ORIGINAL RESEARCH

Hierarchical Bayesian approach for improving weights for solving multi-objective route optimization problem

Romit S. Beed[1] · **Sunita Sarkar**[2] · **Arindam Roy**[3]

Abstract The weighted sum method is a simple and widely used technique that scalarizes multiple conflicting objectives into a single objective function. It suffers from the problem of determining the appropriate weights corresponding to the objectives. This paper proposes a novel Hierarchical Bayesian model based on multinomial distribution and Dirichlet prior to refine the weights for solving such multi-objective route optimization problems. The model and methodologies revolve around data obtained from a small-scale pilot survey. The method aims at improving the existing methods of weight determination in the field of Intelligent Transport Systems as data driven choice of weights through appropriate probabilistic modelling ensures, on an average, much reliable results than non-probabilistic techniques. Application of this model and methodologies to simulated as well as real data sets revealed quite encouraging performances with respect to stabilizing the estimates of weights. Generation of weights using the proposed Bayesian methodology can be used to develop a bona-fide Bayesian posterior distribution for the optima, thus properly and coherently quantifying the uncertainty about the optima.

✉ Sunita Sarkar
sunitasarkar@rediffmail.com

Romit S. Beed
rbeed@yahoo.com

Arindam Roy
arindam_roy74@rediffmail.com

[1] Dept. of Computer Sc., St. Xavier's College (Autonomous), Kolkata, India

[2] Dept. of Computer Sc. and Engineering, Assam University, Silchar, India

[3] Dept. of Computer Sc., Assam University, Silchar, India

1 Introduction

Majority of the real-world complexities generally involve optimizing multiple conflicting objectives. Simply obtaining a solution for their least values concurrently for all the objectives does not guarantee correctness; hence a compromise needs to be made. As these optimization formulations involve multiple objectives, the objective function is formulated as a vector and it is treated as a vector optimization or a multi-objective optimization problem (MOOP) [1]. A MOOP involving multiple, conflicting objectives may be combined into a single-objective scalar function. This approach is named as the weighted-sum method. This is an a priori method established on the "linear aggregation of functions" principle.

The method is alternatively named as Single Objective Evolutionary Algorithm (SOEA). By definition, the weighted-sum method reduces to a positively weighted convex sum of the objectives, as follows:

$$\text{Min} \sum_{i=1}^{n} w_i f_i(x), where \sum_{i=1}^{n} w_i = 1; w_i > 0 \forall i \tag{1}$$

Minimization of this single-objective function is expected to give an efficient solution for the original multi-objective problem. The process involves scalarizing the conflicting objectives into a single objective function. There are various scalarization techniques which have been proposed in the past.

1332

Int. j. inf. tecnol. (August 2021) 13(4):1331–1341

Accounting for ambiguity is important when there are restrictions on data which leads to inaccurate interpretation about choices, sensitivities, and other behavioral characteristics. Bayesian analysis, grounded on Bayes' theorem, is an instrument that assists in this accounting procedure. Although theoretically lucid, it is hard to apply in different real-world problems mainly due to the need of refined estimates. However, this problem was resolved through the advancement of mathematical techniques of iterative calculations largely based on Markov chain Monte Carlo (MCMC) methods. Bayes models free researchers from computational constraints by developing more realistic models of user behavior and decision making by integrating a hierarchical model with Bayesian estimates. The several sub-models are hierarchically integrated using the Bayes theorem which manages the uncertainty aspect, hence the name Bayesian Hierarchical model.

Bayesian Learning is based on the simple logic that one can achieve better choices by including their recent interpretations and beliefs obtained through previous knowledge and experience. Bayesian learning is also effective where frequentist statistics is not implementable. It possesses supplementary features like iterative upgrade of the posterior while analyzing a hypothesis to assess the parameters of a machine learning model. This promising learning technique is based on Bayes' theorem to obtain the conditional probability of a hypotheses which is in turn based on certain previous knowledge. Majority of the everyday problems does encompass ambiguity and incremental wisdom, therefore making Bayesian learning more applicable to solve such problems. The Bayesian approach incorporates past belief and gradually refines the prior probabilities based on fresh evidence.

Bayesian data analysis is a process of mathematically modelling data and assigning credibility to individual parameters that are steady with the data and with previous experience. Incorporation of prior belief and experience gives Bayesian learning an edge over frequentist statistics. The Bayesian technique offers greater flexibility in system modeling based on available data. It also provides clarity of demonstrating parameter uncertainty which is accurately inferred and there exists no requirement for building sampling distributions from supporting null hypotheses. Frequentist approaches for calculating ambiguity are inconsistent and tough, whereas Bayesian approaches are essentially intended to offer distinct demonstrations of ambiguity. Frequentist techniques are comparatively more cumbersome for building confidence intervals on parameter. Though there are several advantages of the Bayesian approach, there is an inherent need for an adequately greater set of trials or assigning a confidence to the established hypothesis.

Bayesian reasoning ensures the reorganization of credibility over likelihoods while incorporating fresh data. The main objective of Bayesian estimation is to obtain the most reliable parameter values for the model and this estimation offers a total distribution of credibility over the space of parameter values, not simply one "best" value. The crux of Bayesian estimation is to correctly define how ambiguity changes when fresh data are considered. It is seen that at times that the parameters have significant dependencies on one another. This ordering of dependencies among parameters illustrates a hierarchical model. A hierarchical model specifies dependencies among parameters in an ordered manner based on the semantics of the parameters. Considering data from entities within sets is a salient hierarchical model application. A hierarchical model has the flexibility to possess parameters for every entity that define every discrete entity's characteristics, and the distribution of different parameters inside a set is exhibited by a higher order distribution with its own parameters that define the tendency of the set. The entity level and set level parameters are assessed concurrently. The hierarchical approach is beneficial as it does not merge the entity's data together thereby preventing dilution of trends within entities. To summarize, hierarchical models have parameters that implicitly define the data at several levels and link data within & across levels.

