



Quantifying and reducing the complexity of multi-line charts as a visual aid in multi-criteria decision-making

He Huang^{1,3} · Sajid Siraj²

Received: 4 November 2023 / Accepted: 31 May 2024
© The Author(s) 2024

Abstract

Multi-line charts are commonly used in multi-criteria decision-making (MCDM) to represent multiple data series on the same graph. However, the presence of conflicting criteria or divergent viewpoints introduces the challenge of accurately interpreting these charts, necessitating thoughtful design to improve their comprehensibility. In this paper, we model these multi-line charts as connected perfect matching bipartite graphs. We propose a metric called the Coefficient of Complexity (CoC) to quantify the complexity of these multi-line charts. In order to reduce the visual complexity of these charts, we propose to minimize the CoC by modeling it as an integer linear optimization problem (reminiscent of the traveling salesman problem). We demonstrate our techniques through multiple real-life case studies, wherein multi-line charts serve as data visualization across various MCDM software tools. Additionally, multi-line charts with specific requirements have been optimized using our approach, showcasing the adaptability and efficacy of our technique. We also formulate the radar chart as a specialized form of the multi-line chart, and adapt our technique to improve its comprehensibility. The proposed CoC and its optimization are important contributions to the field of analytics, as a number of methods use multi-line charts for visual aid. Consequently, enhancing their comprehensibility can facilitate the decision-making process and help decision-makers gain insights.

Keywords Multiple criteria analysis · MCDM · MCDA · Optimization · Data visualization

1 Introduction

Data visualization leverages graphical or visual representations to depict quantitative or qualitative data, offering a more intuitive comprehension of the information (Knaffic, 2015). In

✉ He Huang
he.huang@psi.ch
Sajid Siraj
s.siraj@leeds.ac.uk

¹ Laboratory for Energy Systems Analysis, Paul Scherrer Institut, Forschungsstrasse 111, 5232 Villigen, Aargau, Switzerland

² Centre for Decision Research, University of Leeds, Woodhouse, Leeds, West Yorkshire LS2 9JT, UK

³ Mobilise research group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Ixelles, Brussels, Belgium

the realm of multi-criteria decision-making (MCDM), various data visualization methods like tables, graphs, and charts are useful means to assist decision-makers and stakeholders (Miettinen, 2014). A judicious selection of charts can enhance the reader's understanding significantly (Arabnia, 1999). By effectively communicating the strengths and weaknesses of alternatives, these visualization techniques facilitate a more informed decision-making process.

Nonetheless, achieving effective data visualization in MCDM is not an easy task. This difficulty is registered in the context of the description problematic δ of Bernard Roy's four-stage methodology (Roy & Vincke, 1981), aiming to represent alternatives more accurately by fostering a cognitive process through data visualization (Roy, 1996). Its complexity arises from the need to incorporate multiple dimensions of discrete elements-including criteria, alternatives, and multiple decision-makers (DM) in group decision-making (GDM)-which together constitute a multi-dimensional data series. To facilitate comprehension, MCDM methods and their associated software tools have incorporated a variety of chart types, each offering different levels of detail. Multi-line charts are frequently employed in MCDM, yet their comprehensibility remains under-explored, particularly given their propensity for visual complexity and entanglement when data volume increases. In this study, our objective is to propose and apply a novel technique to reduce the visual complexity of these charts, thus providing a better representation of the underlying data for DMs. Our goal is to enhance the readability of multi-line charts that feature a relatively modest number (e.g. 10) of data series, i.e., alternatives, without compromising any of the information they convey. Undertaking this task is far from trivial. Insights from the psychology of comparison literature indicate that the complexity of making comparisons increases substantially as the number of stimuli increases (Thurstone, 1954). This suggests that adding even a single stimulus to an analysis can dramatically alter the comprehensibility of visual displays. Hence, when delving into the intricacies of multi-line charts, the difficulty in interpreting them stems not merely from the quantity of alternatives shown but is also compounded by the diversity of dimensions represented on the x-axis. Consequently, a relatively modest number of data series (i.e., alternatives) across a limited number of dimensions (i.e., criteria, factors, scorers) can still quickly escalate the complexity of the chart due to overlapping alternatives. This complexity increase poses a significant challenge to decision makers (DMs), adding additional visual complexity that can hinder their ability to efficiently interpret and act on the data displayed. Therefore, developing an appropriate technique to overcome this challenge is crucial.

The study is structured as follows: We begin with a literature review on MCDM, group decision-making related data visualization, and the general usage and development of multi-line charts in MCDM. Through this review, we identify the research gap that our study aims to address. We then introduce our approach for enhancing multi-line charts and present our optimization model, which serves as the core technique. The effectiveness of this technique is then demonstrated through both a hypothetical complex numerical example and multiple real-world case studies. Finally, we discuss potential directions for future work and conclude the study.

2 Literature review

MCDM problems inherently involve comparing and selecting from a variety of alternatives based on conflicting criteria (Marttunen et al., 2017). The challenge lies not only in the

multidimensionality of the decision vectors but also in effectively communicating this complex information to the DMs (Korhonen et al., 1992). Data visualization serves as a crucial support mechanism in this regard. They help elucidate similarities and differences among alternatives and enhancing overall comprehension and efficiency in the decision-making process (Miettinen, 2014).

2.1 Are charts properly utilized in MCDM software?

To assist practitioners in addressing MCDM problems, diverse MCDM software tools have been developed (Weistroffer & Li, 2016). Miettinen (2014) extensively surveyed various types of charts, ranging from generic to more specialized visualizations, utilized in MCDM software. MCDM software can be categorised into two main types based on Malczewski's taxonomy on MCDM (Malczewski, 1999): Multi-Criteria Decision Analysis (MCDA)/Multiple Attribute Decision Analysis (MADA) software, and Multiple Objective Programming (MOP)/Multiple Objective Optimization (MOO) software. Due to the distinct characteristics of MCDA and MOO (Vergara-Solana et al., 2019), these software tools offer varied visualization options to meet specific requirements.

Common types of charts, such as bar charts, line charts, and radar charts, are widely employed in MCDA software, including ValueDecision (Haag et al., 2022a) and MAMCA software (Huang et al., 2020), among others. Furthermore, several sensitivity analysis charts proposed by Belton and Vickers (1990) are frequently utilized in software tools like Expert Choice (Forman et al., 1983), etc.

In addition to these general charts, specialized charts for specific methodologies have also been developed. For instance, visualization techniques specific to the Analytic Hierarchy Process (AHP) are presented in Siraj et al. (2015), to PROMETHEE in Visual PROMETHEE (Mareschal & De Smet, 2009), and to FITradeoff in the FITradeoff software (de Almeida et al., 2016), among others. These specialized charts further extend the scope of visual tools available for decision-making analysis.

On the other hand, the parallel coordinate plot is a popular chart extensively used in MOP software tools. In parallel coordinate plots, objective functions are typically represented by vertical axes and solution candidates are represented by polylines. We will discuss it in detail in Sect. 2.3.

Though different charts are utilized, the task of visualizing such complex data is non-trivial. There is a delicate balance between ensuring the simplicity of the visualized data and maintaining its comprehensive nature (Cleveland, 1985). In other words, the challenge is to create visual representations that are easy to interpret without losing critical information or unintentionally introducing misleading information. It is essentially a trade-off among complexity, precision, and comprehensibility. In this study, we delve deeper into the multi-line chart, examining its usage in MCDM and exploring potential avenues for improvements.

2.2 Are multi-line charts preferred within the MCDM context?

Line charts are widely used for data visualization (Spear, 1952). This type of chart presents information through a series of data points, i.e., markers, which are linked by straight line segments (Newman, 1954). Building on the versatility of line charts, multi-line charts, also known as series line charts, extend the functionality by allowing the representation of more than one data series on the same graph. These charts become especially useful when the data sets share a common variable, such as time or an ordinal scale, making the comparison across

Table 1 Multi-line chart usage in MCDM software

Category	Software name	Used*	Source
MCDA software	1000Minds	Yes	Hansen and Ombler (2008)
	Criterion decision plus	No	Murphy (2014)
	DecideIT	No	Preference (2024a)
	Decisionarium	Yes	Hämäläinen (2003)
	DEXI	No	Craheix et al. (2015)
	Diviz	No	Meyer and Bigaret (2012)
	D-Sight	Yes	Hayez et al. (2012)
	Entscheidungsnavi	Yes	Siebert et al. (2021)
	Expert choice	Yes	Forman et al. (1983)
	FITradeoff	Yes	Pessoa et al. (2022)
	Helison	No	Preference (2024b)
	IRIS	No	Dias et al. (2002)
	Logical decisions	No	Chelst and Canbolat (2011)
	MAMCA	Yes	Huang et al. (2020)
	M-MACBETH	Yes	Bana e Costa and Vansnick (1999)
	PriEsT	No	Siraj et al. (2015)
	SOCRATES	No	Munda et al. (2022)
	Transparent choice	Yes	TransparentChoice (2024)
	Value decision	Yes	Haag et al. (2022a)
	Visual PROMETHEE	Yes	Mareschal and De Smet (2009)
MOO software	VIP analysis	No	Dias and Clímaco (2000)
	IND-NIMBUS	Yes	Miettinen (2014)
	Interalg	No	Kroshko (2007)
	ModeFRONTIER	Yes	Parashar et al. (2010)
	RGDB	No	Lotov et al. (2004)
	ParallAX	Yes	Avidan and Avidan (1999)

* If multi-line chart is used as data visualization in the software

different categories or groups straightforward. It effectively displays multi-dimensional data in a sequential format. Thus, the multi-line charts are aligned with the data type and objective of MCDM, which are able to illustrate the performance of different alternatives on different criteria. The process becomes more complicated when decisions are made by groups of DMs, each with varying preferences and judgements, which has been widely researched as multi-criteria group decision-making (MCGDM) (Hwang & Lin, 2012). Therefore, in this context, the multi-line chart is a suitable chart type that can simultaneously illustrate the performances/preferences of alternatives on different criteria and possibly different groups.

To systematically assess the integration of multi-line chart visualizations within MCDM software, we conducted an exhaustive review of both web-based and desktop (Windows-based) applications as catalogued in the studies by Miettinen (2014); Weistroffer and Li (2016); MCDM-Society (2024).¹ The surveyed software is listed in Table 1. It's important to mention that our analysis excluded complementary tools or solvers developed on Excel,

¹ Test Environment Specifications: The evaluations were performed on a system running Windows 10, powered by an Intel® Core™ i7-12800H processor.

