Americas Conference on Information Systems AMCIS2013 Chicago Multidimensional Charts

Research-in-Progress

Andrés Díaz López
University of Mississippi
alopez@bus.olemiss.edu

Xiang Guo
University of Mississippi
xguo@bus.olemiss.edu

David Pumphrey
University of Mississippi
dpumphrey@bus.olemiss.edu

ABSTRACT

This work presents a new visual representation of multidimensional data and compares its usefulness in terms of effectiveness and efficiency with tabular representations. This idea makes two important contributions. First, it shows the feasibility of representing multidimensional data in 2D and 3D charts that are understandable by humans. Second, it builds on the theory of cognitive fit by testing the appropriateness of graphical representations to convey information of complex problems. In particular, it shows multidimensional data representation that has not been tested, perhaps due to the lack of a suitable graphical representation. The charts proposed are better representations of information stored in data warehouses than those provided by data cubes. The proposed format can be used to represent fuzzy variables, and are suitable for implementation in dashboards.

Keywords

Multidimensional, Chart, Dimensions, Attributes, Data, Visual Representation, Cognition, Data Visualization.

INTRODUCTION

The lack of graphical representations for some of the most complex problems with which information technologies' experts deal nowadays is evident and must be addressed. These problems involve data mining processes applied on data warehouses, fuzzy systems, and system dynamics, which, unlike other problems, let humans relying almost completely in the computer manipulations of the data to have a better view of the situation. In opposition, there is a view encouraging us to rely more on the human capabilities in the visual domain (Shneiderman and Maes, 1997). This paper makes a contribution towards this end by delivering a tool that facilitates visualizing multidimensional data which enhance the interaction between humans and computers in decision making processes.

This paper proceeds as follows: The first section of the paper addresses the theoretical background, hypotheses and conceptual model. Then, it shows the development of the multidimensional chart responding to the first question. This development includes background, assumptions and limitations of the chart. The next section proposes a test of the benefits of the new chart, the approach to collect the required data, and the assumptions and limitations of the test. Finally, we explain the contributions of this paper and propose the direction of future research.

THEORETICAL FOUNDATION

Edward Tufte suggests that data visualization should be developed to help retention of information (Zachry & Thralls, 2004). Furthermore, Slovic says that decision makers tend to use only information that is explicitly displayed (i.e. concreteness) (Vessey, 1991). These authors suggest that graphs affect the cognitive processes of humans, which, in turn, affect the performance outcomes of problem solving. This is consistent with the work of Gigerenzer and Hoffrage (1995), and Kahneman and Tversky (1974, 1979, and 1986) who demonstrated that different framings of problems can lead to different outcomes.

There is an ongoing debate between the performances of graphical and of tabular representations. While Lukas says that the graphics system had a dramatic effect on the decision-making process (Lukas, 1981), Ives (1982) agrees with the lack of empirical support for this contention. More recently, Kelton, Pennington and Tuttle (2010) find that the dilemma persists and cite several studies as examples (e.g. DeSanctis 1984; Remus 1984, 1987; Montazemi and Wang 1988). However, they concede that graphs are more useful for complex tasks with large amounts of data, whereas tables are more suitable for precise estimates and simple tasks. Therefore, the multidimensional information of complex systems and the abundant data stored in data warehouses may be better explained by graphs than by tables. This is explained by the theory of cognitive fit (Vessey, 1991), which introduces the mediation of the task characteristics to explain the contradictory results obtained by different authors comparing tabular (symbolic) representations and graphical (spatial) representations. In short, this theory explains that representations should match the task characteristics, i.e., symbolic representations are more appropriate for symbolic tasks, whereas spatial representations are better for spatial tasks.

RESEARCH QUESTIONS

If a multidimensional representation can be created and if many data are representable, such comparison may and should be done. Accordingly, the research questions are:

1. Is it possible to represent multidimensional data, specifically, data that has more than four dimensions in 2D and 3D visualizations?
2. If so, is this representation a better instrument than the tabular representations currently used to represent multidimensional data?

HYPOTHESES

The theoretical reference being challenged is the classical view of three coordinates on which most current charts are based, i.e. the Cartesian system proposed by Rene Descartes in the 17th century. Obviously, there is nothing wrong with using this form of data representation when appropriate. We contend that by freeing ourselves of the orthogonal tridimensional space, we might be able to visually represent more than three dimensions in 2D and 3D physical displays. This leads to the first hypothesis:

Hypothesis 1: It is possible to obtain a visual representation of multiple dimensions suitable to manage abundant data by using an approach that diverges from the Cartesian and Euclidean paradigms.

In addition, we believe that there is evidence of the benefits of this graphical representation over tabular presentations. Therefore, we propose the following second hypothesis.

Hypothesis 2: The more spatial the representation of a complex problem, the more effective and efficient its solution is.

By complex problem we mean a multidimensional problem. The spatial condition suggested is a “graphicalization” of the information; on a continuum ranging from tabular representation on one end to graphical representation on the other.

The effectiveness of this representation of data will be measured in terms of the proportion of correct answers, and the short and long term retention of information. Efficiency will be measured in terms of the speed of selection of the appropriate solution and of the user friendliness of the chart. This produces a two-part hypothesis 2 with additional sub-sets of hypotheses:

H2A: Graphicalization is positively related with the effectiveness of the decision making process.

H2A1: Graphicalization is positively related with the accuracy of the decision making process.

H2A2: Graphicalization is positively related with short-term retention of information.

H2A3: Graphicalization is positively related with long-term retention of information.

H2B: Graphicalization is positively related with the efficiency of the decision making process.

H2B1: Graphicalization is positively related with the speed of making a proper decision.

H2B2; Graphicalization is positively related with user friendliness of the visual representation.