The main objective of this work is to provide a flexible Bayesian nonparametric approach that effectively incorporates various levels of uncertainties in determining the weights, whereas the existing approaches focused on improving the solution set without paying much attention to the appropriate choice of weights. The novelty of the proposed work lies in its coherent approach of considering all the levels of uncertainty for determining the weights and thereby improving the solution set. It is to be noted that the novelty also lies in the technique of generating the weights using the proposed Bayesian methodology which helps develop a bona-fide Bayesian posterior distribution for the optima, thus accurately quantifying the uncertainty about the optima. This work, formulated on the Bayesian framework, considers prior knowledge (realized through prior distribution) of the relative significance of the conflicting objectives while generating the weights. Unlike the existing techniques, this method can be applied with convenience to handle any number of objectives. As this method yields a posterior probability distribution over the weights, the stochastically generated weight vectors can be used to obtain the points on the Pareto front with less computational complications. Yet another novelty of this work is its ability to work with small data samples as it is not feasible at times to have a large sample dataset due to time and associated cost factors.

2 Literature survey

Zadeh popularized the weighted sum technique as a classical approach for solving such problems [2]. This method, as the name suggests, scalarizes a set of conflicting objective functions, by pre-multiplying each of the objective function by predefined weights. The ε-constraints method, introduced by Hamines et al. [3], focused on minimizing the most significant objective function fs(x). Another popular scalarizing technique is the goal attainment technique [4] where the goals are stated for individual objective function fs(x) and the process aims to reduce the overall deviation from the goals. The hierarchical approach [5] and the weighted metrics technique [6] are two more techniques for solving such problems. However, the weighted sum method has gained the most popularity among these due to its simplicity. Although much research has been devoted to the development of different algorithms improving the solution set in multi-objective optimization problems using weighted sum technique, to date, a comprehensive model generating the weights using various sources of uncertainties seem to be lacking. The weighted sum (WS) technique, a commonly used scalarizing technique in multi-objective algorithms, has distinctive advantages of greater search efficiency and easier computational capabilities. Nevertheless, it is frequently critiqued for its inability to predict the logic behind the weight selection as well as its incompetence to deal with nonconvex problems.

It was suggested by Steuer [7] that the weights should scientifically determine the decision-maker's preference for a particular objective. Das and Dennis [8] offered a graphical explanation of this technique to elucidate few of its drawbacks. The delusion between the hypothetical and the realistic interpretation of the weights for the conflicting objectives made the weight selection process quite inefficient. Various approaches have been suggested for weight selection—Yoon and Hwang [9] suggested a ranking technique whereby the objectives are ranked based on their significance. The most important objective received the largest weight with gradual decrease in weights to lesser important objectives. It was quite similar to the categorization technique in which the conflicting objectives were grouped according to their varying degree of importance. Saaty [10] proposed an eigenvalue process for attaining weights, where n (n−1)/2 pair-wise evaluations were made between these objective functions to generate a comparison matrix; from this matrix the eigenvalues yielded the weights. Wierzbicki [11] proposed a method for generating weights, where the comparative significance of the objective functions is vague, based on the utopia and the aspiration points. Another method for weight determination was proposed based on fuzzy set theory by Rao and Roy [12]. Though various techniques exist for weight determination, just the selection of the weights may not necessarily generate a feasible solution. New weights may have to be considered and the process may have to be executed again. It was thus suggested by Messac [13] that weights should be functions of the objectives and not simply constants to simulate a task precisely. According to him, the weights must address the issues related to both scaling and relative preference of the objective functions to reflect the preference appropriately.

Selection of appropriate weights leads to an algorithm's better performance. Timothy Ward Athan [14] proposed a quasi-random weighted criteria system that produces weights covering the Pareto set consistently. The method is based on random probability distribution and involves a large number of computations. Gennert and Yuille [15] proposed a nonlinear weight determination algorithm where an optimal point is obtained that is not in the vicinity of the extreme points. Although a lot of work is available in the literature regarding systematic selection of the weights in solving a Multi-Objective Optimization problem, till date a comprehensive data driven technique determining weights reflecting the relative importance of the conflicting objectives is lacking. Many authors including Das and Dennis [8] have shown that choosing weights uniformly over (0,1) does not guarantee uniform spread of Pareto points on the Pareto front. In many cases it has been observed that the points obtained using a uniform generation of weights are found to be clustered in certain regions of the Pareto set. In their subsequent work, Das and Dennis [16] have proposed a technique based on Normal-Boundary Intersection, of obtaining an even spread of Pareto points. Like many others, their method prioritizes the solution set while deciding upon the choice of weights. J. G Lin [17] point towards the scarcity of the number of Pareto Optimal solutions obtained by the existing methods, in addition to some of the solutions coinciding with extreme points. Lin has proposed a method of solving multi-objective optimization problems by transforming them into Single-objective optimization problem, by changing one of the multi-objectives to proper equality constraints using Lagrange multiplier.

Marler and Arora [18] explicated that the weighted sum method is a simple method that delivers a linear estimate of the preference function and need not necessarily reflect one's primary preferences. It is essentially inept of including multifaceted preferences. In spite of determining satisfactory weights a priori, the end solution need not precisely display original preferences. Rui Wang et al. [19] proposed a multi-objective decomposition-based Evolutionary algorithm based on local application of the weighted sum technique. They proposed that the optimal

result for each of the search routes is obtained from amongst its adjoining results. Experimental outcomes confirmed that the MOEA/D-LWS outperformed the remaining algorithms for majority of the cases. Zhang proposed a dynamic weighted sum (DWS) technique [20] to methodically alter the weights of individual conflicting objectives for solving multi-objective optimization problems (MOO). He studied the search effect of the different dynamic weighted aggregations namely bang-bang, linear, sinusoidal and random weighted aggregations.