Matlab, Python, or similar platforms. However, it's worth acknowledging the existence of notable software such as ASMO (Eichfelder, 2009) and DESDEO (Misitano et al., 2021). These tools were not included in our survey of multi-line chart usage due to their higher flexibility for data visualization on their platforms.

We surveyed 26 software tools and discovered that 14 of them employ multi-line(-like) charts for data visualization purposes.² These charts are utilized in various ways across the surveyed tools. Specifically, some tools use multi-line charts to display the outcomes of data analyses, effectively representing multi-dimensional data sets. Others incorporate these charts as interactive elements, broadening their applicability to activities such as sensitivity analysis. Below, we delve into some of these applications in greater detail, with screenshots in Fig. 1.

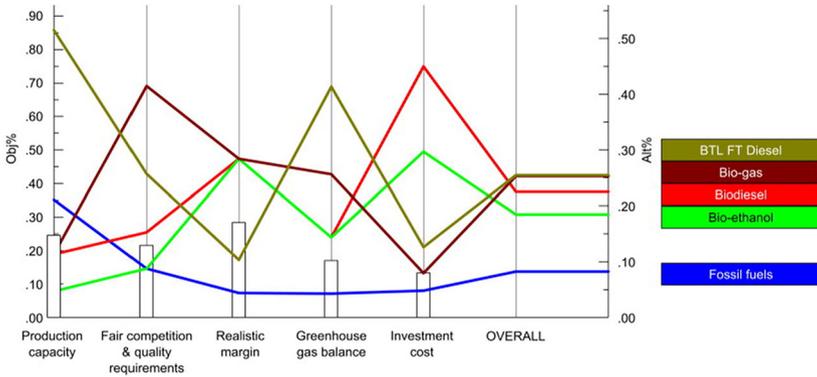
1. In Expert Choice, the multi-line chart is utilized as a form of sensitivity analysis visualization. It is called “all in one chart” (see Fig. 1a). Paired with a bar chart that depicts the weights of various criteria, users can gain a comprehensive understanding of the performance of different alternatives (Forman et al., 1983; Ishizaka & Labib, 2009).
2. In software such as VISA and IND-NIMBUS, multi-line charts are deployed to illustrate “value paths” (Belton & Vickers, 1990; Miettinen & Mäkelä, 2000). As illustrated in Fig. 1b, each line in the chart represents an alternative, reflecting its performance across various criteria, which are depicted as bars. These bars illustrate the range of each criterion within the Pareto optimal set (Miettinen, 1999).
3. In the MAMCA software, multi-line charts are employed to showcase the preferences of various stakeholder groups within a group decision-making framework, called “multi-actor view” (Huang et al., 2020). In Fig. 1c, the x-axis represents different stakeholders, using a nominal scale, while the y-axis portrays the preferences of the alternatives among these various stakeholder groups.
4. In MOP software, parallel coordinates increasingly stand out among the most efficient approaches (Wegman, 1990). The “value paths” can be seen as parallel coordinates with fewer solutions. In particular, parallel coordinates offer space efficiency per criterion, allowing scalable application across many criteria (Fleming et al., 2005). Such diagrams facilitate the comparison of multiple alternatives over a set of numerical variables. An example of this application can be seen in the modeFRONTIER software, as shown in Fig. 1d.

2.3 The enhancement of multi-line chart

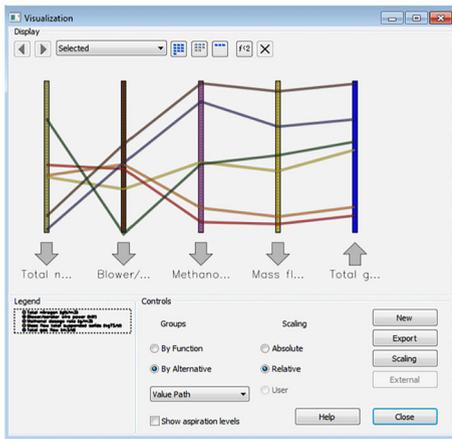
We can see the multi-line charts are commonly used as a method for visual representation in various contexts (Miettinen, 2014). However, multi-line charts, despite their ability to provide rich insights, can also become overly complex, leading to confusion and reducing their interpretability. Each additional line adds a layer of complexity to the chart, which can lead to clutter and confusion, making it harder to distinguish individual lines and derive accurate insights.

This overly complex scenario often happens in the MOP studies. Thus, researchers continue to devise new approaches to facilitate the readability of multi-charts, i.e., parallel coordinate plot in MOP. For example, some suggest that by carefully choosing colors, line styles, line thickness, and markers may help distinguish different lines (representing different

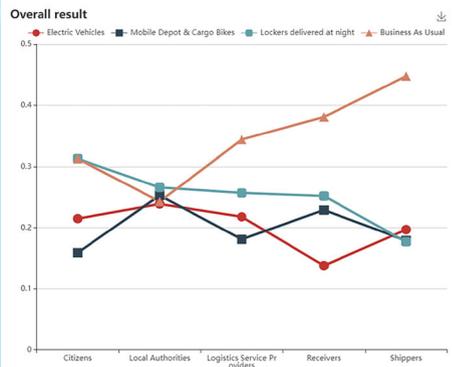
² Despite the different terminologies used in different software to refer to multi-line(-like) charts, they are unified by their underlying purpose of data visualization. For consistency and clarity, we will refer to all such variations as “multi-line charts” in the remainder of this paper.



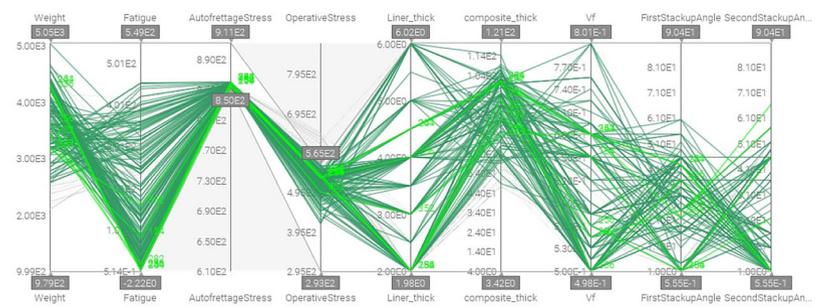
(a) ExpertChoice (Forman et al., 1983)



(b) IND-NIMBUS (Miettinen & Mäkelä, 2000)



(c) MAMCA software (Huang et al., 2020)



(d) modeFRONTIER (Parashar et al., 2010)

Fig. 1 Screenshots of multi-line charts in MCDM software

series of data) (Healey, 1996; Card, 1999). Inselberg (1997) also proposed a systematic way to process parallel coordinate plots. The integration of more interactive techniques has led to the adoption of a technique known as “brushing,” which allows users to manipulate plots actively to produce more insightful data visualizations (Becker & Cleveland, 1987).

A notable advancement in enhancing parallel coordinate plots involves rearranging the dimensions-specifically, the labels on the X-axis-to uncover specific data properties. Ankerst et al. (1998) introduced an algorithm that models the task of rearranging axes as a solution to the Traveling Salesman Problem (TSP) (Lawler, 1985), aiming to position dimensions with similar behaviors adjacent to each other. Zhen et al. (2017) expanded on this concept, developing an algorithm that effectively illustrates the dominance relationship between solutions, thereby shedding light on the evolutionary process. However, they didn’t investigate how to extract more useful information from the solution sets and then present this information in the parallel coordinates. To address this, Saini (2022) proposed an innovative visualization technique known as SCORE bands. This method combines several strategies, including solution clustering, axis ordering, and axis placement, to aid DMs in identifying patterns among solutions and understanding correlations between objectives. In the end, the lines in parallel coordinates plots are clustered and visualized as bands in the charts and axes are reordered to better help DMs understand objective correlations. Nonetheless, such methods typically employed in MOP are generally better suited for analyzing larger datasets, for example, those containing hundreds of lines. The application of these methods might lead to a loss of original information in the visualization, such as when transforming clusters of lines into bands. In the context of MCDA/MCGDM, the scenario is quite different. Here, a handful number of alternatives have already been identified, and the evaluation is conducted with respect to specific criteria. The focus remains on the readability and accuracy of information, that is, reducing the complexity of multi-line charts whilst keeping the information intact.

While the aforementioned approaches may not be directly applicable to this goal, they provide valuable insights that inform our proposed technique. Research conducted by Ghoniem et al. (2005) demonstrates a notable decline in graph readability as the quantity of intersections escalates. For this reason, careful attention must be given to the design and layout of multi-line charts.

In the context of reducing readability complexity, force-directed layout algorithms, which aim to position nodes to minimize edge crossings, represent an intriguing approach for enhancing graph readability (Tamassia, 2013). Despite their effectiveness in network graphs, their direct application to multi-line charts proves challenging due to differing structural complexities. While there is a certain degree of similarity between the two chart types, there currently isn’t a suitable algorithm for reducing visual complexity in multi-line charts using a force-directed layout concept. Consequently, the goal of this study is to develop a technique that adapts the principles of this algorithm to optimize multi-line charts. This objective stands to benefit from the insights gained through extensive research on edge (or arc) crossings in graph theory (Purchase et al., 1996).

2.4 Edge crossings problems in bipartite graphs

Bipartite graphs, which consist of two distinct sets of vertices with edges only connecting vertices from different sets, play a fundamental role in graph theory and have broad applications in fields such as networking, data visualization, and matching algorithms. A critical aspect of research on bipartite graphs focuses on minimizing edge crossings when visualizing

these structures because such crossings significantly impact the readability and interpretability of the graphs (Purchase et al., 1996). An exemplary problem in this domain is Turán's brick factory problem, which queries the minimum number of crossings in a drawing of a complete bipartite graph (Turán, 1977). Following this, numerous studies have introduced various algorithms aimed at reducing edge crossings. Tamassia (2013) provides a comprehensive overview of these techniques, including the layering approach. In this approach, vertices are organized into horizontal layers, and edges are depicted as straight lines between these layers. Although effective, this method often requires complex algorithms to optimize the ordering of vertices within the layers to minimize crossings. The edge crossing problem has been proven to be NP-hard (Garey & Johnson, 1983), highlighting the difficulty of finding an optimal solution. Consequently, researchers have explored heuristic methods. He et al. (2007) discuss the use of heuristic algorithms that, although not ensuring an optimal solution, can significantly reduce the number of crossings in large bipartite graphs.

