss) criteria common in the path selection problems [2].

5.1 Dropbox Scenario

The Dropbox scenario considers a cloud environment where Dropbox services [16] were evaluated. The evaluation of this scenario uses data collected from TCP applications in a university campus, accessing Dropbox facilities. The collected traces contain application network performance values, such as RTT, jitter, retransmissions and duplicates. The evaluation considers a multihomed node with four distinct paths for a Dropbox service. In addition, the collection of data was beyond our control, since data acquisition was performed by Drago et al. [16]. The wireless environment is configured as follows: one path is set according to the IEEE 802.11n and the remaining are configured according to the IEEE 802.11g standard. Moreover, the different paths are configured with different security values, to simulate open networks and networks with security mechanisms. The values used in the evaluation are included in the paper to allow the reproduction of results.

Table 3. Levels of each criteria for the different paths in Dropbox scenario. Levels are represented in the form of min;max.

Paths	Benefits Criteria			Costs Criteria		
	(Sec)	(Cov)	(BW)	(Jitter)	(RTT)	(Loss)
P1	1; 7	0; 250	0; 300	0.20; 575.31	62.48; 171.79	0; 0.40
P2	1; 7	0; 100	0; 54	1.5; 999.1531	46.32; 166.27	0; 0.11
P3	1; 3	0; 100	0; 54	0.20; 10105.49	75.35; 5141.21	0; 0
P4	1; 5	0; 100	0; 54	0; 1126.61	0; 259.78	0; 0.18

5.2 Heterogeneous Scenario

The Heterogenous scenario comprises a multihomed node with three available paths, provided through a wired link (IEEE 802.3ab) and two wireless links, namely IEEE 802.11n and IEEE 802.16e. This scenario was under our control and includes data acquired during several weeks. To collect criteria values, the OWAMP protocol [17] was used, since it allows to gather values according to standardized recommendations from IETF. Owping [18] and bwctl [19] tools were employed, as these implement OWAMP protocol and enable an accurate data acquisition of RTT, jitter, loss and bandwidth, respectively. The clock of machines was synchronized using Network Time Protocol (NTP), to meet the requirements of OWAMP protocol.

Table 4. Levels of each criteria for the different paths in Heterogenous scenario. Levels are represented in the form of min;max.

Paths	Benefits Criteria			Costs Criteria		
	(Sec)	(Cov)	(BW)	(Jitter)	(RTT)	(Loss)
P1	1; 7	0; 54000	0.8821144; 16.81217	0.0; 312.0	0.0; 202.7	0; 0.67
P2	1; 7	0; 250	32.27258; 56.85376	0.1; 6.4	1.1; 21.6	0; 0
P3	1; 7	0; 100	89.99288; 91.26333	0.0; 3.5	0.2; 21.2	0; 0

5.3 Methodology

The different experiments, in light of DoE, were based on ranking determination with different criteria weights. Weights, for both scenarios, were organized in sets to include a full representation of the possible and most representative combinations $dW_{sets}[z, k]$. Table 5 depicts the different combinations of benefits and costs weights, for the $z = 16$ experiments.

Table 5. Configured weights. Weigths sets have been configured regarding the possible and most representative combinations.

Set	W_{Sec}	W_{Cov}	W_{BW}	W_{Jitter}	W_{RTT}	W_{Loss}
1	0.33	0.33	0.33	0.33	0.33	0.33
2	0.33	0.33	0.33	0.6	0.2	0.2
3	0.33	0.33	0.33	0.2	0.6	0.2
4	0.33	0.33	0.33	0.2	0.2	0.6
5	0.6	0.2	0.2	0.33	0.33	0.33
6	0.6	0.2	0.2	0.6	0.2	0.2
7	0.6	0.2	0.2	0.2	0.6	0.2
8	0.6	0.2	0.2	0.2	0.2	0.6
9	0.2	0.6	0.2	0.33	0.33	0.33
10	0.2	0.6	0.2	0.6	0.2	0.2
11	0.2	0.6	0.2	0.2	0.6	0.2
12	0.2	0.6	0.2	0.2	0.2	0.6
13	0.2	0.2	0.6	0.33	0.33	0.33
14	0.2	0.2	0.6	0.6	0.2	0.2
15	0.2	0.2	0.6	0.2	0.6	0.2
16	0.2	0.2	0.6	0.2	0.2	0.6