Jaini and Utyuzhnikov [21] proposed a compromise grading method in a fuzzy multi-criteria choice-making system. The fuzzy quantities symbolize the vague weights of each of the conflicting objectives. The authors have designed a fuzzy trade-off grading technique to rank alternatives by awarding the smallest compromise solution as the finest choice. Most of the work in the available literature has focused on fixing the weights based on some prior beliefs or information. The focus of the existing methods is towards refining the distribution of Pareto solutions provided by the WS technique [22], with less emphasis on the stability and appropriateness of the choice of weights for precise representation of the conflicting objectives.

In contrast to the objective of weight determination of the existing works, which aimed at choosing the set of weights which stabilizes the solution set [23–25], this work proposes to frame a model which determines a much stable set of weights in comparison to that obtained deterministically. The criticisms of the existing methodologies for determination of weights have motivated this work and to propose the Bayesian model based on multinomial and Dirichlet priors. As per the authors' existing knowledge, this work is first of its kind since none of the earlier works had this motivation of searching for stability in weights. Unlike the frequentist approach, the Bayesian modelling is based on treating the uncertainties in the parameters probabilistically. The Frequentist methodologies, not considering prior probabilities, come up with estimates based mostly on the maximum likelihood or confidence intervals while Bayesians, have a complete posterior distribution over the possible parameter values. This allows them to account for the uncertainty in the estimate by integrating the entire distribution, and not just the most likely value.

This work has been based on the Bayesian framework as one can coherently consider any prior knowledge (reflected through the prior distribution) about the relative importance of the conflicting objectives to generate the weights. This prior knowledge (distribution) will then be updated using data from the sample using Bayesian paradigm. The sample data obtained from the pilot survey makes the probability distribution narrower around the parameter's

true but unknown values. The hierarchical Bayesian model has been so developed as to reflect the relative importance of the conflicting objectives through the respective weights, which were stochastically estimated, based on the data obtained from a pilot survey for the given purpose.

3 Statistical prerequisite

The proposed hierarchical Bayesian model methodology for generating weights is based on the Multinomial and Dirichlet distributions as priors, which are conjugate to each other. A brief discussion on the statistical prerequisites is given in this section.

3.1 Bayesian approach

Unlike frequentist approach which does not quantify the uncertainty in fixed but unknown values of the parameters, Bayesian approach, defines probability distributions over possible values of a parameter. Let x denote the data and θ be the parameter of interest which is unknown. Let $\theta \in \Theta$ be the parametric space. Under Bayesian approach one can quantify the prior belief about θ by defining a prior probability distribution over Θ, the set of possible values of θ. The newly collected data makes the probability distribution over Θ narrower by updating the prior distribution to posterior distribution (updated) θ of using Bayes' theorem which states that

$$P(\theta|\text{Data}) = \frac{P(\text{Data}|\theta)P(\theta)}{P(\text{Data})} \qquad (2)$$

where P(Data|θ) is called the likelihood & P(θ|Data) is the posterior distribution of the parameter θ.

3.2 Hierarchical Bayesian model

A model in which the prior distribution of some of the model parameters depend on some other unknown parameters, which are in turn modelled as random variables following some other distribution is a Hierarchical Bayesian model. The level of hierarchy depends on the context and complexity of the problem. Given the observed data x, suppose x follows f(.|θ), with θ being distributed as a prior π(φ). If the parameters φ can further assumed to be following ζ(υ), the use of hierarchical models ensures a more flexible account of data.

3.3 Multinomial distribution

It is a multivariate generalization of Binomial distribution. Suppose an experiment is conducted such that each trial has k (finite & fixed) mutually exclusive & exhaustive

possible outcomes with probabilities $p_1, p_2, \ldots, p_k$ such that $p_i \geq 0 \forall i = 1(1)k$ and $\sum_{i=1}^{k} p_i = 1$. If X_i be the random variable indicating the number of times category I is observed over n independent trials of the experiment, then the vector $\underline{X} = (X_1, X_2, \ldots, X_k)$ follows a Multinomial Distribution with parameters n and $p_1, p_2, \ldots, p_k$. The probability mass function of the multinomial distribution is:

$$
F(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)
$$
$$
= \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \cdots p_{k-1}^{x_{k-1}} (1 - p_1 - p_{k-1})^{n - x_1 - \ldots - x_{k-1}}
$$
(3)

3.4 Dirichlet distribution

Dirichlet distribution is a multivariate generalization of Beta distribution. Dirichlet distribution of order $(k \geq 2)$ with parameters $\alpha_1, \ldots, \alpha_k > 0$, has the following probability density function:

$$
g(x_1, x_2, \ldots, x_k) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k} x_i^{\alpha_i - 1}
$$
(4)

Here X_i's are continuous random variables with $x_i \geq 0 \forall i$ and $\sum x_i = 1$, that is, the support of Dirichlet distribution is the set of k-dimensional vectors whose entries belong to (0,1) and add up to one. The parameter vector $p_1, p_2, \ldots, p_k$ of the Multinomial distribution has the properties of the x_i's above, as $p = \{p_1, p_2, \cdots, p_k\}$, where $0 \leq p_i \leq 1$ for $i \in [1, k]$ and $\sum p_i = 1$ and hence can be modelled using an appropriate Dirichlet distribution. Dirichlet distribution is a family of continuous probability distribution for a discrete probability distribution with k categories. The usefulness of this method is explained with the help of a realistic example. Considering a company produces six faced dice; though manufacturing processes are precise nowadays, they are still not 100% perfect—if one rolls a randomly selected dice, getting an exact relative frequency of one sixth for the outcomes is difficult due to a slight manufacturing defect. As one can always expect a probability distribution over all possible values, 1, 2, 3, 4, 5 and 6; this probability distribution can be modelled using Dirichlet distribution.

