Eurographics Conference on Visualization (EuroVis), Posters Track (2017)
A. Puig Puig and T. Isenberg (Editors)

DcPairs: A Pairs Plot Based Decision Support System

Evanthia Dimara[1], Paola Valdivia[2] and Christoph Kinkeldey[1]

[1] Inria, France
[2] University of São Paulo, Brazil

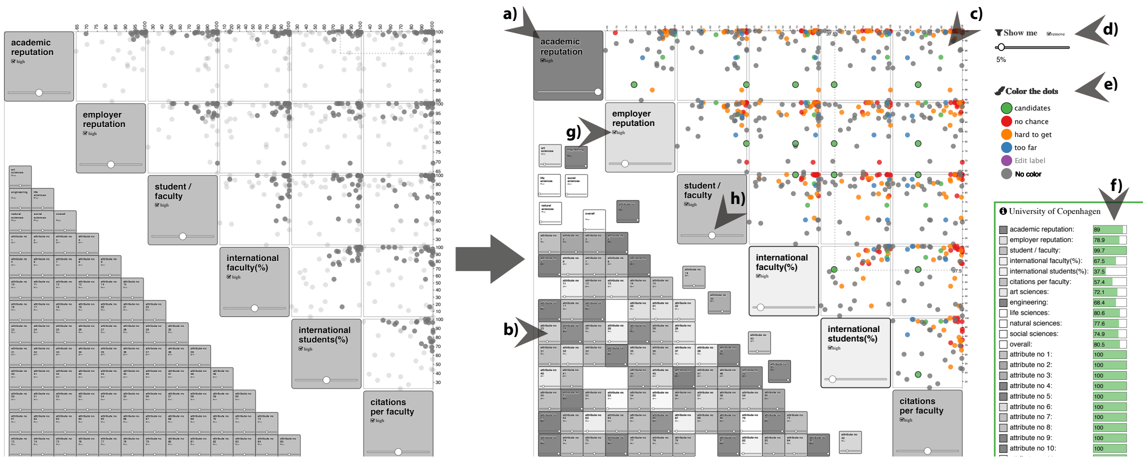

Figure 1: DcPairs *overview: Initial view before the user assigns importance weights (left). Use case scenario described in section 2 (right)*

Abstract

Visualizations designed to support multi-attribute decisions often use colors to encode the identity of the attributes. This approach facilitates mapping of attributes across multiple coordinated views but it has certain limitations: colors often communicate semantics (e.g., red stands for "danger") deemed to influence the user's preference, and qualitative color palettes are of limited scalability. We are currently developing a tool with an alternative approach, DcPairs*: a pairs plot based decision making support tool that employs a compact overview of the decision space and uses visual encodings that communicate uncertainty and suboptimal preference elicitation. Instead of encoding the identity of attributes we use colors for user-authored annotations to support the decision making process. A use case scenario of a prospective undergraduate student choosing a university from the "QS world university ranking" dataset illustrates the functionality of the tool.*

Categories and Subject Descriptors (according to ACM CCS): Visualization System and Toolkit Design, Scalability Issues, Multidimensional Data

1. Introduction and Background

Decision making can be a complex process involving a high number of *attributes* to take into account. An example is a real estate purchase that may be based on price, size, the number of rooms, location (distance to work, schools, public transport, neighborhood quality), etc. Since there is rarely a single optimal alternative, people often need to consider trade-offs between different attributes (e.g., price vs. size) for which personal preferences play an important role. Visualization systems for decision support assist users in finding the best alternative by letting them compare the alternatives according to their preferences [GLG*13, PSTW*17, CL04]. Such preferences are expressed by assigning a personal importance score to each one of the attributes (often called *attribute weights*).

A critical point in the design of decision support tools is that, apart from representing the choice alternatives (data points), design emphasis is also given to the attributes' representation. Most

DOI: 10.2312/eurp.20171165

decision support tools choose to encode the attribute identity with a distinct color to be able to identify and explore them across multiple views [GLG*13, PSTW*17, CL04, AWL*15].