CONCEPTUAL MODEL

The model contains one independent variable, Graphicalization, and one dependent variable, Decision Making Performance. Both are continuous variables measured low to high and are expected to be positively related. Performance has two dimensions, Effectiveness and Efficiency, which are measured with three and two indicators respectively. The indicators of Effectiveness are Accuracy, Short-term Retention and Long-term Retention. The indicators of Efficiency are Speed of Accurate Decision Making and User Friendliness. Details of this model are explained in the methodology section.

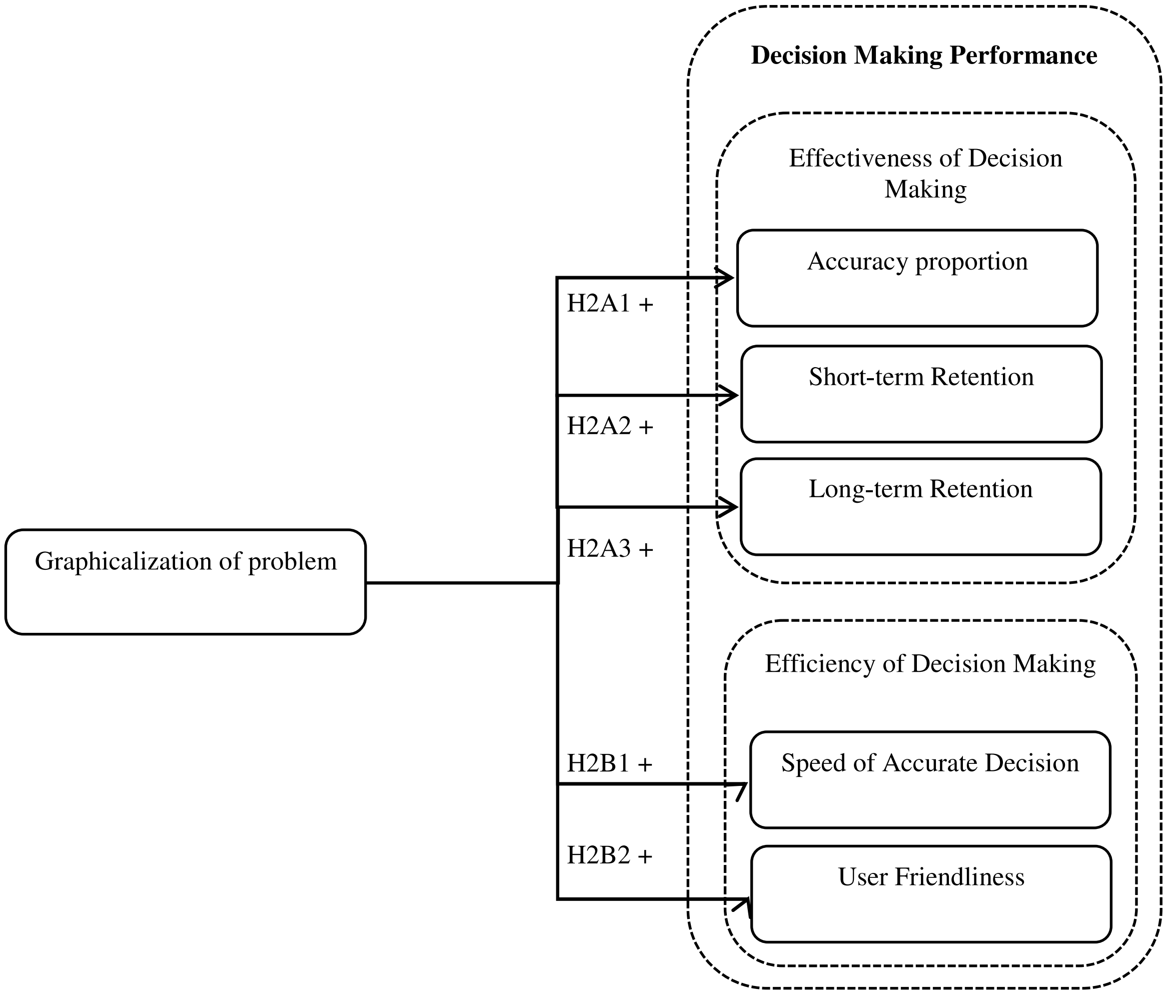

Figure 1. Conceptual Model

METHODOLOGY FOR THE FIRST QUESTION

The present multidimensional chart differs from the Cartesian paradigm in two aspects. First, dimensions do not have to be orthogonal, only divergent. This is because in reality it makes no sense to establish any angular distance among most of the dimensions we deal with in the managerial sciences. Second, it is the object what is fixed, not the dimensions. The rationale here is that in the world we perceive the observation of an object gives us an idea of many more attributes or dimensions than those that are possible to depict in a Cartesian or Euclidean chart. For instance, by looking at a person we have an idea of his or her age, gender, height, weight, skin tone, etc. This approach sets a way to see things in which the observation is a pivot through which dimensions cross in divergent directions. For this reason, we can call this proposal as Divergent Attributes Locked on a Pivot.

CURRENT THEORETICAL PARADIGM

Euclidean space can provide information of up to three attributes of an object. This is possible because in the Euclidean space the three dimensions are orthogonal, thus, an intersection of the three dimensions in a single point can be found by projecting lines that are perpendicular to the three axes. The addition of more dimensions makes us appeal to other visual artifacts such as colors and shapes in order to visually communicate more dimensions of the data. These artifacts, unfortunately, are limited and the result is far from parsimonious. Tables are simple in terms of descriptive information, but not a good way to communicate relative information (Kelton et al. 2010).

Pivot tables represent more than three attributes. However, these attributes should be sub sets of the dimensions and the information must be discretized, which reduces its original richness. Yet, information can be redundant.