3.5 Conjugate prior

In Bayesian probability theory if posterior and prior probability distributions of the parameter θ belong to the same probability distribution family, the prior is then called a conjugate prior. In other words, in Equation (2) if $P(\theta)$ and $P(\theta|\text{Data})$ are in the same family of distributions, they

are called conjugate distributions. It can be shown that Dirichlet distribution acts as a conjugate prior for Multinomial distribution.

3.6 Proposed methodology

As generation of the weights corresponding to different conflicting objectives in a weighted sum problem is the primary interest, the authors have considered the weights w_i in (1) as the unknown parameters.

Let M be a set of conflicting objectives in the objective space defined as follows:

$$
M = \{f_1(x), f_2(x), \ldots, f_l(x); h_i(x) \leq / \geq 0, i = 1, \ldots, p\}
$$
(5)

where $f_1(x), f_2(x), \ldots, f_l(x)$ are the conflicting objective functions, $h_i(x)$ denotes the set of p constraints. The Weighted Sum method scalarizes the vector objective functions, $f = (f_1(x), f_2(x), \ldots, f_l(x)) \epsilon \mathcal{R}^l$, where $\mathcal{R}^l$ is the $\underline{l}$ dimensional Euclidean space, using the appropriately selected vector of weights $\underline{w} = (w_1, \ldots, w_l) \in \mathcal{R}^l$ such that $w_i > 0$ and $\sum_{i=1}^{l} w_i = 1$.

$$
\theta = \underline{w}' \underline{f} = w_1 f_1 + \cdots + w_l f_l
$$
(6)

It is to be noted that, $\theta \in R$ is a scalar. Without any loss of generality one can assume the objective functions $f_i(x)$ $\forall_i = 1(1)l$ to be normalized.

To determine the weights, suppose one obtains data on the preferences of n individuals regarding the choice of different categories (representing different conflicting objectives) through a planned pilot survey. Individuals may be asked to vote for the single most important category out of a finite number of mutually exclusive and exhaustive set of choices. Let n_i= number of individuals who have voted for category i (i = 1, 2,..., l; representing the ith objective function) in the pilot survey. The multinomial distribution is used for modeling the probability of counts in the different categories (representing the different objective functions), as the individuals vote independently for exactly one of the l categories. Then, $(n_1, n_2, \ldots, n_l) \sim$ Multinomial $(n; w_1, w_2, \ldots, w_l)$, where w_i is the population proportion of individuals who will vote for category i or is the probability that a randomly selected individual votes for ith category. Probability mass function of multinomial distribution is given by

$$
f(n_1, n_2, \ldots, n_l | w_1, w_2, \ldots, w_l)
$$
$$
= \frac{n!}{n_1! n_2! \ldots n_l!} w_1^{n_1} w_2^{n_2} \cdots w_{l-1}^{n_{l-1}} (1 - w_1 - \cdots - w_{l-1})^{n - n_1 - \ldots n_{l-1}}
$$
(7)

As w_i's are continuous random variables, where

1336

Int. j. inf. tecnol. (August 2021) 13(4):1331–1341

$w_i \geq 0 \forall i$, and $\sum w_i = 1$ it can be further assumed that,$(w_1, w_2,...,w_l) \sim$ Dirichlet $(\alpha_1, \alpha_2,...,\alpha_l)$ having the following form of density.

$$g(w1, w2,..., wl) = \frac{\Gamma(\sum_{i=1}^{l} w_i)}{\prod_{i=1}^{l} \Gamma(w_i)} \prod_{i=1}^{l} w_i^{\alpha_i - 1} \qquad (8)$$

Here, Dirichlet distribution, being a distribution over a probability simplex, is most appropriate for modelling $(w_1, w_2,...,w_l)$. Dirichlet distribution is a multivariate generalization of Beta distribution a acts as a conjugate prior to multinomial where $\alpha_1, \alpha_2,...,\alpha_l$ are the concentration parameters such that $\alpha_i > 0 \; \forall \; i= 1 \,(1)\, l$.

The marginal likelihood function is given by,

$$h(n_1, n_2,..., n_l | \alpha_1, \alpha_2,..., \alpha_l) =$$

$$\int_{w1,w2,...,wl}^{\cdot} f(n_1, n_2,..., n_l | w_1, w_2,..., w_l).$$

$$g(w_1, w_2,..., w_l) dw_1 dw_2...dw_l = \qquad (9)$$

$$\frac{\Gamma \sum_{j=1}^{l}(\alpha_j)}{\prod_{j=1}^{l} \Gamma(\alpha_j)} \cdot \frac{n!}{\prod_{j=1}^{l} n_j!} \cdot \frac{\prod_{j=1}^{l} \Gamma(n_j + \alpha_j)}{\Gamma(\sum_{j=1}^{l} \alpha_j + n)}$$

Equation (9) gives the conditional probability of observing the data given $\alpha_1, \alpha_2,...,\alpha_l$. The values of $\alpha_1, \alpha_2,...,\alpha_l$ which maximizes (9) are considered as the estimates. Now it can be shown that $[w_1, w_2,...,w_l | n_1, n_2,...,n_l \sim$ Dirichlet $(\alpha_1+n_1, \alpha_2+n_2 ,..., \alpha_l+n_l)$, i.e. the posterior distribution of the weights given the data, follows Dirichlet distribution with concentration parameters $(\alpha_1+n_1, \alpha_2+n_2,...,\alpha_l+n_l)$.