The n^k factorial design was chosen, as many parameters had values of zeros in both scenarios. The factorial design matrices rely on maximum values for each criteria of the distinct paths, depicted in Table 3 and Table 4 for Dropbox and Heterogenous scenarios, respectively. Indeed the matrices for these scenarios were $dF_{Drop}[4^6, 6]$ and $dF_{Het}[3^6, 6]$. The input matrix $dI[a, k + z]$ for ANOVA considers the defined experiments (Table 5) and factorial design matrices. In this evaluation, $dI_{Drop}[4^6, 6+16]$ and $dI_{Het}[3^6, 6+16]$ matrices were set for Dropbox and Heterogenous scenarios, respectively.

6 Results and Discussion

This section presents and discusses the results achieved with the evaluation performed. All the evaluation has been performed using R-project [20], and models are compared using model completeness, effects significance, R^2 and F-statistics metrics. The beta terms of the ANOVA regression model (recall Eq. 3) are not specified in the models obtained to simplify comparison between MADM techniques.

6.1 Dropbox Scenario

$$Y_{lmTOP} = BW + RTT + Jitter + Loss + Cov \tag{8}$$

The model obtained by TOPSIS (lmTOP) using the methodology presented in this paper includes all the criteria, and is specified according to Eq. 8. DiA also results in the same model. This model, lmTOP, does not include any interactions, and defines score as a function of bandwidth, RTT, jitter, loss and coverage (e.g., all criteria). The model is not fully complete, as interactions are not detected.

$$\begin{aligned} Y_{lmMeth} = {} & BW + RTT + Jitter + Loss + Cov + \text{BW:Cov} + \text{BW:RTT:Cov}+ \\ & \text{BW:Jitter:Cov} + \text{BW:Loss:Cov} + \text{BW:RTT:Jitter:Cov}+ \\ & \text{BW:RTT:Loss:Cov} + \text{BW:Jitter:Loss:Cov} \end{aligned} \tag{9}$$

MeTH, our proposed MADM technique, outputs a different model (lmMeTH) and besides including all the criteria, it also includes interactions between them, as per Eq. 9. lmMeTH demonstrates that criteria has relations, and can be considered as a complete model, in comparison to lmTOP, since criteria and respective interactions are included.

Table 6. Results of Dropbox

method	model	signif	interactions	R^2	F-statistic
TOPSIS	lmTOP	yes	no	0.5274	14624.2727
DiA	lmTOP	yes	no	0.4452	10518.2098
MeTH	lmTOP	yes	no	0.7240	34376.5185
TOPSIS	lmMeth	no	yes	0.5274	6093.3300
DiA	lmMeth	no	yes	0.4452	4382.2384
MeTH	lmMeth	yes	yes	0.7413	15649.5765