Data Warehouses data are usually explained with a tridimensional cube, as in Datta and Thomas (1999); Messaoud, Boussaid, Rabaséda, and Missaoui, (2006); and Gray, Chaudhuri, Bosworth, Layman, Reichart, Venkatrao, Pellow and Pirahesh (1997). These representations are adequate to explain the drill down and roll up abilities of OLAP applied on data warehouses. However, warehouses manipulate many more dimensions. Thus, the data cubes are not the adequate visual representation of Data warehouses data.

Other attempts to represent more than three dimensions try to work with bended planes that clutter the spaces to a point that makes the graphs less clear than the raw data. The addition of dimensions in these cases confuses rather than clarify.

PROPOSED PARADIGM - DIVERGENT ATTRIBUTES LOCKED ON A PIVOT

As explained before, one thing is to think that dimensions exist and any new object has a location in such a multidimensional space. Another idea is to think that those dimensions appear just when an object is realized. Thus, looking at the object is the way to realize a specific value of each dimension by which we can characterize the object. We do not know the relationship among those dimensions, but we do know the value of each dimension that can be attributed to the object. Therefore, we can represent these dimensions not by fixed axes in which we locate the value for the object, but by axes that slide along a fixed object. As the relationship among dimensions is unknown, we just need to differentiate them by using different angles among them, not necessarily right angles.

Let us take, for example, the data of a transaction from a data mining model designed to predict the likelihood of fraud with credit cards (Bacon, Diaz, Neeley, 2011). This example is taken from a model developed for a finance institution in the United States. In this case, our unit of analysis or what has been called so far "the object" is a transaction. Let us take four dimensions characteristic of any transaction, e.g. the amount of dollars transacted, the day of the week of occurrence, the zip code of the location where the transaction took place, and the consecutive number identifying the transaction. Let us say that the values for one specific transaction are:

Dimension	Value	Min.	Max.
ID	1004	0	5000
Amount	5000	-10000	10000
Zip Code	38655	500	99999
Week Day	Tuesday	Monday	Sunday

Table 1. Data for Transaction A

In order to be able to represent values in the proposed chart, we must follow these general steps:

1. Choosing the attributes to be represented (showed in the table)

2. Setting the highest and lowest limits of each attribute (showed in the table)

3. Order the categorical attributes in a continuum or group them in a dimension

4. Locate the attributes at different (preferably symmetric) angles and at our convenient visualization

5. Draw the line representing each attribute over a fixed point (the pivot) in the center of the visual area with the value of the attribute matching the pivot.

Thus, the data of our example can be represented as follows:

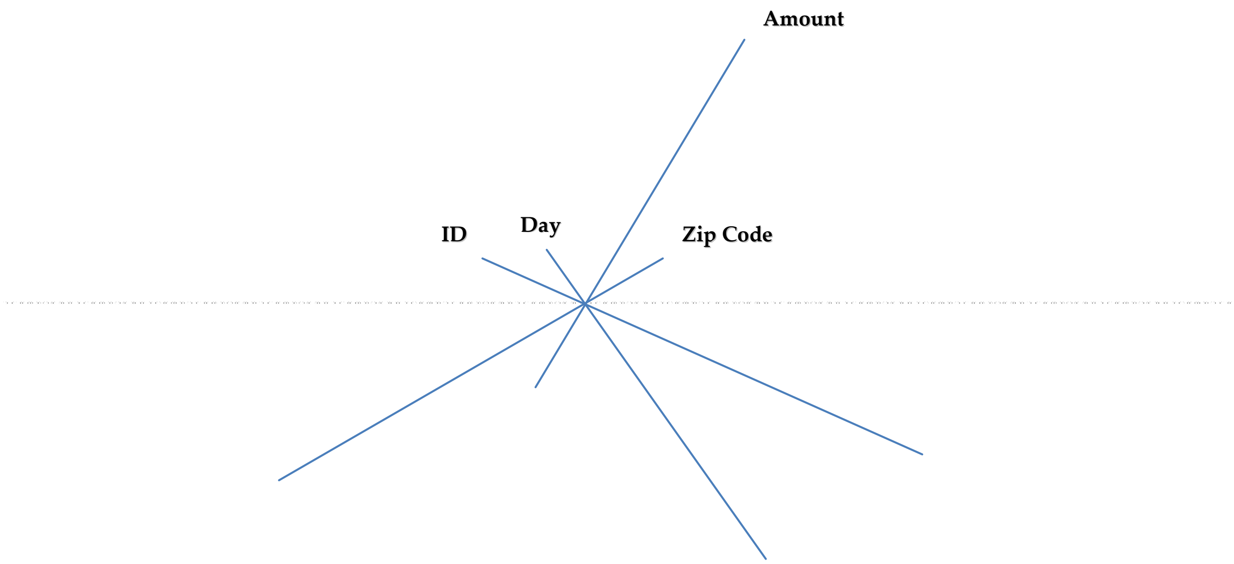

Figure 2. Multidimensional Chart for One Observation

The explanation of this chart is:

The transaction value of A is $5000, which is $15000 away from the minimum possible value of negative $10000; the total length of this dimension is $20000, so $15000 represents ¾ of it. This is why ¾ of the segment are above the pivot. Likewise, it is $5000 away from the maximum of $10000 (note that maximums are located at the bottom and minimums at the top); thus, ¼ of the segment is below the pivot.

The dimension day of the week is divided into 7 equal segments, thus, it follows that the value of a day lies in the middle of the respective sub segment. Therefore, the value Tuesday is 1 ½ segments away from the extreme. According to the recommended convention, Monday is located above the line because it is the minimum of that attribute.