Posterior expectations of the weights are given by,

$$W_i* = E(w_i | n_1, n_2, n_3) = \frac{\alpha_i + n_i}{\sum_{i=1}^{3}(\alpha_i + n_i)}, \; i = 1, 2, 3 \qquad (10)$$

Hence, estimates of weights can be taken as

$$\hat{W}_i* = \frac{\hat{\alpha}_i + n_i}{\sum_{i=1}^{3}(\hat{\alpha}_i + n_i)} \qquad (11)$$

where $\hat{\alpha}_1, \hat{\alpha}_2$ and $\hat{\alpha}_3$ are the values maximizing (9).

Hence, the objective function gets modified as follows:

$$\text{Minimization of } f = \hat{W}_1 * f_1 + \hat{W}_2 * f_2 + \hat{W}_3 * f_3 \qquad (12)$$

The above modeling technique incorporates the uncertainties in determination of the weights through a Bayesian hierarchical model based on multinomial distribution with Dirichlet prior. As observed ithe existing literature w_i's have been estimated simply by the proportion of preference in the respective categories,

$$\hat{W}_i = p_i = n_i / n; \; i = 1, 2,..., l; \; n = \sum_{i=1}^{l} n_i \qquad (13)$$

The proposed algorithm for weight determination is as follows:

Step 1: Read $n_1,. ....n_l$(votes for individual categories).

Step 2: Calculate *sum* $n = n_{1+}. .... + n_l$.

Step 3: Calculate *Probability Mass Function* $f(n_1, n_2,...,n_l | w_1, w_2,...,w_l)$ where w_i is the population proportion of individuals (unknown) who will vote for category i

Step 4: Calculate *Probability Density Function* $g(w_1, w_2,..., w_l | \alpha_1, \alpha_2,...,\alpha_l)$ w_i's are continuous random variables and $\alpha_1, \alpha_2,...,\alpha_l$ are the concentration parameters.

Step 5: Calculate *Marginal Likelihood Function* $h(n_1, n_2,..., n_l | \alpha_1, \alpha_2,...,\alpha_l)$: The values of $\alpha_1, \alpha_2,...,\alpha_l$ which maximizes this expression are considered as the estimates.

For three variables, **h=**

$$\frac{\text{gamma}(a_1 + a_2 + a_3)}{\text{gamma}(a_1) * \text{gamma}(a_2) * \text{gamma}(a_3)}$$

$$\times \frac{\text{factorial}(n)}{\text{factorial}(n_1) * \text{factorial}(n_2) * \text{factorial}(n_3)}$$

$$\times \frac{\text{gamma}(a_1 + n_1) * \text{gamma}(a_2 + n_2) * \text{gamma}(a_3 + n_3)}{\text{gamma}(a_1 + a_2 + a_3 + 3n)}$$

Step 6: Calculate *Posterior Expectations of weights* W_i*

Step 7: Calculate *Estimates of weights* $\hat{W}_i*$

Step 8: *Exit*

The corresponding MATLAB pseudo-code for obtaining the weights is as follows:

```
nv = input('input sample values [n1 n2 n3]: ');
n = sum(nv);
v = [1, 1, 1];
options = optimset('Display', 'off', 'MaxFunEvals', 100000,
'MaxIter', 100000);
a = fminsearch(@three_var,v,options);
weights = drchrnd(a, 100)
mean_weights = mean(weights());
function r = drchrnd(a,n)
p = length(a);
r = gamrnd(repmat(a,n,1),1,n,p);
r = r ./ repmat(sum(r,2),1,p);
function func = three_var(v)
func  =  (gamma(sum(v))/(gamma(v(1))*gamma(v(2))*
gamma(v(3))))*(factorial(n)/(factorial(nv(1))*factorial(nv(
2))*factorial(nv(3))))*((gamma(v(1)+nv(1))*gamma(v(2)
+nv(2))*gamma(v(3)+nv(3)))/gamma(v(1)+nv(1)+v(2)+n
v(2)+v(3)+nv(3)));
func = -func;
end
```

Table 1 Comparison of the error variance under the Frequentist and Bayesian techniques for different sample sizes