While our study shares the general objective of minimizing visual complexity, our approach to simplifying multi-line graphs by drawing inspiration from edge crossing problems in bipartite graphs differs from traditional edge crossing problems. The classical edge crossing problem is defined in the context of complete bipartite graphs, where every vertex from one set is connected to every vertex in the other set (Gould, 2012). In contrast, our investigation of multi-line graphs concerns a set of bipartite graphs where each graph is perfectly matched, indicating a one-to-one correspondence between vertices in the two sets (Tanimoto et al., 1978). Moreover, in our context, the connections within the bipartite structures are predetermined, with each line representing a set of data, such as the performance metrics of alternatives across different criteria. This introduces unique properties and challenges specific to our study, which are discussed in the following section. Consequently, there is a need to develop a novel method specifically tailored to reducing the visual complexity inherent in multi-line charts.

3 Visual complexity in the multi-line chart

As aforementioned, our objective is to enhance graph readability by minimizing edge crossings (i.e. intersections). Our goal is to optimize a specific chart type, namely a multi-line chart, where one axis represents a nominal scale and the other axis corresponds to a numerical scale. The chart depicts scores associated with a set of labels, denoted as $N = \{n_1, n_2, \dots, n_J\}$, which are displayed on the nominal scale.

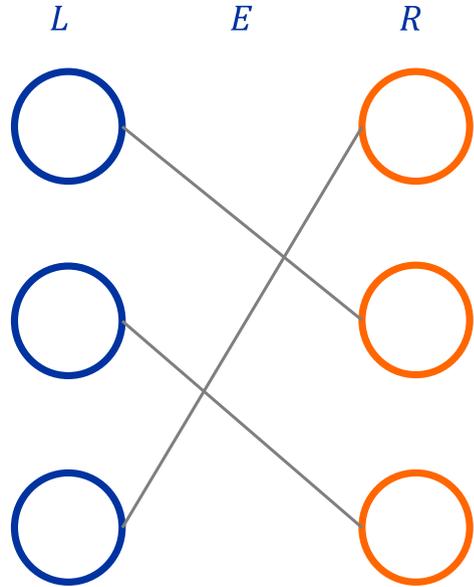
We use a sequence of connected perfect matching bipartite graphs to illustrate this type of multi-line chart.

Definition 1 (Bipartite Graph) A bipartite graph is a graph $G = (V, E)$ consisting of a set of vertices V , and a set of edges E . V can be divided into two disjoint sets L and R , where L represents the vertices on the left side of the edges, and R denotes the vertices on the right side of the edges. Every edge $e \in E$ connects a vertex in L to a vertex in R .

Formally, this means that there are no edges connecting vertices within the same set. The bipartite graph can also be denoted as:

$$G = (L \cup R, E) \text{ where } L \cap R = \emptyset, \text{ and } \forall e \in E, e = \{l, r\} \text{ with } l \in L, \text{ and } r \in R. \quad (1)$$

Fig. 2 A perfect matching bipartite graph



Definition 2 (Perfect Matching Bipartite Graph) If a bipartite graph L and R have equal cardinality, meaning they have the same number of vertices, the graph is called a balanced bipartite graph. A perfect matching bipartite graph is a balanced bipartite where every vertex in V is adjacent to exactly one edge in E .

This implies that a perfect matching pairs each vertex in L with exactly one vertex in R and vice versa, leaving no vertex unmatched. An example of a perfect matching bipartite graph is illustrated in Fig. 2, which can be mathematically formulated as:

$$B = (L \cup R, E) \quad \text{where} \quad |E| = |L| = |R|, \quad \text{and} \quad \forall v \in L \cup R, \exists ! e \in E \text{ such that } v \in e. \tag{2}$$

Now, we define the items to be displayed on the multi-line chart. Consider a set of items $I = \{i_1, i_2, \dots, i_M\}$. Each item is allocated a score based on the labels; consequently, for each $i_m \in I$, an associated score $S_m = \{s_{m,n_1}, s_{m,n_2}, \dots, s_{m,n_j}\}$ is assigned. The score matrix can be illustrated as:

$$\begin{matrix}
 & n_1 & \dots & n_j \\
 i_1 & \left[\begin{matrix} s_{1,n_1} & \dots & s_{1,n_j} \end{matrix} \right] \\
 \vdots & \left[\begin{matrix} \vdots & \ddots & \vdots \end{matrix} \right] \\
 i_M & \left[\begin{matrix} s_{M,n_1} & \dots & s_{M,n_j} \end{matrix} \right]
 \end{matrix}, \tag{3}$$

where the line segments for i_m are subsequently generated by connecting the scores for each label $\{s_{m,n_1}, s_{m,n_2}, \dots, s_{m,n_j}\} (\forall m \in M)$.

Definition 3 The multi-line chart can be represented by a sequence of $J - 1$ connected perfect matching bipartite graphs $B = \{B_1, B_2, \dots, B_{J-1}\}$.

For each graph $B_j = (L_j \cup R_j, E_j)$, line segments are drawn for items on the labels n_j and n_{j+1} . The set L_j contains the scores on label n_j , i.e., $\{s_{1,n_j}, s_{2,n_j}, \dots, s_{M,n_j}\}$, while $R_j = \{s_{1,n_{j+1}}, s_{2,n_{j+1}}, \dots, s_{M,n_{j+1}}\}$. The scores $\{s_{m,n_j}, s_{m,n_{j+1}}\}$ are connected by edge $e_{m,n_j} \in E_j$. These bipartite graphs are connected such that the right-side vertices of one bipartite graph

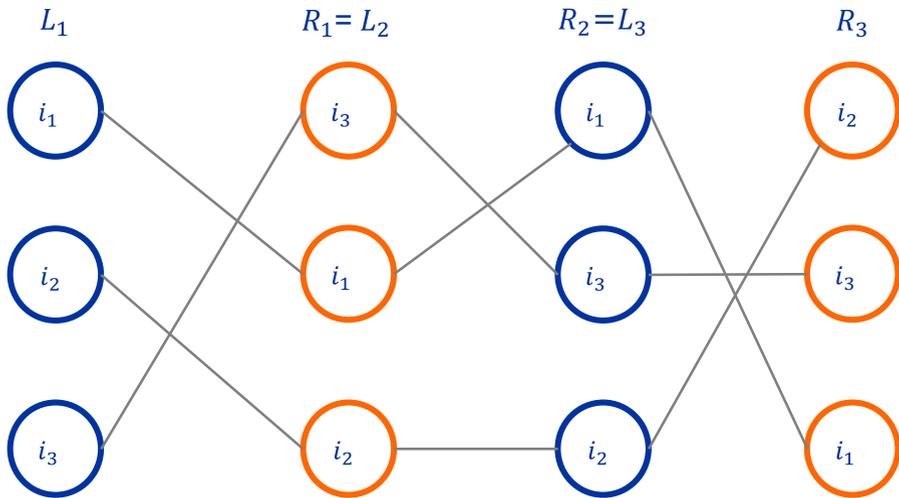


Fig. 3 Connected perfect matching bipartite graphs $\{B_1, B_2, B_3\}$

is the left-side vertices of the next bipartite graph, ensuring that for all B_j, B_{j+1} , we have $R_j = L_{j+1}$. Thus we define $\mathcal{V}'_j = R_j = L_{j+1}$. As illustrated in Fig. 3, a series of connected perfect matching bipartite graphs $\{B_1, B_2, B_3\}$ demonstrates this relationship. In total, $M \times N$ line segments are drawn to represent the relationships among the items.

3.1 Definition of the visual complexity of bipartite

Now, we introduce the concept of visual complexity in the context of the bipartite graph. As the number of intersections increases, it becomes more challenging to comprehend the connections between vertices. A study by Ghoniem et al. (2005) found that the readability of a graph significantly decreased when the number of intersections increased. Thus, minimizing intersections in a bipartite graph can improve the readability and overall effectiveness of a graph.

As a result, it is essential to determine the number of intersections within a given bipartite graph. To accomplish this, we first need to establish the conditions under which two edges in the bipartite graph intersect. To illustrate this, let us consider two edges connecting two random items in an adjacent bipartite in a multiple bipartite graph, i_m and $i_{m'}$, with their respective labels n_j and n_{j+1} , where the line segments/edges e_{m,n_j}, e_{m',n_j} connect items. It is important to note that n_j and n_{j+1} represent nominal labels, and as such, it is not permissible to compute their difference directly. Nevertheless, for the purpose of this analysis, we assume an undefined distance $\mathcal{D} > 0$ exists between the two sides of the bipartite graph, such that $|n_j - n_{j+1}| = \mathcal{D}$ (see Fig. 4). The subsequent theorem demonstrates that the value of \mathcal{D} has no impact on the results of our analysis.

Theorem 1 For a given adjacent bipartite in connected bipartite graphs, two non-overlapped line segments are intersecting as long as it satisfies $(s_{m,n_j} - s_{m',n_j}) \cdot (s_{m,n_{j+1}} - s_{m',n_{j+1}}) \leq 0$.

Proof Antonio (1992) developed an algorithm that determine two given line segments in 2-D space whether they intersect or not. The line segments can be defined in terms of first degree