By the same token, the value of ID, which is approximately 1/5 of the total, and the value of Zip Code, which is approximately 1/3 of its total are represented by 1/5 and 1/3 segments above the pivot respectively.

Some assumptions have to be considered for a correct construction and interpretation of the chart. Likewise, some conventions are proposed in order to facilitate visualization and common understanding.

ASSUMPTIONS

1. Dimensions are independent from each other. Therefore, the angular positions do not indicate any measurable relationship among dimensions.
2. The value of the object in a dimension is given by the displacement from the pivot to one extreme of the line.
3. Categorical attributes have a fixed number of factors.
4. Categorical variables are organized as ordinal variables whose order is assigned at the convenience of the user.
5. Levels of variables can be considered as an independent attribute. In the variable color, green can be an attribute by itself and an observation is qualified in terms of its greenness. This allows representation of fuzzy variables.
6. Time can be seeing as a continuous infinite variable or as a categorical variable with sub segments months, days, etc.

RECOMMENDED CONVENTIONS

1. All lines representing dimensions should have equal length. Hence, each dimension represents a hundred percent of its own scale, not comparable with any other dimension.
2. Minimum values of dimensions should be set to the extreme above the horizontal axis, so, the segment portion above this axis represents the value of the object in such a dimension.

3. Attributes of a dimension are represented by sub segments of dimensions. But, if attributes overlap or are not seen as continuous, they may be treated as independent attributes, perhaps as different dimensions. For example, the attributes months and days of dimension time are better represented such that days are sub segments of months and months sub segments of years. But, if the interest is in seasonal phenomena, months can be an independent dimension.

The analyst can freely organize the categorical zip codes as consecutive numbers, but he or she can also organize them by the distance to a zip code of reference (List of ZIP code prefixes, 2012), hence, getting information of likelihood of fraud in terms of remoteness.

The ability to set levels of attributes as independent attributes by themselves, permits the more realistic representation of attributes such as ethnicity (because one person can belong to different ethnicities at the same time). For this reason, it is expected that those working with fuzzy variables take particular advantage of this visualizations.

3D CHART

An idealistic 3D representation of this chart can be attained by considering the pivot as the frontal view of another dimension. In our case, the attribute Day is a categorical variable indicating the day of the week in which the transaction takes place. But, it does not represent a chronological placement of the event in a continuous dimension. The exact moment in time can be represented by a time dimension hidden in the pivot. We can reveal this dimension by pushing our plane, as shown below.

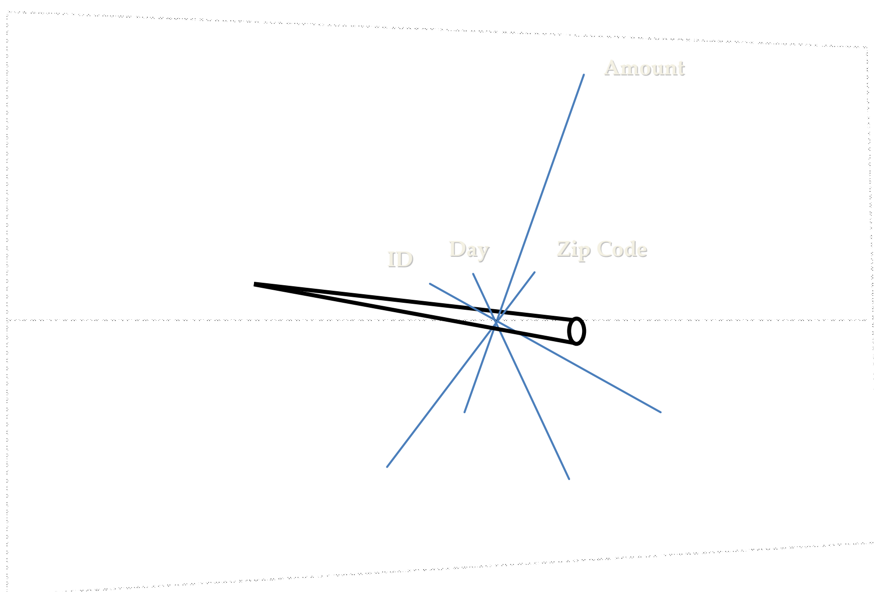

Figure 3. Pivot as the Time Dimension

In this graph, the line representing time is drawn with volume to indicate that the circular area at the end of this line was our pivot when we saw the plane from the front. If we consider this dimension as time, that end of the line will represent the present moment. Therefore, we can have transactions in different points along this line. Notice that if we continue pushing our 2D plane up to 90 degrees, and rotate our original dimensions as a propeller we obtain sort of a Cartesian chart of whichever attribute is in the vertical axis versus time. Thus, we can see the trend line of that attribute by marking other observations along the timeline.

Although time is a sound dimension to use the pivot as a dimension in a 3D chart, other attributes can be chosen to be represented in the pivot-as-dimension view according to the user's preference.