Table 6 summarizes the statistical values obtained in the dropbox scenario. With lmTOP model, the TOPSIS technique can explain $\approx$ 53% of variation of data, since $R^2 = 0.5274$. DiA is only able to explain $\approx$ 45% of the variance, nonetheless, MeTH is able to explain $\approx$ 72% of the score variance. Recall that values close to 1 are fully explained by the model. The F-statistic also reports higher values in MeTH, namely, $F_5 = 34376.5185, p < 0.005$, which means that the variation of score is higher between experiments than in experiments. Thus, MeTH considers more properly the weights sets configured on the diverse experiments, when compared to TOPSIS or DiA approaches. F-statistic for TOPSIS follows MeTH model, namely, $F_5 = 14624.2727, p < 0.005$. MeTH within the lmTOP model is the technique with more satisfactory statistical values, followed by TOPSIS. The main issue with this model is that is lacks interactions, that is, it does not consider the relations between criteria (e.g., if one criterion increases the other criterion will increase as well, or vice-versa). In this context the lmMeth model is more complete and is obtained with MetH, mainly due to the distance function (recall Table 1) that correlates data of the distinct paths. In the lmMeTH model, MeTH presents, again, the best performance regarding statistical values, since R^2 is higher and F-statistic is also higher $F_{12} = 15649.5765, p < 0.005$, in comparison to TOPSIS and DiA results. In addition, when comparing both models, (the main difference relies on the interactions), lmMeTH model with MeTH technique is able to explain $\approx 74\%$ of score variance, against the $\approx 72\%$ of lmTOP. F-statistic in the lmMeTH model is not higher in comparison to lmTOP model, but the 14 terms in Eq 9 against 5 in Eq. 8 justify such fact. It is also relevant to point out for the lmMeTH model that with TOPSIS and DiA not all the effects are significant, which means that these techniques are not able to find relations between criteria.

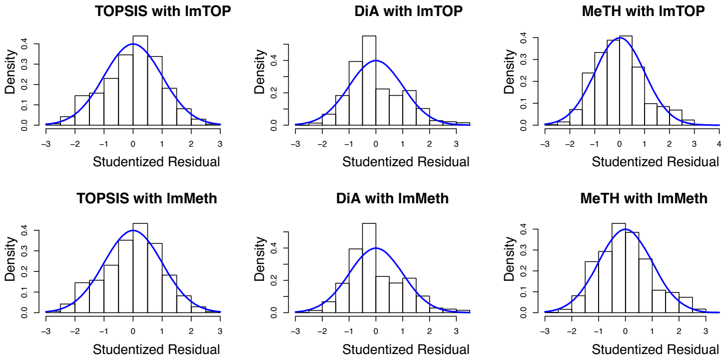

Fig. 1. Normality for analysed MADM in Dropbox scenario

According to the Step 7 of the evaluation methodology, assumptions for ANOVA need to be checked, in order to guarantee that the results have higher confidence. Fig. 1 depicts a graphical test to assess normality of lmTOP and lmMeTH models within the different MADM techniques evaluated, relying on histograms and normal curve. At a first glance, DiA is the only technique violating normality in lmTOP and lmMeTH models, which may indicate that the distance or scoring functions (recall Table 1) perform transformations that break such assumption. MeTH and TOPSIS are able to present normality in the scoring for both models. In these techniques bars follow the trend of the normal curve (pictured in blue), that is, there is a pattern of ascending and descending "stairs", without any exception.

The results in this scenario demonstrate that the distance and associated score functions lead to different results, mainly in terms of supporting interactions and statistical importance.

6.2 Heterogenous Scenario

The model obtained by TOPSIS (lmTOP) using the methodology presented in this paper for heterogenous scenario includes all the criteria, and is similar to Dropbox scenario. The DiA technique has the same model. Despite having fewer data (3^6 rows when compared to 4^6 in Dropbox scenario), MeTH is also able to provide interactions, as per Eq. 10. In particular, the lmMeTH model in the heterogenous scenario is more complete with 14 effects, in comparison to the Dropbox model, which has only 12 effects, compare Eq. 10 and Eq. 9.

$$\begin{aligned} Y_{lmMeTH} &= BW + RTT + Jitter + Loss + Cov + \text{BW:Jitter} + \text{BW:Loss}+ \\ &\text{BW:Cov} + \text{BW:RTT:Cov} + \text{BW:Jitter:Cov} + \text{BW:Loss:Cov} + \\ &\text{BW:RTT:Jitter:Cov} + \text{BW:RTT:Loss:Cov} + \text{BW:Jitter:Loss:Cov} \end{aligned} \tag{10}$$