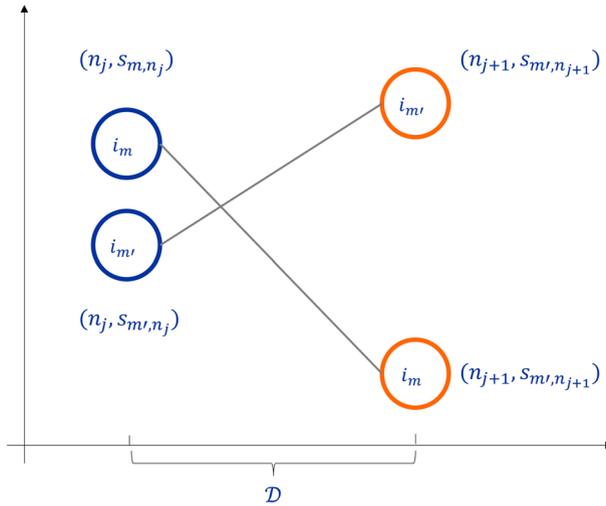


Fig. 4 Two intersecting line segments

Bézier parameters (Mortenson, 1999):

$$e_{m,n_j} = \begin{bmatrix} n_j \\ s_{m,n_j} \end{bmatrix} + t \begin{bmatrix} n_j - n_{j+1} \\ s_{m,n_j} - s_{m,n_{j+1}} \end{bmatrix}, e_{m',n_j} = \begin{bmatrix} n_j \\ s_{m',n_j} \end{bmatrix} + u \begin{bmatrix} n_j - n_{j+1} \\ s_{m',n_j} - s_{m',n_{j+1}} \end{bmatrix}, \quad (4)$$

where t and u are real numbers. The intersection point of the line segments is found with one of the following values of t or u range $[0, 1]$ (Antonio, 1992). In our case:

$$t = \frac{\begin{vmatrix} n_j - n_j & n_j - n_{j+1} \\ s_{m,n_j} - s_{m,n_{j+1}} & s_{m,n_{j+1}} - s_{m',n_{j+1}} \end{vmatrix}}{\begin{vmatrix} n_j - n_{j+1} & n_j - n_{j+1} \\ s_{m,n_j} - s_{m',n_j} & s_{m,n_{j+1}} - s_{m',n_{j+1}} \end{vmatrix}} = \frac{(n_j - n_j)(s_{m,n_{j+1}} - s_{m',n_{j+1}}) - (s_{m,n_j} - s_{m,n_{j+1}})(n_j - n_{j+1})}{(n_j - n_{j+1})(s_{m,n_{j+1}} - s_{m',n_{j+1}}) - (s_{m,n_j} - s_{m',n_j})(n_j - n_{j+1})}, \quad (5)$$

$$u = \frac{\begin{vmatrix} n_j - n_j & n_j - n_{j+1} \\ s_{m,n_j} - s_{m,n_{j+1}} & s_{m,n_j} - s_{m',n_j} \end{vmatrix}}{\begin{vmatrix} n_j - n_{j+1} & n_j - n_{j+1} \\ s_{m,n_j} - s_{m',n_j} & s_{m,n_{j+1}} - s_{m',n_{j+1}} \end{vmatrix}} = \frac{(n_j - n_j)(s_{m,n_j} - s_{m',n_j}) - (s_{m,n_j} - s_{m,n_{j+1}})(n_j - n_{j+1})}{(n_j - n_{j+1})(s_{m,n_{j+1}} - s_{m',n_{j+1}}) - (s_{m,n_j} - s_{m',n_j})(n_j - n_{j+1})}. \quad (6)$$

As there is an undefined distance \mathcal{D} exists between the two sides of the bipartite graph, we have:

$$t = \frac{0 - \mathcal{D}(s_{m,n_j} - s_{m,n_{j+1}})}{\mathcal{D}(s_{m,n_{j+1}} - s_{m',n_{j+1}}) - \mathcal{D}(s_{m,n_j} - s_{m',n_j})} = \frac{-s_{m,n_j} + s_{m,n_{j+1}}}{-s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j}} = u. \quad (7)$$

Thus, the intersection can be found if:

$$0 \leq \frac{-s_{m,n_j} + s_{m,n_{j+1}}}{-s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j}} \leq 1. \quad (8)$$

When $-s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j} \neq 0$, we have:

$$\begin{aligned} -s_{m,n_j} + s_{m,n_{j+1}} &\leq -s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j}, \text{ if } -s_{m,n_j} + s_{m,n_{j+1}} \geq 0, \\ -s_{m,n_j} + s_{m,n_{j+1}} &\geq -s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j}, \text{ if } -s_{m,n_j} + s_{m,n_{j+1}} \leq 0 \end{aligned} \quad (9)$$

i.e.

$$s_{m',n_{j+1}} \leq s_{m',n_j}, \text{ if } s_{m,n_{j+1}} \geq s_{m,n_j}; s_{m',n_{j+1}} \geq s_{m',n_j}, \text{ if } s_{m,n_{j+1}} \leq s_{m,n_j}.$$

Eventually,

$$(s_{m,n_j} - s_{m',n_j}) \cdot (s_{m,n_{j+1}} - s_{m',n_{j+1}}) \leq 0, \text{ if } -s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j} \neq 0.$$

□

Now, we define the visual complexity of the bipartite graph. Recall that intersections of line segments contribute to the difficulty of readability. Additionally, readability is also challenged when two line segments overlap. Overlapping occurs in the special case when $-s_{m,n_j} + s_{m,n_{j+1}} - s_{m',n_{j+1}} + s_{m',n_j} = 0$, i.e., $(s_{m,n_j} - s_{m',n_j}) = (s_{m,n_{j+1}} - s_{m',n_{j+1}})$. In this situation, the two line segments are either parallel or overlapping; trivially, $(s_{m,n_j} - s_{m',n_j}) \cdot (s_{m,n_{j+1}} - s_{m',n_{j+1}}) \geq 0$. Overlapping occurs specifically when $(s_{m,n_j} - s_{m',n_j}) = (s_{m,n_{j+1}} - s_{m',n_{j+1}}) = 0$. With this understanding, we can define under which circumstances the visual complexity increases in the bipartite graph. And then we can further identify the visual complexity of a given bipartite.

Definition 4 (Visual complexity). For two edges/line segments in a given bipartite, it can be seen as an increase of visual complexity when $(s_{m,n_j} - s_{m',n_j}) \cdot (s_{m,n_{j+1}} - s_{m',n_{j+1}}) \leq 0$.

Theorem 2 For a given perfect matching bipartite graph with M edges, the maximum number of intersections and overlaps is $\binom{M}{2}$.

Proof For every edge $e_j \in E_j$, it can intersect or overlap with all the other edges. Consequently, to determine the number of intersections between every two edges from a fixed set of M edges, we utilize the binomial coefficient $\binom{M}{2} = \frac{M!}{2!(M-2)!}$. □

To better quantify the visual complexity in the perfect matching bipartite, we define a coefficient called coefficient of complexity (CoC):

Definition 5 (Coefficient of complexity). For a perfect matching bipartite with M edges, the coefficient of complexity (CoC) is formulated as follows:

$$\begin{aligned} \text{CoC} &= \frac{\# \text{Intersections and overlaps}}{\text{Maximum possible}} \# \text{intersections and overlaps} \\ &= \frac{|(s_{m,n_j} - s_{m',n_j}) \cdot (s_{m,n_{j+1}} - s_{m',n_{j+1}}) \leq 0|}{\binom{M}{2}}, \forall m < m' \leq M, \end{aligned} \quad (10)$$

where CoC is a value in the range of $[0, 1]$. It compares the number of intersections and overlaps of edges in a given bipartite graph to the maximum possible number of intersections and overlaps for an M -edge perfect matching bipartite graph. When the CoC is 1, the bipartite graph exhibits maximal readability difficulty; conversely, when the CoC is 0, the bipartite graph is considered to have no readability difficulty.

4 Optimization model for decreasing the visual complexity of the multi-line chart

Expanding upon the proposed coefficient of complexity, we propose an optimization model with the objective of minimizing visual complexity. As previously noted, the type of multi-line charts targeted for optimization can be conceptualized as a series of interconnected perfect matching bipartite graphs. These charts exhibit a unique characteristic: they present the scores of various items across distinct labels, and rearranging the order of the labels has no impact on the chart’s meaning. Consequently, the strategy for reducing the visual complexity of these multi-line charts involves adjusting the order of the labels in order to minimize the number of intersections and overlaps. Building on Theorem 1, the intersections of line segments are unaffected by the nominal scale distances on the labels. Trivially, overlaps are also unrelated to these distances. Thus, the first step of the optimization is to compute the CoC for all possible single bipartite graphs corresponding to different pairs of labels. For a given dataset represented as matrix (3), the CoC of items can be pairwise compared across different labels and is depicted as follows:

$$\begin{bmatrix}
 \text{CoC}_{n_1, n_1} & \text{CoC}_{n_1, n_2} & \dots & \text{CoC}_{n_1, n_{J-1}} & \text{CoC}_{n_1, n_J} \\
 \text{CoC}_{n_2, n_1} & \text{CoC}_{n_2, n_2} & \dots & \text{CoC}_{n_2, n_{J-1}} & \text{CoC}_{n_2, n_J} \\
 \vdots & \vdots & \ddots & \vdots & \vdots \\
 \text{CoC}_{n_{J-1}, n_1} & \text{CoC}_{n_{J-1}, n_2} & \dots & \text{CoC}_{n_{J-1}, n_{J-1}} & \text{CoC}_{n_{J-1}, n_J} \\
 \text{CoC}_{n_J, n_1} & \text{CoC}_{n_J, n_2} & \dots & \text{CoC}_{n_J, n_{J-1}} & \text{CoC}_{n_J, n_J}
 \end{bmatrix}, \tag{11}$$

where $\text{CoC}_{n_j, n_{j'}}$ signifies the coefficient of complexity for a perfect matching bipartite graph based on the scores of $I = \{i_1, i_2, \dots, i_M\}$ on the labels $n_j, n_{j'}$. The coefficients along the diagonal, i.e., CoC_{n_j, n_j} , are not required to be calculated.

Upon computing coefficients for all pairwise comparisons, our next step is to identify an ordering to construct the new multi-line chart that minimizes the aggregated CoCs. Let’s define the reordered connected bipartite graphs as $B' = \{B'_1, B'_2, \dots, B'_{n_{j-1}}\}$. These graphs still adhere to the properties outlined in Definition 3. Thus, the reordered connected bipartite graphs is still a series of perfect matching bipartite graphs. For each vertex set $V' \in \{L'_j, R'_j = L'_{j+1}, R'_{j+1}\} (j = 1, 2, \dots, J - 1)$ within the total vertex set of the connected bipartite graphs V^C , the condition $\forall V' \in V^C, \exists ! V'$ holds. This rule ensures that a set of vertices within one label only appears once throughout the entire connected bipartite graphs. Consequently, by calculating each CoC of bipartite within the reordered connected bipartite graphs, we obtain a set of CoCs, $C' = \{CoC_{n'_1, n'_2}, CoC_{n'_2, n'_3}, \dots, CoC_{n'_{j-1}, n'_j}\}$. Then we aggregate all the CoCs of C' :

$$O = \sum_{j=1}^{J-1} \text{CoC}_{n'_j, n'_{j+1}}. \tag{12}$$

Given the unique existence property of B' , i.e., $\forall V' \in V^C, \exists ! V'$, we can actually determine O values for all possible orderings of the connected bipartite graphs, using matrix (11). This is enabled by the interconnection of all labels. With the exception of the initial and final labels, the item scores for all other labels must uniquely connect to two other distinct sets of item scores. The optimization process thus aims to identify the minimum O within the realm of possible orderings. Intriguingly, this optimization bears resemblance to the (symmetric) traveling salesman problem (TSP), as the CoCs can be considered analogous to distances between cities, with O representing the total tour length in TSP (Pop et al., 2023). Therefore,

we can construct an integer linear optimization model, drawing inspiration from the principles of the TSP (Papadimitriou & Steiglitz, 1998):

$$\min O = \min \sum_{j=1}^J \sum_{j \neq j', j'=1}^J \text{CoC}_{n_j, n_{j'}} \cdot x_{n_j, n_{j'}} \quad (13)$$

$$x_{n_j, n_{j'}} = \begin{cases} 1, & \text{if vertex sets on label } j \text{ and } j' \text{ are connected;} \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

where $x_{n_j, n_{j'}}$ are the binary variables that determine whether the vertex sets on different labels are connected. Thus, we have the constraints that except for the first and the last labels, the vertex sets on all the other labels are exactly connected to two other distinct sets. To generalize this as a conventional TSP problem, where the salesman is required to return to the original city (i.e., all labels must connect to two other labels), we introduce a dummy label. This dummy label has zero CoC to every other label, thereby allowing for the return to the origin, aligning with conventional TSP formulations (Lawler, 1985). Thus, the optimization problem (13) changes to:

$$\min O = \min \sum_{j=1}^{J+1} \sum_{j \neq j', j'=1}^{J+1} \text{CoC}_{n_j, n_{j'}} \cdot x_{n_j, n_{j'}}, \quad (15)$$

where

$$\begin{aligned} \text{CoC}_{n_j, n_{J+1}} &= 0, \forall n_j \in N; \\ \text{CoC}_{n_{J+1}, n_j} &= 0, \forall n_j \in N. \end{aligned} \quad (16)$$