MULTIPLE OBSERVATIONS CHART

Representing several observations is achieved by using one object's chart as a reference and then marking the value of a second object on each of the already depicted dimensions. For example, we can compare a new transaction B with the already drawn transaction A. Let us say that the data for transaction B are:

Dimension	**Value**	**Min.**	**Max.**
ID	2500	0	5000
Amount	-10000	-10000	10000
Zip Code	79015	500	99999
Week Day	Sunday	Monday	Sunday

Table 2. Data for Transaction B

The chart for these two transactions will look like this:

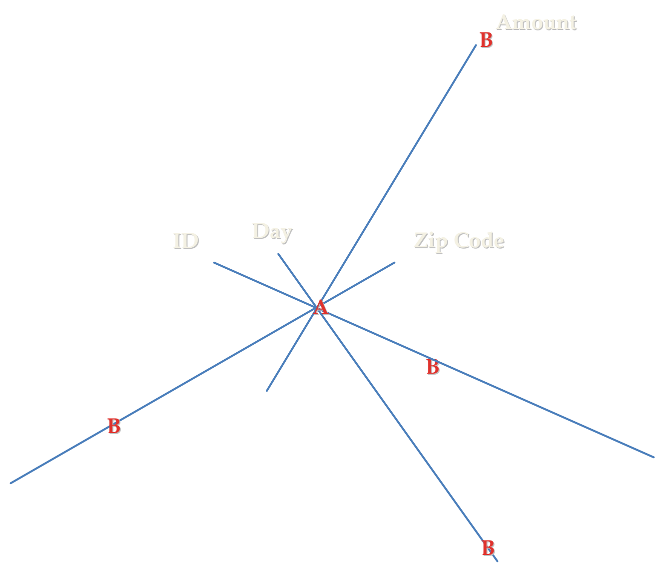

Figure 4. Multidimensional Chart with One Observation as the Pivot

In this case, the values of B located above the horizontal line indicate lower values for B than for A (the pivot), and values below the pivot are bigger values for B than for A. This should be read as "more line is above B than above A, hence, B is bigger".

More transactions can be added with reference to A, obtaining a plot of multidimensional data of which, heretofore, no reference has been found in the literature reviewed. However, the multidimensional plot does not need to be in reference to one specific observation. A generalization of this chart can assume that the reference point is the mean, the media, the bisection or even a desired or expected value for each dimension. For instance, if there is an ideal value for every dimension, as shown in the second column of table 3 below, we can obtain the chart shown if Figure 5.

Dimension	Ideal	Min.	Max.
ID	2500	0	5000
Amount	10000	-10000	10000
Zip Code	70000	500	99999
Week Day	Wednesday	Monday	Sunday

Table 3. Expected (Ideal) Values

All the observations are marked on the dimensions and the final chart looks like this in Figure 5.

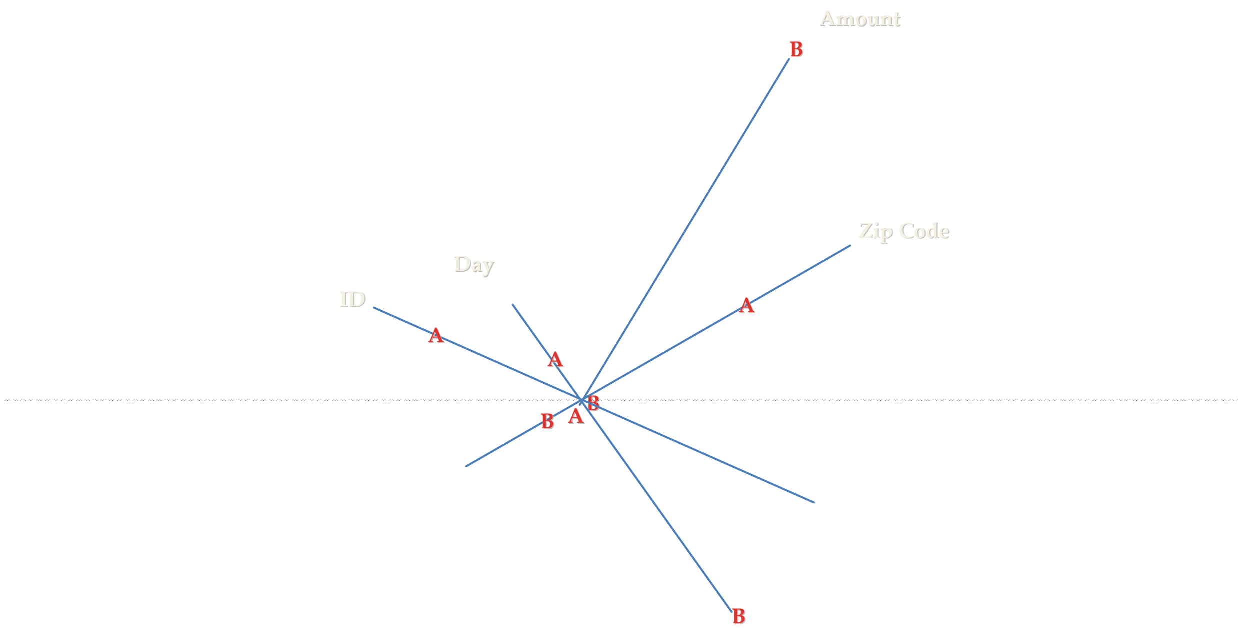

Figure 5. Multidimensional Chart with Ideal Values as the Pivot

Here the dispersion of the data from the pivot shows that desired values are not being achieved whereas a high concentration towards the center shows good prediction or good performance. The ability of zooming out and in is desirable in a software application developed to produce these charts. But, other tricks may also help when high density of observations close to the pivot are found, for example hiding some attributes or expanding the diameter of the pivot point.

LIMITATIONS

As in other charts, visual space is limited. It is easy to see that coincident points towards the pivot may clutter the chart. Thus, different colors, numbers or symbols can be used to differentiate a moderate amount of different observations. However, specific observations are hardly of more interest than the shapes that abundant data can take. For this reason differentiating observations is not as important as detecting agglomerations and dispersions. Still, a software application with zooming functions should help solving this limitation. The most salient limitation at this moment then is not having the software application allowing the construction of the proposed chart.