Table 7. Results of Heterogenous

method	model	signif	interactions	R^2	F-statistic
TOPSIS	lmTOP	yes	no	0.5352	2684.5152
DiA	lmTOP	yes	no	0.4313	1768.3257
MeTH	lmTOP	yes	no	0.7514	7046.4885
TOPSIS	lmMeth	no	yes	0.5352	958.0181
DiA	lmMeth	no	yes	0.4313	631.0595
MeTH	lmMeth	yes	yes	0.7963	3253.4246

Table 7 summarizes the statistical values obtained in the heterogenous scenario. With lmTOP model, the TOPSIS technique can explain $\approx 53\%$ of variation of data, since $R^2 = 0.5352$. DiA is only able to explain $\approx 43\%$ of the variance. MeTH is able to explain $\approx 75\%$ of the score variance. The F-statistic also reports higher values in MeTH, namely, $F_5 = 7046.4885, p < 0.005$, what means that the variation of score is higher between experiments than inside the respective experiment. TOPSIS follows the MeTH performance in terms of F-statistic. This indicates that DiA is the technique that less impacts scoring regarding weights configurations. Considering weights as applications preferences (i.e., one might prefer more security other prefers higher bandwidths), DiA may not provide a scoring adapted to the requirements of distinct applications. Regarding the lmMeTH model, MeTH technique is able to explain $\approx 80\%$ of score variation. Thus, contrasts with TOPSIS and DiA techniques that do not increment values of R^2 in the lmMeTH.

Fig. 2 depicts a graphical test to assess normality of lmTOP and lmMeTH models within the different techniques, relying on histograms and normal curve. With the lmTOP model, normality is supported only by MeTH technique, as bars follow the trend of the normal curve (pictured in blue). DiA and TOPSIS present some exceptions to the normality assumption.

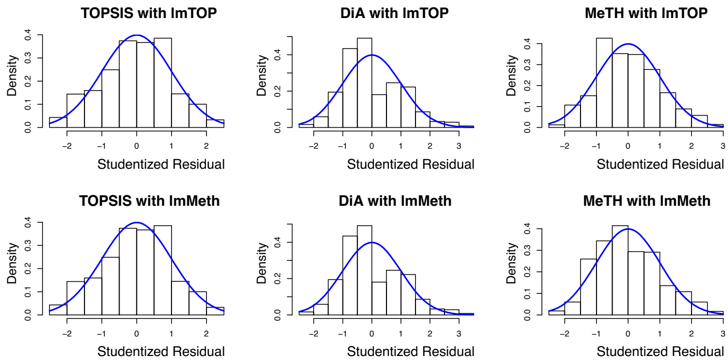

Fig. 2. Normality for analysed MADM in heterogenous scenario

The values in the heterogenous scenario regarding R^2 are higher for MeTH and TOPSIS techniques in comparison to the Dropbox scenario. The reason for such performance increase relies on the complexity of the scenario, with 3 paths against 4 paths. This fact indicates that TOPSIS and MeTH adapt more efficiently to the problem size in comparison to DiA. In fact, MeTH is able to explain $\approx 80\%$ of the values of score with all the criteria and respective interactions.

7 Conclusion

This paper specified MeTH as a MADM technique best suited for path selection problems, as it does not assume distance to be a straight line or grid-based, but, instead a composite function between benefits and costs criteria. In MeTH distance is considered for ideal values and relevant ranges. The evaluation methodology relying on statistical analysis, specified in this paper, has the advantage of being easily reproducible. Results based on data from controlled and uncontrolled path selection scenarios and with different number of paths, demonstrate that MeTH is able to perform optimal path selection more efficiently regarding the configured weights and multihoming nodes configurations.

Acknowledgment. The first author acknowledges the support of the PhD grant SFRH/BD/61256/2009 from Ministério da Ciência, Tecnologia e Ensino Superior, FCT, Portugal. This work is supported by CoFIMOM project PTDC/EIA-EIA/116173/2009 and TRONE project CMU- PT/RNQ/0015/2009.