The complete optimization model will be designed utilizing the Dantzig-Fulkerson-Johnson (DFJ) formulation (Dantzig et al., 1954). The selection of the DFJ formulation is attributed to its relative simplicity and ease of implementation for the TSP, particularly when dealing with a problem of our scale which does not present a large optimization challenge (Gutin & Punnen, 2006):

$$\underset{x, u}{\text{minimize}} O = \sum_{j=1}^{J+1} \sum_{j \neq j', j'=1}^{J+1} \text{CoC}_{n_j, n_{j'}} x_{n_j, n_{j'}}, \quad (\text{Objective}), \quad (17)$$

$$\sum_{j=1, j' \neq j}^{J+1} x_{n_j, n_{j'}} = 1, \forall j = 1, \dots, J+1, \quad (18)$$

$$\sum_{j=1, j \neq j'}^{J+1} x_{n_j, n_{j'}} = 1, \forall j' = 1, \dots, J+1, \quad (19)$$

$$\sum_{n_j \in Q} \sum_{n_{j'} \in Q, j \neq j'} x_{n_j, n_{j'}} \leq |Q| - 1, \forall Q \subsetneq \{n_1, \dots, n_{J+1}\}, |Q| \geq 2, \quad (\text{DFJ constraint}), \quad (20)$$

$$x_{n_j, n_{j'}} \in \{0, 1\}, \forall j, j' = 1, \dots, J+1, j \neq j', \quad (\text{Binary decision variables}). \quad (21)$$

Where constraint (20) is the subtour elimination constraint. Q is a subset of the labels, and this constraint enforces that, the resultant solution forms a singular, uninterrupted tour rather than a collective amalgamation of smaller, disparate tours (Grötschel & Padberg, 1975).

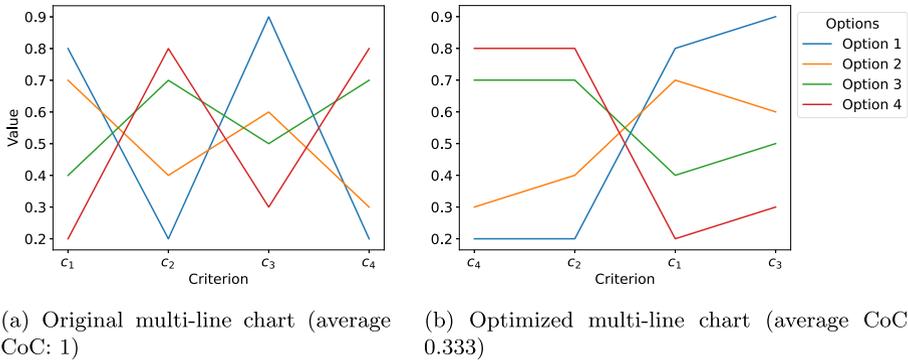


Fig. 5 Illustrating the performance of 4 alternatives on 4 criteria from a hypothetical numerical illustration

While the subtour elimination constraint can potentially give rise to an exponential quantity of constraints, our scenario is relatively straightforward, implying that computational complexity should not pose a significant issue. Upon completion of the optimization process, the dummy label is subsequently disregarded. Consequently, this process allows us to ascertain the path of the TSP without necessitating a return to the originating point, i.e., find out the initial label and final label in the line chart in our case. It should be noticed that, the objective function (17) is possible to be modified, as the intention of set the current objective function is to mimic the TSP problem. For example, it can be modified as:

$$\underset{x,u}{\text{minimize}} \frac{O}{J - 1} \tag{22}$$

where we can obtain a normalized value to evaluate the visual complexity in a range of [0, 1].

5 Algorithm demonstration

To showcase the efficacy of our optimization algorithm, we begin by applying it to a hypothetical MCDM problem featuring four alternatives and four criteria. Figure 5a presents a complex multi-line chart, characterized by numerous intersections among the lines, resulting in an average CoC value of 1 (i.e., the highest complexity, each bipartite has maximum possible intersections over edges). This intricate pattern suggests that the Decision Maker (DM) has varying preferences for different alternatives across the multiple criteria, making it challenging to understand their overall preference structure.

To reduce this complexity, we apply our optimization algorithm on this chart. The resultant chart, depicted in Fig. 5b reveals a significant improvement in readability. By optimizing the ordering of the criteria, the visual representation becomes more straightforward, allowing for a clearer understanding of the DM’s preference trends. The data clearly shows that Option 1 performs well on criteria c_1 and c_3 , while Option 2 performs better on c_2 and c_4 . Assuming equal weighting across criteria, the graph shows that Option 3 outperforms Option 2 overall. This is evidenced by the fact that Option 3’s superior criteria outperform those of Option 2, and its lesser performing criteria still outperform the corresponding lesser performing criteria of Option 2. An optimized graph enhances the clarity of these distinctions, making it easier to compare the performance of the options. Also, notably, the ‘sawtooth’ fluctuations

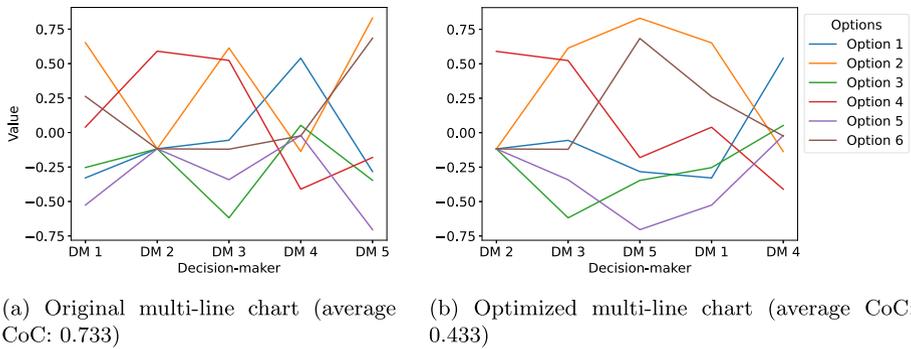


Fig. 6 Illustrating the performance of 6 alternatives appraised by 5 decision-makers from Huang et al. (2021)

in the lines are eliminated. Additionally, it allows for the discernment of trends in alternative preferences across various criteria. The application of the algorithm successfully reduces the average CoC to a more manageable value of 0.333, thereby facilitate the decision-making process.

5.1 Real-life case study: reaching consensus among DMs

Our real-life analysis begins with a case extracted from a consensus-reaching problem rooted in a transportation decision-making case (Huang et al., 2021). This particular case was selected due to its relative simplicity and clarity. In addition, as the main objective of this paper revolves around consensus-reaching issues, i.e., identifying the best performed options (items) based on different decision-makers (labels), we actually explore another mean of achieving “consensus” from a visual optimization perspective.

Figure 6a presents an initial multi-line chart generated from the Brussels case. The primary goal of this analysis is to determine the most sustainable alternatives according to the scores assessed by various DMs (Huang et al., 2021). The figure illustrates the performance of 6 alternatives appraised by 5 stakeholder groups based on Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) (Behzadian et al., 2010). Although the original case has its intrinsic value, in order to concentrate on our focus of visual improvement, we’ve chosen to represent the actual alternative options and DMs as numbers, rather than revealing their specific identities.

The calculated average CoC for the original chart is 0.733. The chart is notably cluttered with numerous intersecting lines, escalating its visual complexity. Consequently, capture the performance of individual options becomes a challenging task due to this intricate presentation of data.

We therefore implement the optimization model detailed in (17) with the goal of reducing visual complexity. The outcome of this process is demonstrated in Fig. 6b, which presents the optimized multi-line chart.

The improvements in visualization become discernible through the reduced intersection count. Noticeably, the average CoC has declined to 0.433 (A detail investigation can be found in Table 2). This improved view allows us to already visually identify the performance of individual options. Option 2 stands out with its relatively better performance: Option 2 is the favored choice among decision-makers DM1, DM3, and DM5. Furthermore, the analysis simplifies the task of pinpointing the least successful alternative. Options 3 and 5 demonstrate

Table 2 The CoC comparison before and after the optimization from Huang et al. (2021)

	Before	After
$\text{CoC}(B_1/B'_1)$	0.800	0.733
$\text{CoC}(B_2/B'_2)$	0.733	0.200
$\text{CoC}(B_3/B'_3)$	0.733	0.067
$\text{CoC}(B_4/B'_4)$	0.667	0.733
Average CoC	0.733	0.433

comparatively weaker preferences. The performance of “middle” alternatives also becomes more apparent, facilitating a straightforward assessment of their standings. Notably, The preferences for Option 6, which were difficult to identify in Fig. 6a, are clarified in Fig. 6b. Here it appears as the second choice for DM1 and DM5, and the fourth choice for DM3, providing a clearer perspective on its relative preference among decision makers.

5.1.1 Tailoring visualization to reflect specific needs

In practical applications of data visualization, the requirements often extend beyond simply generating a clear chart. For instance, it might be necessary to maintain a fixed order for certain dimensions of the data. This need could arise from the desire to compare two specific criteria more closely or to explore the preferences of two different DMs in greater detail. These specialized requirements can actually be incorporated as constraints in the proposed optimization algorithm to meet these practical needs effectively.