Many Data Example

In the following hypothetical example, there is a multidimensional chart of attributes that different cities are following closely in order to measure their abilities to attract investors who are interested in developing business related to information technologies. In this case one city A is interested in eight attributes and compares its performance against that of other fifteen cities that are considered as direct competitors in attracting such investors. The pivot represents the values in each attribute for city A. the attributes are Amount of employees and students in IT, Number of IT corporations, Number of Apps provided by public entities, Size of the city in inhabitants, a Multilingual Index indicating the ability of workers to speak several languages, and Coverage of Free WiFi.

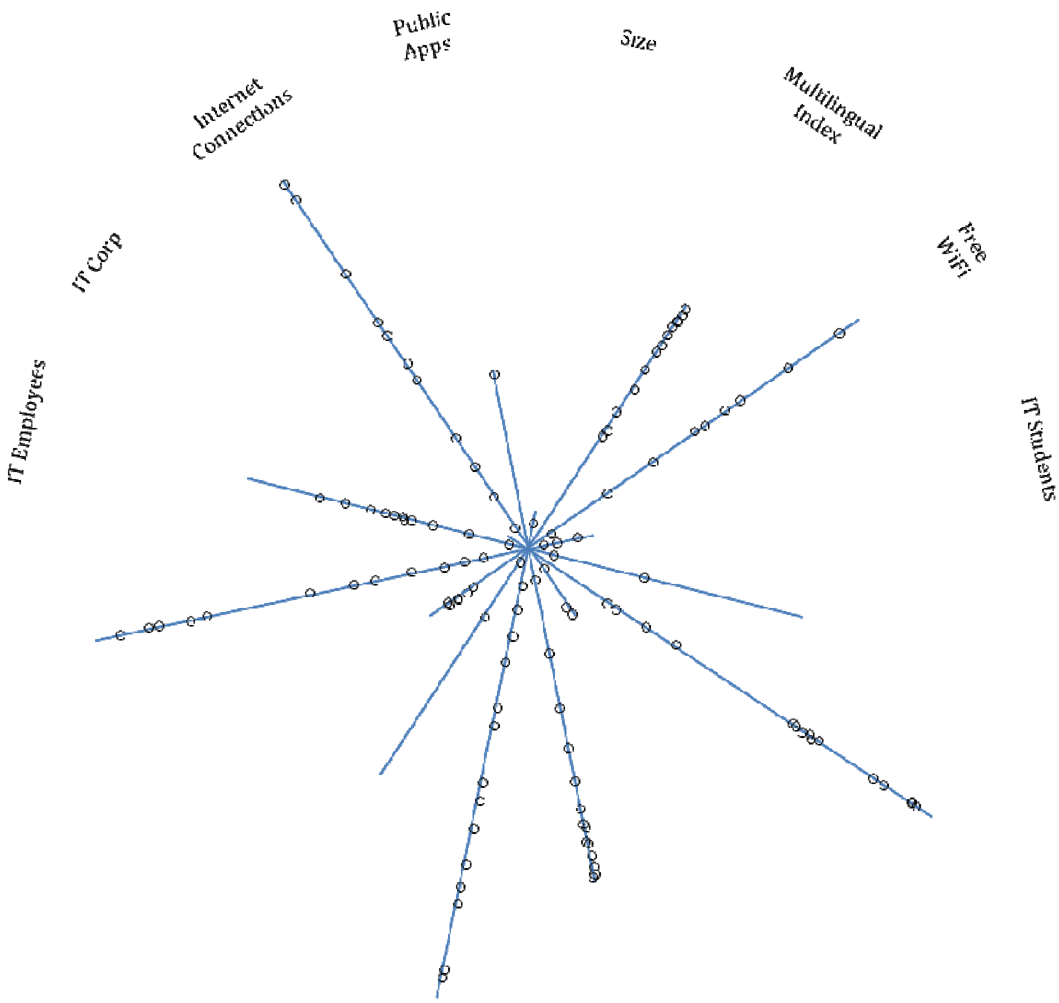

Figure 6. Multidimensional Chart with City A Values as the Pivot

As shown in the chart, city A knows its performance in each dimension, but also the distribution of other cities in the attributes of interest. For example, the city officials using this chart may be able to see that in their small city there are few IT corporations despite their comparative advantage in connectivity. They may consider highlighting this advantage among investors. They can hide or reveal attributes and observations therefore obtaining benchmarking information in reference to other selected cities in selected attributes. Presumably, these officials will set a fixed angular location of the attributes and get used to the dispersion of observations. These dispersions and the sliding movement of attributes gives them a comprehensive, simultaneous and immediate information of their performance in several dimensions in reference to their competitors, which makes this chart particularly suitable for dashboards.

An Alternative Chart

The limitation for high density of observations towards the pivot is solved with one variation of this chart, under the same principle of dimensions sliding on a value of reference, the Parallel Attributes on a Line, wherein the vertical lines represent the attributes and their intersections with the horizontal line represent the values of reference that were previously located in the pivot. This alternative chart is shown in Figure 7.