Color is a straightforward way of a uniform attribute encoding, but the number of qualitative color categories is very limited [War12, Mun14] and so is the number of attributes that can be handled by the decision support systems using this approach. At the same time color always implies semantics (e.g. red stands for "danger" [War12]) which can bias the perceived importance of an attribute and affect the quality of decisions [Ber16].

If color is not used to represent attribute identities, its semantics can instead be utilized to support user-authored *annotations*. Annotations in information visualization are considered as new attributes or missing information that users add to the data [Mun14, BM13]. However, even though visualization tools have explored the benefits of annotations in visual analysis [ZGB*17, EB12, CBY10] we are not aware of their use in decision support visualization tools.

We are currently developing DCPAIRS, a visualization tool that assists decisions involving a high number (up to hundreds) of attributes and lets users to annotate alternatives during the decision process.

2. DCPAIRS tool

DCPAIRS is a pairs plot based decision support tool supporting user-authored annotations. It shows decision alternatives in a *generalized pairs plot* [EGS*13] in which quantitative attributes are shown in scatterplots, [Cle93] (Figure 1c) and categorical attributes can be displayed as mosaic plots [HK84] or side-by-side boxplots [Tuk77]. All attributes are encoded in shaded boxes; six in the diagonal for a detailed view (Figure 1a) and the rest in a pixel map overview (Figure 1b). The attributes' importance weights are encoded with a continuous gray scale to express that preferences are often relative tendencies rather than precise values [PSTW*17]. In the following we describe the interactive features of DCPAIRS in a use case scenario illustrating how the use of a pairs plot for attribute pairs can provide critical insights, supported by user-authored annotations (Figure 1e) using a qualitative color scheme [HB03].

We downloaded the "University Rankings" from [lin], containing the "QS world rankings 2013" of 906 institutions. The ranking is based on the weighted sum of the attributes: "academic reputation" (40%), "faculty student ratio" (20%), "citations" (20%), "employer reputation" (10%), "international faculty" (5%), and "international students" (5%) (Figure 1a). The dataset contains additional attributes not considered in the ranking: performance in"arts", "humanities", "engineering", and "life and natural sciences" (Figure 1b). The extra boxes in Figure 1b demonstrate the tool's scalability and are not part of the real dataset.

Use Case: Vangelis, a prospective undergraduate student from Athens, is searching for universities to apply for. He loads the dataset into DCPAIRS and takes a look at the gray values and slider position (Figure 1h) of the attributes to understand which attributes are considered in the default ranking score. The attribute pixel map (Figure 1b) gives a quick overview of their overall number and relative importance. He sees that most emphasis is laid on academic reputation, about half on citations and on the faculty-student ratio, less on employer reputation, and a lot less on international faculty and students. Even though the subject areas are deemed unimportant (signified by the white color), Vangelis is mainly interested in engineering and, to a lesser extent, in art subjects, so he changes the respective sliders to the maximum and to a low value.

Having a limited budget for application fees, Vangelis can only select three universities to apply for and has the strategy to choose a bit less competitive yet suitable institutions to increase his chances of acceptance. He moves the *threshold slider* (Figure 1d) to filter all but the top 1% institutions assuming that these are the most competitive ones and tags them with red and a newly defined label "no chance" (Figure 1e, c) . He moves the slider to 2% and labels the yet untagged ones with orange and "hard to get". Vangelis finally sets a 5% threshold to see institutions he considers as more realistic. Since most attributes do not seem to be much correlated (Figure 1c) (apart from "international faculty" and "international students"), he understands that he needs to carefully consider each one individually. By hovering over the dots, the *inspector* (Figure 1f) allows Vangelis to retrieve more detailed information, and he tags the interesting institutions in green color (label "candidates"). He also uses the orange and red dots as baseline to identify candidates similar to the top institutions.

Another aspect Vangelis cares about is to find an institution at a location that will allow him to travel to his family from time to time. This information is not in the dataset, but he extracts it from the institution names. He tags the ones extremely far from his hometown (e.g., in China) with blue and the label "too far". Using the color tags he finds his three favorite "safe" choices: "University of Copenhagen", "University of Manchester" and "Kings College", but he decides to also apply to "University of Oxford" (in red) trying his luck with at least one of top institutions.