Consider the transportation decision-making case as detailed by Huang et al. (2021). Suppose DM1 and DM2 represent two distinct groups with potentially divergent interests yet collaborate closely. In such a scenario, we would like to keep these two DMs closely in the chart in order to compare the preferences of these two DMs. Therefore, the model (17) can be adapted by introducing an additional constraint:

$$x_{1,2} + x_{2,1} = 1, \quad (23)$$

This constraint ensures a direct connection between DM1 and DM2, either from DM1 to DM2 or vice versa, thereby guaranteeing their adjacency in the solution. Such a modification adeptly accommodates the potential requirements mentioned, effectively tailoring the optimization process to suit specific analytical goals. The multi-line chart optimized with the Eq. (23) is depicted in Fig. 7a, shown along with the unconstrained optimisation in Fig. 7b (for side-by-side comparison).

This visualization maintains a direct connection between DM1 and DM2. Remarkably, the average CoC declined, underscoring the feasibility of meeting specific real-life requirements while concurrently reducing complexity. Such findings highlight the adaptability and versatility of the proposed optimization model: it not only streamlines the intricacies of the visualization but also flexibly aligns with practitioners' specific needs. The subsequent sections will delve into more specialized scenarios, offering a deeper exploration of the model's applicability.

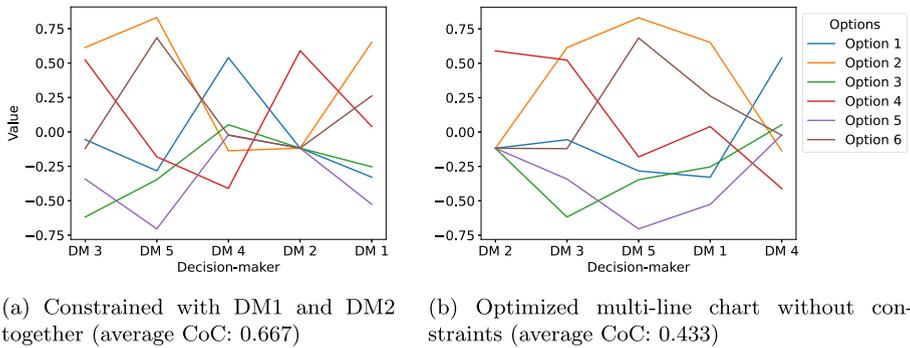


Fig. 7 Comparing the optimisation of CoC with and without constraint from Huang et al. (2021)

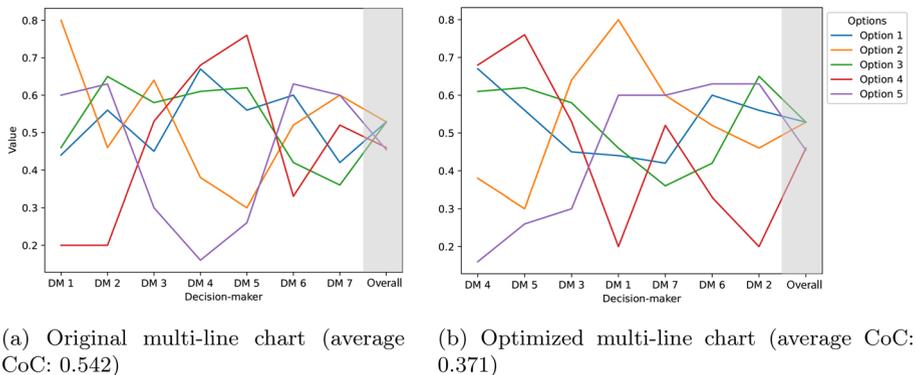


Fig. 8 Illustrating the performance of 5 alternatives appraised by 7 decision-makers from Macharis et al. (2012)

5.2 A special multi-line chart

We will now delve into a more specialized case, which is characterized by a higher level of complexity in the data structure.

The subsequent case study is a decision-making problem that assesses various biofuel options (Macharis et al., 2012), depicted in Fig. 8a. The chart displays the performance evaluation of five alternatives as assessed by seven stakeholder groups using the Analytic Hierarchy Process (AHP) method (Saaty, 1989). In addition, it includes an aggregated representation of each alternative's overall performance. We've chosen this example for several reasons. Firstly, its objective is similar to our previous case: to identify the best performing options, or in this instance, the optimal biofuel alternative. Secondly, the addition of DMs in the figure escalates the level of complexity. Furthermore, the original case study is presented through one of the famous decision-making software, Expert Choice (Ishizaka & Labib, 2009). This affords an opportunity to assess whether our optimization approach could enhance the visualization capabilities of such software (Ishizaka & Siraj, 2018). Lastly, this case study differs from the standard multi-line chart. Here, the overall scores, which is illustrated in gray area in Fig. 8a, reflects the cumulative scores calculated from the preceding DMs, representing the final outcome. Consequently, these overall scores will invariably be positioned last.

Table 3 The CoC comparison before and after the optimization from Macharis et al. (2012)

	Before	After
CoC(B_1/B'_1)	0.400	0.100
CoC(B_2/B'_2)	0.600	0.400
CoC(B_3/B'_3)	0.500	0.400
CoC(B_4/B'_4)	0.100	0.400
CoC(B_5/B'_5)	0.900	0.400
CoC(B_6/B'_6)	0.400	0.300
CoC(B_7/B'_7)	0.900	0.600
Average CoC	0.543	0.371

In order to optimize this variant of the multi-line chart, we need to reconfigure the pairwise CoC matrix (11). In a generic case, we introduce a dummy label. However, for this specific scenario, we'll adjust the CoC from this dummy label, transitioning from equation (16) to the following revised formula:

$$\text{CoC}_{n_j, n_{j+1}} = \varepsilon, j = 1, 2, \dots, J - 1; \tag{24}$$

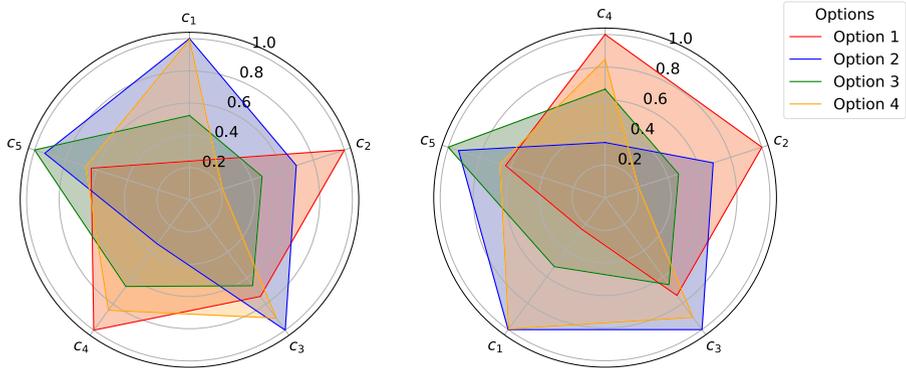
$$\text{CoC}_{n_{J+1}, n_j} = \varepsilon, j = 1, 2, \dots, J - 1. \tag{25}$$

where ε is an arbitrary large value, i.e., $\varepsilon \gg 1$. While $\text{CoC}_{n_j, n_{j+1}}$, $\text{CoC}_{n_{j+1}, n_j}$, $\text{CoC}_{n_{j+1}, n_{j+1}}$ remain 0. By doing so, the dummy data will always be connected to the last label. The optimized result is illustrated as Fig. 8b.

Despite the cumulative scores already being displayed at the end, the optimized visualization offers a more discernible perspective on the performance across DMs. Option 1, 2, and 3, with their higher cumulative scores, also exhibit stable lines in the optimized chart. On the contrary, Option 4 and 5 demonstrate markedly low scores, which contribute to a significantly lower cumulative result. The optimization process significantly enhances our ability to discern the preferences or “trends” among the decision-makers (DMs). Specifically, it clarifies that early in the optimized chart, DM4 and DM5 show a preference for Option 4, whereas towards the end, DM2 and DM6 exhibit a clear aversion to it. This optimization also reveals the rankings of options with greater precision. For instance, Option 1’s position in the preferences of the DMs becomes more clear after optimization. It was initially challenging to determine its rank across all DMs, as it was never positioned as either the top or bottom choice. Yet, through optimization, we can accurately determine Option 1’s standing among the DMs: it secures the second position for DM4 and DM6, the third position for DM2 and DM5, and the fourth position for DM1, DM3, and DM7. The enhanced visualization hence highlights these variations more distinctly. For the comparison of the CoCs before and after optimization, check Table 3.

5.3 The radar chart

The radar chart can be seen as a special form of the multi-line chart (that we aim to optimize), with the exception that every label is connected with two others. Some widely used variants of radar charts are spider charts and star plots (Tague, 2004). In fact, the optimization process for a radar chart can be viewed as a direct application of the original symmetric TSP, without the necessity of constructing a dummy label. By altering the objective from Eqs. (17) to (13),



(a) Original multi-line chart (average CoC: 0.600)

(b) Optimized multi-line chart (average CoC: 0.467)

Fig. 9 Illustrating the performance of 4 alternatives on 5 criteria from Roselli et al. (2019)

Table 4 The CoC comparison before and after the optimization from Roselli et al. (2019)

	Before	After
$\text{CoC}(B_1/B'_1)$	0.667	0.500
$\text{CoC}(B_2/B'_2)$	0.500	0.500
$\text{CoC}(B_3/B'_3)$	0.667	0.167
$\text{CoC}(B_4/B'_4)$	0.833	0.333
$\text{CoC}(B_5/B'_5)$	0.333	0.833
Average CoC	0.600	0.467

and disregarding items on $(J + 1)$ th label in the constraints, we are able to optimize the radar chart.

We will take a radar chart from another decision-making case study, which conducted an extensive neuroscience experiment to gauge how decision makers interpret the visualizations and select the most optimal alternatives from the FITradeoff method decision support system (DSS) (Roselli et al., 2019). Various types of graphs were included to evaluate multiple possibilities and ascertain biases when taken out of their specific contexts. We've chosen this particular case study because it aligns with our study's central aim: enhancing data visualization for decision-making. Despite having a similar objective, our primary focus is to explore methods for reducing visual complexity within a single chart. We selected one dataset from the study: 4A5C, which denotes four alternatives and five criteria. The radar chart for this dataset is illustrated in Fig. 9a. And the radar chart after the optimization is illustrated in Fig. 9b.