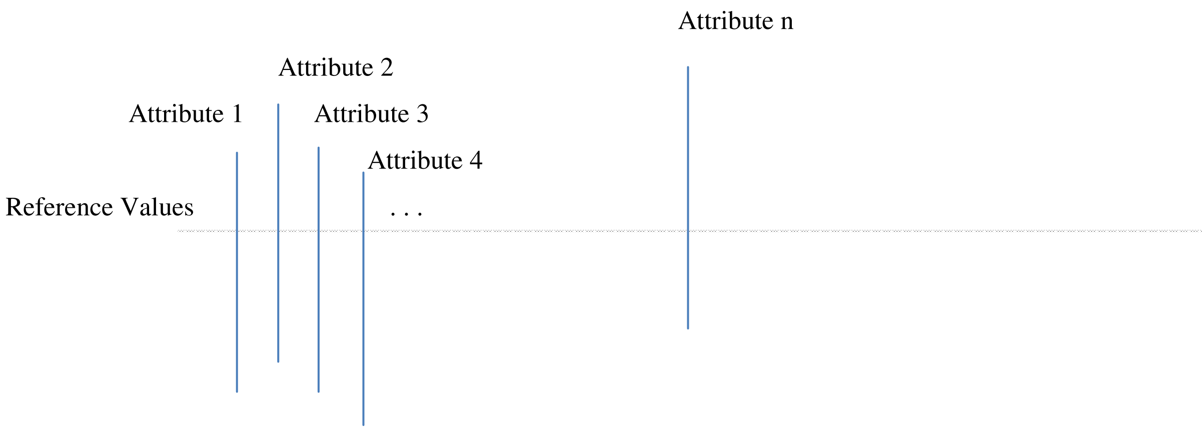

Figure 7. Multidimensional Chart of Parallel Attributes on a Line

METHODOLOGY FOR THE SECOND QUESTION

The proposed test for the second hypothesis consists of an experiment where randomly selected individuals will be given a complex problem, i.e. a multidimensional problem, with abundant data in two formats: tabular and graphic. These individuals will be randomly assigned to two groups. The control group will have access to tabular information only. The experimental group will have access to graphical information only. Both groups will answer questions that can be responded accurately with the sole use of their respective presentation of the problem.

Two dimensions of performance are measured: effectiveness and efficiency. Effectiveness will be measured beyond accuracy as is common in previous researches (Vessey, 1990). It will include retention of information (short and long terms) as recommended by Tufte (Zachry and Thralls, 2004). On the other hand, efficiency will be measured in terms of the speed of accurate answers to questions and in terms of the user friendliness of the data representation.

The inclusion of user friendliness is taken from Dennis (1988) who includes ease of learning and ease of use as qualifying dimensions of graphs. This is relevant because charts are conventions that usually require explanations in order to offer commonly accepted interpretations. Lukas (1981) shows an experiment where users of graphs are outperformed by users of other types of presentations, and attributes this to the lack of experience of decision makers with graphics and CRT terminals. He suggests that users of these systems need to work with the graphics presentations for a while to become familiar with this mode of presentation.

User Friendliness will be measured first by comparing the time it takes individuals to give a general explanation of the characteristics of a population described by the multidimensional chart in contrast to the time it takes them to give the same description when the data is represented by tables. Only individuals not familiar with a specific topic will be tested.

Speed of Accurate Decision will be measured after users manifest understanding of the chart and the tables. Time is expected to decrease as graphicalization increases. Less time means more speed and more efficiency. Thus, efficiency is measured through an experiment comparing the average time taken in giving accurate responses about the information provided in tables with the time taken by those using the new chart when they know how to use it. Thus, at this stage, it is assumed that individuals are familiar with the use of tables as well as with use of the new chart. However, it demands controlling the familiarity of both users of tables and users of charts with the topic represented by the data, as shown in Figure 8. The same design will be used to the test the Accuracy Proportion. In this case, participants will be given a fixed time to answer the set of questions and a measurement will be set to the proportion of correct answers as an indicator of accuracy.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Graphicalization Level</td></tr>
<tr><td>New Chart</td><td>Tables</td></tr>
<tr><td rowspan="2">Topic Familiarity</td><td>Familiar</td><td>Familiar with topic using New Chart</td><td>Familiar with topic using Tables</td></tr>
<tr><td>Unfamiliar</td><td>Unfamiliar with topic using New Chart</td><td>Unfamiliar with topic using Tables</td></tr>
</table>

Figure 8. Test Design for Accuracy and Speed of Accurate Decision

Finally, the same individuals will be asked questions about the given information the day after the experiment and a couple of weeks later. A percentage of accurate responses will give an indicator of short-term retention in the first case and long-term retention in the second case. Retention types and chart types complete a factorial analysis represented in a 2x2 matrix. However, we will include the familiarity with the topic in order to detect any difference in retention associated with this variable. This design delivers eight groups, as shown in Figure 9.

<table>
<tr><td colspan="2" rowspan="2"></td><td rowspan="2">Topic Familiarity</td><td colspan="2">Graphicalization Level</td></tr>
<tr><td>Tabular</td><td>New Chart</td></tr>
<tr><td rowspan="4">Retention Type</td><td rowspan="2">Short-term</td><td>Familiar</td><td>Short-term Retention using Tables and Familiar with topic</td><td>Short-term Retention using New Chart and Familiar with topic</td></tr>
<tr><td>Unfamiliar</td><td>Short-term Retention using Tables and Unfamiliar with topic</td><td>Short-term Retention using New Chart and Unfamiliar with topic</td></tr>
<tr><td rowspan="2">Long-term</td><td>Familiar</td><td>Long-term Retention using Tables and Familiar with topic</td><td>Long-term Retention using New Chart and Familiar with topic</td></tr>
<tr><td>Unfamiliar</td><td>Long-term Retention using Tables and Unfamiliar with topic</td><td>Long-term Retention using New Chart and Unfamiliar with topic</td></tr>
</table>

Figure 9. Test Design for Retention

Data Collection and Sample

The hypotheses of this work will be tested among individuals whose daily activities occur in the business environment. The individuals will be selected randomly and be divided into two groups, those using the new chart and those using tables as source of information. These experiments will help responding the four first sub hypotheses of hypothesis 2. The fifth, user friendliness, will be tested first and will involve comparing users of the new chart with users of tables. Testing user friendliness in a first stage is necessary to guarantee that the users have not had contact with the new chart before undertaking the test comparing the new chart with tabular visualizations.

In order to control by familiarity with the topic in the test of user friendliness, a material about a topic will be prepared an explained to a randomly assigned half of the group. The comparison of groups at two levels of graphicalization (tables and charts), two levels of familiarity with the topic, and two levels of retention, delivers eight groups. We will collect information from at least 10 individuals from each group in order to gain statistical representativeness.