| Sample size	Input values			Frequentist weights			Bayesian weights			Error variance frequentist			Error variance Bayesian			Difference in error variance			Gain in efficiency
Total	n_1	n_2	n_3	w_1	w_2	w_3	w_1'	w_2'	w_3'	evd_1	evd_2	evd_3	evs_1	evs_2	evs_3	d_1	d_2	d_3	
10	2	3	5	0.2000	0.3000	0.5000	0.2012	0.3003	0.4986	0.0160	0.0210	0.0250	0.0000	0.0000	0.0000	0.0159	0.0209	0.0249	0.9965
19	4	6	9	0.2105	0.3158	0.4737	0.2116	0.3159	0.4725	0.0087	0.0114	0.0131	0.0001	0.0001	0.0001	0.0085	0.0112	0.0129	0.9876
21	7	7	7	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.0106	0.0106	0.0106	0.0001	0.0001	0.0001	0.0104	0.0104	0.0104	0.9850
75	17	25	33	0.2267	0.3333	0.4400	0.2276	0.3333	0.4391	0.0023	0.0030	0.0033	0.0004	0.0005	0.0006	0.0018	0.0024	0.0026	0.8053
80	20	40	20	0.2500	0.5000	0.2500	0.2507	0.4986	0.2507	0.0023	0.0031	0.0023	0.0005	0.0006	0.0005	0.0017	0.0024	0.0017	0.7781
100	20	30	50	0.2000	0.3000	0.5000	0.2038	0.2965	0.4997	0.0016	0.0021	0.0025	0.0005	0.0007	0.0008	0.0010	0.0013	0.0016	0.6556
100	23	48	29	0.2300	0.4800	0.2900	0.2309	0.4787	0.2904	0.0018	0.0025	0.0021	0.0006	0.0008	0.0006	0.0011	0.0016	0.0014	0.6650
143	39	45	59	0.2727	0.3147	0.4126	0.2731	0.3151	0.4118	0.0014	0.0015	0.0017	0.0010	0.0010	0.0012	0.0004	0.0005	0.0005	0.2857
167	20	90	57	0.1198	0.5389	0.3413	0.1213	0.5388	0.3399	0.0006	0.0015	0.0013	0.0006	0.0014	0.0013	0.0000	0.0001	0.0000	0.0000
168	55	56	57	0.3274	0.3333	0.3393	0.3274	0.3333	0.3393	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0000	0.0000	0.0000	0.0000
187	47	78	62	0.2513	0.4171	0.3316	0.2526	0.4158	0.3316	0.0010	0.0013	0.0012	0.0010	0.0013	0.0011	0.0000	0.0000	0.0001	0.0000
229	84	90	55	0.3668	0.3930	0.2402	0.3664	0.3922	0.2414	0.0010	0.0010	0.0008	0.0010	0.0010	0.0008	0.0000	0.0000	0.0000	0.0000
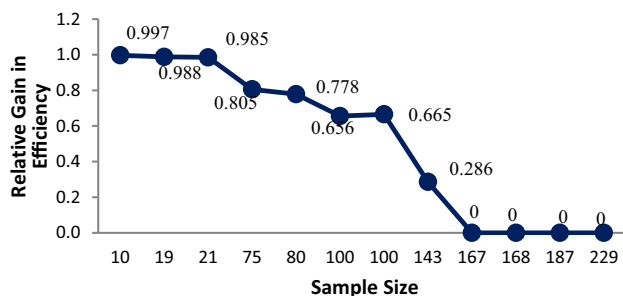
Fig. 1 Relative Gain in Efficiency vs Sample Size

4 Results and discussion

Comparison of estimate of weights has been performed for the Frequentist and Bayesian models. Results obtained from the pilot survey are as follows:

$n_1 = 24$, $n_2 = 11$, $n_3 = 12$, $n = 47$.

Under Frequentist setup, the estimated weights are

$$\hat{W}_i = n_i / n \qquad (14)$$

As $n_i \sim$ Binomial (n, p_i) for $i = 1, 2, 3$, the error variance is given by:

$$V(\hat{W}_i) = Var(n_i / n) = \hat{W}_i(1 - \hat{W}_i)/n \qquad (15)$$

Under Bayesian setup, estimates of weights are given in Equation (11). Expression for variance with respect to posterior Dirichlet $(\alpha_1 + n_1, \alpha_2 + n_2, \ldots, \alpha_l + n_l)$ distribution:

$$V(\hat{W}i*) = Var\left(\frac{\hat{\alpha}_i + n_i}{\sum_{i=1}^{3}(\hat{\alpha}_i + n_i)}\right) = \frac{n\hat{Wi*}(1 - \hat{Wi*})}{\left\{\sum_{i=1}^{3}(\hat{\alpha}_i + n)\right\}^2} \qquad (16)$$

In order to compare the performance of proposed Bayesian model with the existing frequentist method, one needs to consider samples with varying sizes. The results are shown in Table 1.

Although the weights seem to be close, it is clear from the results that the new model outperforms the frequentist one with respect to stability under small sample sizes.

Efficiency of estimator T_2 with respect to T_1 is given by

$$E = \frac{V(T_1)}{V(T_2)} \qquad (17)$$

Figure 1 depicts the performance of the two estimators with respect to the relative gain in efficiency for varying sample sizes. Suppose there are two estimators T_1 and T_2, relative gain in efficiency of T_2 with respect to T_1 is given by,

$$G = \frac{(V(T_1) - V(T_2))}{V(T_1)} \qquad (18)$$

Note that $G \approx 0$, indicates that the two estimators are equally efficient. An estimator T_2 is more efficient than T_1 if $V(T_2) \leq V(T_1)$, $G > 0$. Calculating the relative gain in efficiency in Table 1, it is observed that the gain in efficiency due to the proposed method over the existing one is quite high for small sample sizes.

It can be observed that the proposed estimator outperforms the existing one with respect to gain in efficiency for small sample sizes. With the increase in sample sizes there is a steady decrease in the gain in efficiency, indicating that with respect to the given data, the two estimators become equally competent for large sample sizes. But in reality, it may be difficult to have a large sample data, thus the effectiveness of the proposed method gains prominence. Bayesian determination of weights is highly recommended in such cases where conducting a large-scale survey is time consuming, difficult to implement as well as expensive.

5 Application in parking route problem

The proposed methodology has been applied in the field of Intelligent Transport System (ITS). Smart transportation is the need of the hour for sustainable development in a growing economy. Smart transportation supported by a strong communication network and based on sound statistical techniques is a key for tomorrow's smart cities. This route optimization tool promotes environment conservation and sustainable development by providing the most optimal route to a parking lot thereby saving time, energy, and fuel. Discovering the most optimal parking lot is a serious problem in the cities and it tends to aggravate during peak hours of the day and at congested places. Selecting the route depends on multiple conflicting objectives, namely minimizing the distance to the parking lot, maximizing the speed of the car, and lastly maximizing the parking availability at the lot. The detailed problem definition, formulation, design methodology and implementation are available at [26]. A pilot survey was conducted among 50 drivers. The Bayesian and frequentist weights

Table 2 Weights-Frequentist and Bayesian

Description	Count	Frequentist weight	Bayesian weight
Highest priority to distance to parking lot:	16	0.32	0.29
Highest priority to travel speed on the route:	14	0.28	0.30
Highest priority to parking availability at parking lot:	20	0.40	0.41