The radar chart displays the performance evaluation of four alternatives via the Flexible and interactive trade-off (FITradeoff) method across five distinct criteria: C_1, C_2, \dots, C_5 (de Almeida et al., 2016). Given the simplicity of the radar chart's dataset, it allows us to delve into a detailed investigation of the CoCs. As demonstrated in Table 4, we observe that the average CoC does not exhibit a significant improvement. This can be expected due to the stricter constraints inherent in radar chart optimization, which makes the decreasing of average CoC more challenging. In this case, the sole change following the optimization is the swapping of the positions of C_1 and C_4 . Although this shift doesn't result in a drastic improvement,

it enables an easier identification of the performances of alternatives on different criteria. Since the nature of the optimization model is to reduce the intersections, and in the spider chart all values of alternatives are connected on criteria, we can expect that the optimization will try to connect criteria where alternatives have similar performances together to reduce the intersections. Therefore, we can quickly identify, for example, that option 2 performs well on c_1 , c_3 , and c_5 , and relatively poorly on c_4 and c_2 . It is not obvious from the original graph in Fig. 9a. Furthermore, dominant alternatives—those that outperform all others can be effortlessly identified after the optimization. We can discern that Option 1 performs best in C_4 and C_2 , Option 2 in C_1 (achieving a tied top 1) and C_3 , and Option 3 in C_5 , respectively.

5.4 Discussion

Our proposed technique demonstrates several advancements over existing visualization strategies for reducing the complexity of multi-line charts used in MCDM. By introducing the CoC as a metric to quantify visual complexity and applying an optimization algorithm to minimize the CoC, our method improves the readability and interpretability of multi-line charts without compromising data integrity.

Our method offers distinct advantages over existing visualization techniques. Existing techniques, such as those that focus on optimizing parallel coordinate plots by rearranging dimensions (Ankerst et al., 1998; Zhen et al., 2017), or more advanced optimization techniques such as SCORE bands proposed by Saini (2022), perform better with larger datasets by highlighting specific data points. However, our approach is more effective in situations with fewer alternatives, ensuring that visual complexity remains manageable and the decision-making process is facilitated without loss of information. This balance between readability and information integrity makes our technique a valuable addition to the variety of visualization tools available in the MCDM domain. Moreover, our technique is highly flexible and can be applied to various types of multi-line charts, including radar charts. By adjusting different constraints, the optimization can be tailored to different situations. This tailored visualization optimization increases the applicability and effectiveness of our technique in different scenarios.

Nevertheless, when optimizing a spider chart, it is critical to validate whether the performance of alternatives may be chart dependent. For example, in certain methodologies, such as Multi-Criteria Sustainability Assessment (MCSA), scores are derived by calculating the area of the spider chart generated by the performance of alternatives across criteria (Nzila et al., 2012). In this context, the score is influenced by the order in which the criteria are arranged, which means that changing the order of the criteria can potentially lead to rank reversals (Dias & Domingues, 2014). The potential bias introduced by reordering needs to be considered during the optimization process. This awareness is essential to ensure that interpretations of the data remain accurate and unbiased, recognizing that changes in the visual representation can affect the perceived performance of alternatives.

In conclusion, our method enhances data interpretation and supports decision-makers in making more informed decisions, thereby leading to better decision-making outcomes. Additionally, these improvements have potential applications beyond the field of MCDM, extending their utility to various other domains. When data needs to be illustrated to identify trends and correlations across different metrics and categories, multi-line charts can be effectively employed (Peebles & Ali, 2015; Ratwani & Gregory Trafton, 2008). Our approach can be then optimizing these charts in diverse areas such as healthcare (Kudyba, 2010), engineering (Kamba et al., 1996), manufacturing (Oakland & Oakland, 2007), and market

research (Hutchinson et al., 2010). As long as the X-axis does not require a specific order, such as chronological sequencing, our technique can be applied to improve the readability and interpretability of these data visualizations.

6 Conclusion and future work

In this study, we define the visual complexity of multi-line charts in the spirit of that of connected perfect matching bipartite graphs. We then introduce the Coefficient of Complexity (CoC), a novel metric designed to quantify the complexity of multi-line charts. We further propose an algorithm aimed at minimizing the CoC, modeling it as an integer linear optimization problem analogous to the traveling salesman problem. Through an illustrative example and real-world case studies, we demonstrate how our technique effectively decreases visual complexity in these charts. Moreover, we expand the technique's utility by applying it to other specialized forms of multi-line charts, underscoring its value in bolstering the clarity and readability of these visualizations. For instance, we investigate specific dimensions of interest by strategically positioning them next to each other in the chart, all while aiming for an overall reduction in complexity.

In summary, we regard the optimization of the proposed CoC as significant contributions to the field of MCDM. This significance is underlined by the extensive reliance on multi-line charts within various MCDM methodologies as a means of visual aid. Our enhancements to the comprehensibility of these charts serve a crucial role, not merely as academic exercises, but as practical tools that can facilitate the decision-making process. By improving the comprehensibility of multi-line charts, we empower decision-makers to extract more meaningful insights from the data, thus facilitating more informed and effective choices. Furthermore, these improvements have potential applications beyond the field of MCDM, extending their utility to various other fields.

While the proposed optimization technique effectively reduces visual complexity, there are several directions to explore for future research. We can enhance the validation of our technique by conducting an experimental analysis. This could involve an A/B test where a group of people are tasked with identifying the performance of different data series on multi-line charts both before and after optimization. Such a study could provide additional empirical evidence for the efficacy of our technique. Meanwhile, the optimization model itself could be improved by considering additional chart attributes. For instance, we could examine not only the intersections of adjacent labels, but also the trends in overall performance across different data series. This would help avoid the creation of 'saw-tooth' lines in the chart, which could mislead decision-makers in their assessment of overall performance.

Funding Open Access funding provided by Lib4RI - Library for the Research Institutes within the ETH Domain: Eawag, Empa, PSI & WSL. No funding was received for conducting this study.

Availability of data and materials The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ankerst, M., Berchtold, S., & Keim, D. A. (1998). Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Proceedings IEEE symposium on information visualization (cat. no. 98tb100258)* (pp. 52–60).
- Antonio, F. (1992). Faster line segment intersection. *Graphics gems iii (ibm version)* (pp. 199–202). Elsevier.
- Arabnia, H. (1999). Reading in information visualization: Using vision to think [media review]. *IEEE MultiMedia*, 6(4), 93.
- Avidan, T., & Avidan, S. (1999). Parallax-a data mining tool based on parallel coordinates. *Computational Statistics*, 14(1), 79–89.
- Bana e Costa, C. A., & Vansnick, J. C. (1999). The macbeth approach: Basic ideas, software, and an application. *Advances in decision analysis* (pp. 131–157). Springer.
- Becker, R. A., & Cleveland, W. S. (1987). Brushing scatterplots. *Technometrics*, 29(2), 127–142.
- Behzadian, M., Kazemzadeh, R. B., Albadvi, A., & Aghdasi, M. (2010). Promethee: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, 200(1), 198–215.
- Belton, V., & Vickers, S. (1990). Use of a simple multi-attribute value function incorporating visual interactive sensitivity analysis for multiple criteria decision making. In *Readings in multiple criteria decision aid Readings in multiple criteria decision aid* (pp. 319–334). Springer.
- Card, M. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- Chelst, K., & Canbolat, Y. B. (2011). *Value-added decision making for managers*. CRC Press.
- Cleveland, W. S. (1985). *The elements of graphing data*. Wadsworth Publ. Co.
- Craheix, D., Bergez, J. E., Angevin, F., Bockstaller, C., Bohanec, M., Colomb, B., Dore, T., Fortino, G., Guichard, L., Pelzer, E., & Messean, A. (2015). Guidelines to design models assessing agricultural sustainability, based upon feedbacks from the dexi decision support system. *Agronomy for Sustainable Development*, 35, 1431–1447.
- Dantzig, G., Fulkerson, R., & Johnson, S. (1954). Solution of a large-scale traveling-salesman problem. *Journal of the Operations Research Society of America*, 2(4), 393–410.
- de Almeida, A. T., de Almeida, J. A., Costa, A. P. C. S., & de Almeida-Filho, A. T. (2016). A new method for elicitation of criteria weights in additive models: Flexible and interactive tradeoff. *European Journal of Operational Research*, 250(1), 179–191.
- Dias, L. C., & Clímaco, J. N. (2000). Additive aggregation with variable interdependent parameters: The vip analysis software. *Journal of the Operational Research Society*, 51(9), 1070–1082.
- Dias, L. C., & Domingues, A. R. (2014). On multi-criteria sustainability assessment: Spider-gram surface and dependence biases. *Applied Energy*, 113, 159–163.
- Dias, L. C., Mousseau, V., Figueira, J., Clímaco, J., & Silva, C. G. (2002). Iris 1.0 software. Newsletter of the European Working Group. *Multicriteria Aid for Decisions*, 3(5), 4–6.
- Eichfelder, G. (2009). An adaptive scalarization method in multiobjective optimization. *SIAM Journal on Optimization*, 19(4), 1694–1718. <https://doi.org/10.1137/060672029>
- Fleming, P. J., Purshouse, R. C., & Lygoe, R. J. (2005). Many-objective optimization: An engineering design perspective. In *International conference on evolutionary multi-criterion optimization* (pp. 14–32).
- Forman, E., Saaty, T., Selly, M., & Waldron, R. (1983). Expert choice, decision support software. McLean, VA. URL <http://expertchoice.com/products-services/expert-choice-desktop/>.
- Garey, M. R., & Johnson, D. S. (1983). Crossing number is np-complete. *SIAM Journal on Algebraic Discrete Methods*, 4(3), 312–316.
- Ghoniem, M., Fekete, J. D., & Castagliola, P. (2005). On the readability of graphs using node-link and matrix-based representations: A controlled experiment and statistical analysis. *Information Visualization*, 4(2), 114–135.
- Gould, R. (2012). *Graph theory*. Courier Corporation.
- Grötschel, M., & Padberg, M. W. (1975). Partial linear characterizations of the asymmetric travelling salesman polytope. *Mathematical Programming*, 8, 378–381.