References

1. Hou, R., Lui, K., Leung, K., Baker, F.: Approximation Algorithm for QoS Routing With Multiple Additive Constraints. In: ICC 2009, pp. 1–5. IEEE (2009)
2. Li, X., Mehani, O., Agüero, R., Boreli, R., Zaki, Y., Toseef, U.: Evaluating User-Centric Multihomed Flow Management for Mobile Devices in Simulated Heterogeneous Networks. In: Timm-Giel, A., Strassner, J., Agüero, R., Sargento, S., Pentikousis, K. (eds.) MONAMI 2012. LNICST, vol. 58, pp. 84–98. Springer, Heidelberg (2013)

3. Choque, J., Agüero, R., Muñoz, L.: Optimum Selection of Access Networks Within Heterogeneous Wireless Environments Based on Linear Programming Techniques. Mob. Netw. Appl. 16(4), 412–423 (2011)
4. Kucharzak, M., Walkowiak, K.: On modelling of fair throughput allocation in overlay multicast networks. In: Balandin, S., Koucheryavy, Y., Hu, H. (eds.) NEW2AN 2011 and ruSMART 2011. LNCS, vol. 6869, pp. 529–540. Springer, Heidelberg (2011)
5. Behzadian, M., Otaghsara, S.K., Yazdani, M., Ignatius, J.: A State-of The-Art Survey of TOPSIS Applications. Expert Systems with Applications 39(17), 13051–13069 (2012)
6. Sun, C.-C.: A Performance Evaluation Model by Integrating Fuzzy AHP and Fuzzy TOPSIS Methods. Expert Systems with Applications 37(12), 7745–7754 (2010)
7. Nacef, B., Montavont, N.: A Generic End-Host Mechanism for Path Selection and Flow Distribution. In: Proc. PIMRC 2008. IEEE (September 2008)
8. Tran, P.N., Boukhatem, N.: The Distance to The Ideal Alternative (DiA) Algorithm for Interface Selection in Heterogeneous Wireless Networks. In: Proc. MobiWac 2008 (October 2008)
9. Lahby, M., Cherkaoui, L., Adib, A.: New Optimized Network Selection Decision in Heterogeneous Wireless Networks. International Journal of Computer Applications 54(16) (September 2012)
10. Ren, L., Zhang, Y., Wang, Y., Sun, Z.: Comparative Analysis of a Novel M-TOPSIS Method and TOPSIS. Applied Mathematics Research eXpress 2007, 1–10 (2010)
11. Chakraborty, S., Yeh, C.-H.: A Simulation Comparison of Normalization Procedures for TOPSIS. In: CIE 2009, pp. 1815–1820. IEEE (2009)
12. Sousa, B.: Evaluating MADM techniques, `http://mcoa.dei.uc.pt/doe/` (last Visit: May 12, 2013)
13. Montgomery, D.C.: Design and Analysis of Experiments. John Wiley & Sons (2008)
14. İç, Y.T.: An Experimental Design Approach Using TOPSIS Method for The Selection of Computer-Integrated Manufacturing Technologies. Robot. Comput.-Integr. Manuf. 28(2), 245–256 (2012)
15. Sandanayake, Y., Oduoza, C., Proverbs, D.: A systematic modelling and simulation approach for jit performance optimisation. Robotics and Computer-Integrated Manufacturing 24(6), 735–743 (2008)
16. Drago, I., Mellia, M., Munafò, M.M., Sperotto, A., Sadre, R., Pras, A.: Inside dropbox: Understanding personal cloud storage services. In: Proceedings of the 12th ACM Internet Measurement Conference, IMC 2012 (2012)
17. Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., Zekauskas, M.: A One-way Active Measurement Protocol (OWAMP). IETF Request for Comments: 4656 (September 2006)
18. Boote, J., Karp, A.: Internet 2 One-Way Ping (OWAMP) (2013), `http://www.internet2.edu/performance/owamp/index.html` (last visit: May 12, 2013)
19. Boote, J., Brown, A.: Internet 2 Bandwidth Test Controller (BWCTL) (2013), `http://www.internet2.edu/performance/bwctl/index.html` (last visit: May 12, 2013)
20. R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2012), ISBN 3-900051-07-0