Table 3 Fitness values across Generations for Frequentist and Bayesian Weights for two time slots 12AM–4AM and 12Noon–4PM

Gen	Freq 12–4 am	Bayes 12–4 am	Freq 12–4 pm	Bayes 12–4 pm
1	0.742	0.729	0.808	0.789
2	0.73	0.719	0.808	0.789
3	0.711	0.696	0.808	0.75
4	0.688	0.696	0.761	0.75
5	0.688	0.675	0.747	0.75
6	0.671	0.672	0.718	0.732
7	0.671	0.672	0.718	0.726
8	0.671	0.63	0.718	0.709
9	0.654	0.63	0.702	0.709
10	0.654	0.61	0.702	0.697
11	0.628	0.599	0.702	0.691
12	0.628	0.599	0.695	0.678
13	0.604	0.597	0.695	0.678
14	0.604	0.597	0.676	0.678
15	0.588	0.596	0.676	0.678
16	0.588	0.577	0.676	0.678
17	0.588	0.566	0.676	0.654
18	0.572	0.561	0.663	0.654
19	0.572	0.554	0.663	0.643
20	0.572	0.55	0.663	0.632
21	0.558	0.548	0.663	0.632
22	0.558	0.533	0.649	0.632
23	0.558	0.533	0.649	0.615
24	0.536	0.518	0.637	0.608
25	0.536	0.498	0.623	0.608
26	0.529	0.498	0.623	0.596
27	0.529	0.485	0.623	0.596
28	0.503	0.485	0.623	0.596
29	0.503	0.485	0.623	0.596
30	0.503	0.485	0.623	0.596

Fig. 2 Plot of Fitness vs. Generations for two time slots 12AM–4AM and 12 Noon–4PM showing how Fitness values change across 30 generations using frequentist (green) and Bayesian (red) weights (Color figure online)

were calculated respectively using Equations (11) and (14) and summarized in the Table 2.

Genetic Algorithm has been used to solve the Multi-Objective optimization problem. The algorithm has been designed to run for 30 generations as the fitness values have stabilized by then in most cases. The fitness values obtained across generations have been plotted and graphically represented in Table 3 and Figure 2.

It is observed that as the generations increases, the value of the fitness function tend to decrease till it stabilizes at an optimal value. In both cases it was seen that the Bayesian weights produced lower fitness values consistently. The process was then repeated for thirty different executions. The fitness values obtained for both the Frequentist and Bayesian weights were noted in Table 4 and plotted in Figure 3.

Table 3, combined with Fig. 2, shows that the fitness value exhibits a consistently decreasing trend as the number of generations increase across all time zones. Focusing on the fitness values corresponding to a typical time zone, here 12 am to 4 am, it can further be noticed from Table 5 that both the average fitness value and the best fitness value were lower for the Bayesian weights than Frequentist weights in 30 executions. Secondly the routes as well as the parking lot vary depending on the time zone. This simulates a real-life scenario where parking lots and routes are bound to change as the values for the different factors changes. Although distance remains constant but the average speed and availability of parking lots changes with time which gets finally reflected in the fitness function.

6 Conclusion

The Bayesian Hierarchical model provides a posterior distribution on weights and is suitable for generating weights to check the nature of the solution set. Moreover, generation of weights using the proposed Bayesian methodology can be used to develop a bona-fide Bayesian posterior distribution for the optima, thus properly and coherently quantifying the uncertainty about the optima. It has been shown that the proposed estimator outperforms the existing ones with respect to efficiency for small sample sizes. In practice, as it is difficult to have a large sample data, the effectiveness of the proposed method gains prominence. Bayesian determination of weights finds high applicability in cases where conducting a large-scale survey is time consuming, difficult to implement as well as expensive. This proposed model is designed to adequately derive information from the collected data, rendering highly efficient estimators for small data sizes. This technique has been analyzed for error variances thereby quantifying the reliability of the estimates.
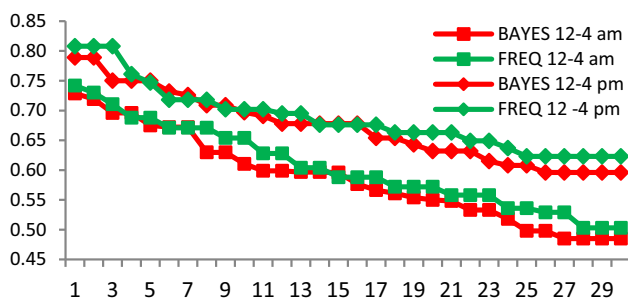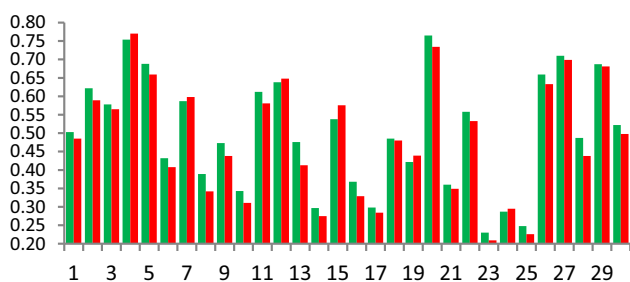
1340

Int. j. inf. tecnol. (August 2021) 13(4):1331–1341

Table 4 Fitness values across 30 executions for Frequentist and Bayesian Weights for time slot 12AM–4AM