- Gutin, G., & Punnen, A. P. (2006). *The traveling salesman problem and its variations*. Springer Science & Business Media.
- Haag, F., Aubert, A. H., & Lienert, J. (2022). Valuedecisions, a web app to support decisions with conflicting objectives, multiple stakeholders, and uncertainty. *Environmental Modelling & Software*, *150*, 105361. <https://doi.org/10.1016/j.envsoft.2022.105361>
- Hämäläinen, R. P. (2003). Decisionarium-aiding decisions, negotiating and collecting opinions on the web. *Journal of Multi-Criteria Decision Analysis*, *12*(2–3), 101–110.
- Hansen, P., & Ombler, F. (2008). A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-Criteria Decision Analysis*, *15*(3–4), 87–107.
- Hayez, Q., De Smet, Y., & Bonney, J. (2012). D-sight: A new decision making software to address multi-criteria problems. *International Journal of Decision Support System Technology (IJDSST)*, *4*(4), 1–23.
- He, H., Sýkora, O., & Mäkinen, E. (2007). Genetic algorithms for the 2-page book drawing problem of graphs. *Journal of Heuristics*, *13*, 77–93.
- Healey, C. G. (1996). Choosing effective colours for data visualization. In *Proceedings of seventh annual IEEE visualization '96* (pp. 263–270).
- Huang, H., De Smet, Y., Macharis, C., & Doan, N. A. V. (2021). Collaborative decision-making in sustainable mobility: identifying possible consensuses in the multi-actor multi-criteria analysis based on inverse mixed-integer linear optimization. *International Journal of Sustainable Development & World Ecology*, *28*(1), 64–74.
- Huang, H., Lebeau, P., & Macharis, C. (2020). The multi-actor multi-criteria analysis (mamca): New software and new visualizations. In *International conference on decision support system technology*(pp. 43–56).
- Hutchinson, J. W., Alba, J. W., & Eisenstein, E. M. (2010). Heuristics and biases in data-based decision making: Effects of experience, training, and graphical data displays. *Journal of Marketing Research*, *47*(4), 627–642.
- Hwang, C. L., & Lin, M. J. (2012). *Group decision making under multiple criteria: Methods and applications (VOL 281)*. Springer Science & Business Media.
- Inselberg, A. (1997). Multidimensional detective. In *Proceedings of viz'97: Visualization conference, information visualization symposium and parallel rendering symposium* (pp. 100–107).
- Ishizaka, A., & Labib, A. (2009). Analytic hierarchy process and expert choice: Benefits and limitations. *OR Insight*, *22*(4), 201–220.
- Ishizaka, A., & Siraj, S. (2018). Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. *European Journal of Operational Research*, *264*(2), 462–471.
- Kamba, T., Elson, S. A., Harpold, T., Stamper, T., & Sukaviriya, P. (1996). Using small screen space more efficiently. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 383–390).
- Knafllic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. Wiley.
- Korhonen, P., Moskowitz, H., & Wallenius, J. (1992). Multiple criteria decision support-a review. *European Journal of Operational Research*, *63*(3), 361–375.
- Kroshko, D. (2007). Openopt: Free scientific-engineering software for mathematical modeling and optimization. URL <http://www.openopt.org>.
- Kudyba, S. P. (2010). *Healthcare informatics: improving efficiency and productivity*. CRC Press.
- Lawler, E. L. (1985). *The traveling salesman problem: A guided tour of combinatorial optimization*. Wiley-Interscience Series in Discrete Mathematics.
- Lotov, A. V., Kistanov, A. A., & Zaitsev, A. D. (2004). Visualization-based data mining tool and its web application. In *Chinese academy of sciences symposium on data mining and knowledge management* (pp. 1–10).
- Macharis, C., Turcksin, L., & Lebeau, K. (2012). Multi actor multi criteria analysis (MAMCA) as a tool to support sustainable decisions: State of use. *Decision Support Systems*, *54*(1), 610–620.
- Malczewski, J. (1999). *Gis and multicriteria decision analysis*. Wiley.
- Mareschal, B., & De Smet, Y. (2009). Visual promethee: Developments of the promethee & gaia multicriteria decision aid methods. In *2009 IEEE international conference on industrial engineering and engineering management* (pp. 1646–1649).
- Marttunen, M., Lienert, J., & Belton, V. (2017). Structuring problems for multi-criteria decision analysis in practice: A literature review of method combinations. *European Journal of Operational Research*, *263*(1), 1–17.
- MCDM-Society (2024). Software related to MCDM | Multiple Criteria Decision Making—mcdmsociety.org. <https://www.mcdmsociety.org/content/software-related-mcdm-0>. Accessed 07 Mar 2024.
- Meyer, P., & Bigaret, S. (2012). Diviz: A software for modeling, processing and sharing algorithmic workflows in MCDA. *Intelligent Decision Technologies*, *6*(4), 283–296.
- Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Springer Science & Business Media.

- Miettinen, K. (2014). Survey of methods to visualize alternatives in multiple criteria decision making problems. *OR Spectrum*, 36(1), 3–37.
- Miettinen, K., & Mäkelä, M. M. (2000). Interactive multiobjective optimization system www-nimbus on the internet. *Computers & Operations Research*, 27(7–8), 709–723.
- Misitano, G., Saini, B. S., Afsar, B., Shavazipour, B., & Miettinen, K. (2021). Desdeo: The modular and open source framework for interactive multiobjective optimization. *IEEE Access*, 9, 148277–148295.
- Mortenson, M. E. (1999). *Mathematics for computer graphics applications*. Industrial Press Inc.
- Munda, G., Azzini, I., Cerreta, M., & Ostlaender, N. (2022). Socrates manual. KJ-NA-31-327-EN-N (online), [https://doi.org/10.2760/015604\(online\)](https://doi.org/10.2760/015604(online))
- Murphy, P. J. (2014). Criterium decisionplus. *Making Transparent Environmental Management Decisions: Applications of the Ecosystem Management Decision Support System*, 35–60.
- Newman, E. B. (1954). *Experimental psychology*. American Psychological Association.
- Nzila, C., Dewulf, J., Spanjers, H., Tuigong, D., Kiriamiti, H., & Van Langenhove, H. (2012). Multi criteria sustainability assessment of biogas production in Kenya. *Applied Energy*, 93, 496–506.
- Oakland, J., & Oakland, J. S. (2007). *Statistical process control*. Routledge.
- Papadimitriou, C. H., & Steiglitz, K. (1998). *Combinatorial optimization: Algorithms and complexity*. Courier Corporation.
- Parashar, S., Clarich, A., Geremia, P., & Otani, A. (2010). Reverse multi-objective robust design optimization (r-mordo) using chaos collocation based robustness quantification for engine calibration. In *13th AIAA/ISSMO multidisciplinary analysis optimization conference* (p. 9038).
- Peebles, D., & Ali, N. (2015). Expert interpretation of bar and line graphs: The role of graphicacy in reducing the effect of graph format. *Frontiers in Psychology*, 6, 165725.
- Pessoa, M. E. B. T., Roselli, L. R. P., & de Almeida, A. T. (2022). Using the fitradeoff decision support system to support a Brazilian compliance organization program. *Information Systems Frontiers*, 1–16.
- Pop, P. C., Cosma, O., Sabo, C., & Sitar, C. P. (2023). A comprehensive survey on the generalized traveling salesman problem. *European Journal of Operational Research*.
- Preference, A. B. (2024a). Decideit - a tool for decision analysis and interpretation. <https://www.preference.nu/decideit/>. Accessed 07 Mar 2024.
- Preference, A. B. (2024b). Helision: Complex decision support, risk management tool, policy framework with customized consultancy. <https://helision.com/index.html>. Accessed 07 Mar 2024.
- Purchase, H.C., Cohen, R.F., & James, M. (1996). Validating graph drawing aesthetics. In *Graph drawing: Symposium on graph drawing, Gd'95 Passau, Germany, September 20–22, 1995 proceedings 3* (pp. 435–446).
- Ratwani, R. M., & Gregory Trafton, J. (2008). Shedding light on the graph schema: Perceptual features versus invariant structure. *Psychonomic Bulletin & Review*, 15(4), 757–762.
- Roselli, L. R. P., de Almeida, A. T., & Frej, E. A. (2019). Decision neuroscience for improving data visualization of decision support in the fitradeoff method. *Operational Research*, 19, 933–953.
- Roy, B. (1996). Problematics as guides in decision aiding. *Multicriteria methodology for decision aiding* (pp. 57–74). Boston: Springer.
- Roy, B., & Vincke, P. (1981). Multicriteria analysis: Survey and new directions. *European Journal of Operational Research*, 8(3), 207–218.
- Saaty, T. L. (1989). Group decision making and the ahp. *The Analytic Hierarchy Process: Applications and Studies*, 59–67.
- Saini, B. S. (2022). Pioneering techniques to tackle challenges of interactive multiobjective optimization. JYU dissertations.
- Siebert, J. U., Kunz, R. E., & Rolf, P. (2021). Effects of decision training on individuals' decision-making proactivity. *European Journal of Operational Research*, 294(1), 264–282.
- Siraj, S., Mikhailov, L., & Keane, J. A. (2015). Priest: An interactive decision support tool to estimate priorities from pairwise comparison judgments. *International Transactions in Operational Research*, 22(2), 217–235.
- Spear, M. (1952). *Charting statistics*. McGraw-Hill.
- Tague, N. (2004). *The quality toolbox*. Quality Press.
- Tamassia, R. (2013). *Handbook of graph drawing and visualization*. CRC Press.
- Tanimoto, S. L., Itai, A., & Rodeh, M. (1978). Some matching problems for bipartite graphs. *Journal of the ACM (JACM)*, 25(4), 517–525.
- Thurstone, L. L. (1954). The measurement of values. *Psychological Review*, 61(1), 47.
- TransparentChoice Ltd (2024). Project prioritization and decision support software. <https://www.transparentchoice.com/> Accessed 07 Mar 2024.
- Turán, P. (1977). A note of welcome. *Journal of Graph Theory*, 1(1), 7–9.

- Vergara-Solana, F., Araneda, M. E., & Ponce-Díaz, G. (2019). Opportunities for strengthening aquaculture industry through multicriteria decision-making. *Reviews in Aquaculture*, *11*(1), 105–118.
- Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, *85*(411), 664–675.
- Weistroffer, H. R., & Li, Y. (2016). Multiple criteria decision analysis software. In *Multiple criteria decision analysis: State of the art surveys* (pp. 1301–1341).
- Zhen, L., Li, M., Cheng, R., Peng, D., & Yao, X. (2017). Adjusting parallel coordinates for investigating multi-objective search. In *Simulated evolution and learning: 11th International Conference, seal 2017, Shenzhen, China, November 10–13, 2017, Proceedings* (Vol. 11, pp. 224–235).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.