Vangelis is browsing the pixel map overview to identify other attributes that could affect his decision. Since he is interested in low tuition fees, he changes the default *direction-check box* (Figure 1g) of the "tuition fee" attribute (hypothetical attribute, not part of the real dataset) from a "high" to a "low" direction and drags it to the diagonal detailed view to see it paired with "academic reputation". He observes that "fee" is correlated to "reputation" in most cases, but that a few top institutions do not follow this trend. Vangelis interprets this as a sign that the latter have a friendlier policy towards student's budget and reconsiders his choices.

3. Status and Future Work

DCPAIRS is a pairs plot based decision making support tool that is currently under development. A working prototype of the tool with the functionality described above already exists. However, its performance has to be improved to handle higher amounts (hundreds) of decision alternatives fluently. An important part of the development will be a user study to assess the advantage of a pairs plot approach and the potential benefits of providing user-authored annotations during the decision making process.

4. Acknowledgments

We thank Pierre Dragicevic, Anastasia Bezerianos and our reviewers for their precious feedback on the paper.

References

[AWL*15] Aseniero B. A., Wun T., Ledo D., Ruhe G., Tang A., Carpendale S.: Stratos: Using visualization to support decisions in strategic software release planning. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), ACM, pp. 1479–1488. 2

[Ber16] Bera P.: How colors in business dashboards affect users' decision making. *Communications of the ACM 59*, 4 (2016), 50–57. 2

[BM13] Brehmer M., Munzner T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (Dec 2013), 2376–2385. doi:10.1109/TVCG.2013.124. 2

[CBY10] Chen Y., Barlowe S., Yang J.: Click2annotate: Automated insight externalization with rich semantics. In *2010 IEEE Symposium on Visual Analytics Science and Technology* (Oct 2010), pp. 155–162. doi:10.1109/VAST.2010.5652885. 2

[CL04] Carenini G., Loyd J.: Valuecharts: analyzing linear models expressing preferences and evaluations. In *Proceedings of the working conference on Advanced visual interfaces* (2004), ACM, pp. 150–157. 1, 2

[Cle93] Cleveland W. S.: *Visualizing data*. Hobart Press, 1993. 2

[EB12] Elias M., Bezerianos A.: Annotating bi visualization dashboards: Needs & challenges. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 1641–1650. 2

[EGS*13] Emerson J. W., Green W. A., Schloerke B., Crowley J., Cook D., Hofmann H., Wickham H.: The generalized pairs plot. *Journal of Computational and Graphical Statistics 22*, 1 (2013), 79–91. 2

[GLG*13] Gratzl S., Lex A., Gehlenborg N., Pfister H., Streit M.: Lineup: Visual analysis of multi-attribute rankings. *IEEE transactions on visualization and computer graphics 19*, 12 (2013), 2277–2286. 1, 2

[HB03] Harrower M., Brewer C. A.: Colorbrewer. org: an online tool for selecting colour schemes for maps. *The Cartographic Journal 40*, 1 (2003), 27–37. 2

[HK84] Hartigan J. A., Kleiner B.: A mosaic of television ratings. *The American Statistician 38*, 1 (1984), 32–35. 2

[lin] Lineup - legatum prosperity index. https://lineup.caleydoapp.org. Accessed: 2017-02-24. 2

[Mun14] Munzner T.: *Visualization Analysis and Design*. CRC Press, 2014. 2

[PSTW*17] Pajer S., Streit M., Torsney-Weir T., Spechtenhauser F., Muller T., Piringer H.: Weightlifter: Visual weight space exploration for multi-criteria decision making. *IEEE transactions on visualization and computer graphics 23*, 1 (2017), 611. 1, 2

[Tuk77] Tukey J. W.: Exploratory data analysis. 2

[War12] Ware C.: *Information visualization: perception for design*. Elsevier, 2012. 2

[ZGB*17] Zhao J., Glueck M., Breslav S., Chevalier F., Khan A.: Annotation graphs: A graph-based visualization for meta-analysis of data based on user-authored annotations. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 261–270. 2