Run	Frequentist fitness	Bayesian fitness	Run	Frequentist fitness	Bayesian fitness
1	0.503	0.485	16	0.368	0.329
2	0.622	0.589	17	0.298	0.284
3	0.578	0.565	18	0.485	0.48
4	0.754	0.77	19	0.422	0.439
5	0.688	0.659	20	0.765	0.734
6	0.432	0.408	21	0.36	0.349
7	0.587	0.598	22	0.558	0.533
8	0.389	0.342	23	0.23	0.209
9	0.473	0.438	24	0.287	0.295
10	0.343	0.311	25	0.248	0.226
11	0.612	0.581	26	0.659	0.633
12	0.638	0.648	27	0.71	0.699
13	0.476	0.413	28	0.487	0.438
14	0.297	0.275	29	0.687	0.681
15	0.538	0.576	30	0.522	0.498

Fig. 3 Graph showing how fitness varies for 30 different runs using frequentist (green) weights and Bayesian weights (red) for time slot 12AM–4AM (Color figure online)

Table 5 Descriptive Statistics for comparing the Fitness values across 30 executions for Frequentist and Bayesian Weights for time slot 12AM–4AM

Methodology	Mean	Worst	Best
Frequentist Approach	0.5005	0.7650	0.2300
Bayesian Approach	0.4828	0.7700	0.2090

When applied in the domain of route optimization in discovering the most suitable parking lot, the proposed methodology has produced results which display close resemblance to the phenomenon observed in real life situations. This work relied on sound statistical techniques to improve the weights representing the relative importance of the possibly conflicting objective functions of the route optimization process rather than improving the solution set directly. It has also been observed that on an average the fitness values obtained under weights generated by the proposed methodology outperforms that obtained by

frequentist approach. If implemented in reality, this would certainly ensure saving of time, energy and fuel, thus a greener world.

References

1. Arulmozhiyal R, Jubril AM. (2012) A nonlinear weights selection in weighted sum for convex multi-objective optimization. FACTA UNIVERSITATIS (NI_S) Ser. Math. Inform.; 27(3): 357–372
2. Zadeh LA (1963) Optimality & non-scalar-valued performance criteria. IEEE Trans. Automat. Control, AC-8:59–60
3. Haimes Y, Lasdon L, Wismer D (1973) On a bi-criterion formulation of the problems of integrated system identification and system optimization. IEEE Trans Syst Man Cybern 1:296–297
4. Lin JG (1991) On multiple-objective design optimization by goal methods. American Control Conference. MA, USA, Boston, pp 372–373
5. Waltz FM (1967) An engineering approach: Hierarchical optimization criteria. IEEE Trans. Auto. Control, AC-12: 179–180
6. Guan J, Lin G (2005) On min-norm and min-max methods of multi-objective optimization. Math. Program, Ser. A(103): 1–33
7. Steuer RE (1989) Multiple criteria optimization: theory, computation & application. Robert E. Krieger Publishing, Malabar
8. Das I, Dennis JE (1997) A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multi-criteria optimization problems. Struct Optim 14:63–69
9. Yoon KP (1995) Hwang CL. Multiple attribute decision making, an introduction. London: Sage Publications.
10. Saaty TL (1977) A scaling method for priorities in hierarchies, multiple objectives and fuzzy sets. J Math Psych 15:234–281
11. Wierzbicki AP (1986) A methodological approach to comparing parametric characterizations of efficient solutions. Large-Scale

Modeling and Interactive Decision Analysis, Lecture Notes in Economics and Mathematical Systems, Springer 273:27–45

12. Rao JR, Roy N (1989) Fuzzy set theoretic approach of assigning weights to objectives in multi criteria decision making. Int J Syst Sci 20:1381–1386

13. Messac A (1996) Physical programming: effective optimization for computational Design. AIAA J 34:149–158

14. Athan TW, Papalambros PY (1996) A quasi-montecarlo method for multi-criteria optimization". Eng Opt 27:177–198

15. Gennert MA, Yuille AL (1998) "Determining the optimal weights in multiple objective function optimization". In Second International conference on computer vision, IEEE, Los Alamos, CA, September 1998, 87–89

16. Das I, Dennis JE (1998) Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM J Optim 8(3):631–657

17. Lin JG (1976) Multiple-objective problems: Pareto-optimal solutions by method of proper equality constraints. IEEE Transactions on Automatic Control, AC-21(5): 641–650

18. Marler R, Arora J (2010) The weighted sum method for multi-objective optimization: new insights". Struct Multidiscip Optim 41:853–862

19. Wang R, Zhou Z, Ishibuchi H, Liao T, Zhang T (2018) Localized weighted sum method for many-objective optimization. IEEE Trans Evol Comput 22(1):3–18

20. Zhang H (2013) A trial of the dynamic weighted sum method for multi-objective optimization. IAENG Trans Electrical Eng 1:201–214

21. Jaini NI, Utyuzhnikov SV (2018) A fuzzy trade-off ranking method for multi-criteria decision-making. Axioms 7(1):1–21

22. Zitzler E (1999) Evolutionary algorithms for multi-objective optimization: methods and applications. Doctoral Thesis of Technical Sciences Diss. ETH No. 13398, Swiss Federal Institute of Technology Zurich

23. Athan TW, Papalambros PY (1996) A note on weighted criteria methods for compromise solutions in multi-objective optimization. Eng Opt 27:155–176

24. Ryu JH, Kim S, Wan H (2009) Pareto front approximation with adaptive weighted sum method in multi-objective simulation optimization. In: Proceedings of the 2009 Winter Simulation Conference, 623–633

25. Schmaranzer D, Braune R, Doerner KF (2019) Multi-objective simulation optimization for complex urban mass rapid transit systems. Annals of Operations Research, 1–38

26. Beed RS, Sarkar S, Roy A (2020) Bayesian hierarchical multi-objective optimization for vehicle parking route discovery. Innovations Syst Softw Eng. https://doi.org/10.1007/s11334-020-00373-4