ASSUMPTIONS AND LIMITATIONS

The assumptions have been addressed directly in the development of the chart. However, only experimental cases and rationales can be presented at this moment. Ideally, this evidence can be achieved by providing instructors and analysts of multidimensional data with an application allowing them to build these charts, asking them about their perceptions and recording their uses of the charts. Because particularities of these charts assume that users will locate dimensions freely, give an order to the categories of variables, and even decide on whether separating or not some of those categories as independent attributes, it is uncertain what the users' reactions will be unless an application allowing them to set this preferences is in place.

CONTRIBUTIONS AND FUTURE RESEARCH

This work makes two important contributions. First, it gives a unique support to the decision making process by offering a way to represent more-than-three-dimensional data in the two-dimensional (2D) plane and the tridimensional (3D) space, which has been a challenging enterprise heretofore. By doing this, it disrupts the paradigm of visualization of data by changing the perspective from which observations are viewed and the perception of the relation among the dimensions through which an item is measured.

The second contribution aims to extend the validity of the theory of cognitive fit proposed by Vessey (1990) by testing the better fit of a graphical representation over a tabular presentation to convey information of complex problems in the multidimensional context. There is no evidence of this test up to this moment perhaps because there has not been a graphical representation of multidimensional data.

Future research must be undertaken in the application of these diagrams for representing the many-dimensional, abundant data of data warehouses. These charts can also be appropriate to represent statistical data from variables that are considered orthogonal. This is because the divergent attributes depicted are not assumed to use a common Euclidean space, thus, in the new chart any angular distance can be assumed among attributes, in this case 90 degrees, even though visually they do not look so. This same characteristic makes it suitable for implementation in Dashboards. Finally, this new paradigm encourages the consideration of fuzzy variables used in complex systems.

ACKNOWLEDGMENTS

We want to thank the anonymous reviewers and the mini-track chair for the valuable comments and suggestions to this paper, as well as all authors, committee members, and volunteers for their hard work and contributions to the conference.

REFERENCES

1. Dennis, A. R. (1988). The use of business graphics. ACM Sigmis, Vol. 19, Issue 2. 17–28.
2. Ives, B. (1892, Dec. 01). Graphical user interfaces for Business Information Systems. MIS Quarterly Special Issue 6(4), 15-47.
3. Ebert, D. S., Favre, J. M., and Peikert, R. (2002, Apr.). Data visualization. Computers & Graphics. 26, 2, 207-208.
4. Forrester, J. W. (2007). System Dynamics – The Next Fifty Years. System Dynamics Review Vol. 23, No. 2/3, (Summer/Fall 2007): 359–370
5. Felten, P. (2008, Dec. 31). Visual Literacy. ChangeMag.org. Retrieved from http://0-web.ebscohost.com.umiss.lib.olemiss.edu/ehost/pdfviewer/pdfviewer?sid=cbc2ab95-7e83-4b47-84ab-9ece85c286b5%40sessionmgr112&vid=2&hid=126
6. Friendly, M. (2009, Aug. 24). Milestones in the history of thematic cartography, statistical graphics, and data visualization. Retrieved from http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf
7. Gigerenzer, G. and Hoffrage, U. (1995) How to Improve Bayesian Reasoning Without Instruction: Frequency Formats. Psychological Review, 102 (4), 1995, 684–704.
8. Herman, I., Melancon, G., and Marshall, M.S.(2000, Jan.). Graph visualization and navigation in information visualization: A survey. IEEE Transactions on Visualization and Computer Graphics. 6, 1. 24-43. Retrieved from doi: 10.1109/2945.841119
   URL: http://0-ieeexplore.ieee.org.umiss.lib.olemiss.edu/stamp/stamp.jsp?tp=&arnumber=841119&isnumber=18188
9. Kahneman and Tversky (1974, Sep. 27). Judgment under Uncertainty: Heuristics and Biases. Science, New Series, Vol. 185, No. 4157 (Sep. 27, 1974), pp. 1124-1131.

10. Kahneman and Tversky (1979, mar.). Prospect Theory: An Analysis of Decision under Risk. Econometrica, Vol. 47, No. 2, pp. 263-292.
11. Kahneman and Tversky (1986). Rational Choice and the Framing of Decisions. The Journal of Business, Vol. 59, No. 4, Part 2: The Behavioral Foundations of Economic Theory (Oct., 1986), pp. S251-S278.
12. Kelton, A. S., Pennington, R. R. & Tuttle, B. M. (2010). The effects of information presentation format on judgment and decision making: a review of the information systems research. Journal of Information Systems, Vol. 24, No. 2.
13. Lohr, S. (2012) Sure, Big Data Is Great. But So Is Intuition. New York Times (Dec. 29).
14. Lukas (1981, Jul.) An experimental investigation of the use of computer-based graphics in decision making. Management Science, Vol. 27. No, 7.
15. Messaoud, R. B., Boussaid, O., and Rabaséda. S. L. and Missaoui, R. (2006). Enhanced Mining of Association Rules from Data Cubes. ACM. DOLAP'06, November 10, 2006, Arlington, Virginia, USA.
16. Meyer, R. D. and Cook, D. (2000). Visualization of Data. Current Opinion in Biotechnology. 11, 1, 89-96.
17. Savitz, E. (2012, Oct. 23). Gartner: Top 10 Strategic Technology Trends For 2013. Forbes. Retrieved from http://www.forbes.com/sites/ericsavitz/2012/10/23/gartner-top-10-strategic-technology-trends-for-2013/
18. Scaife, M., and Rogers, Y. (1996). External cognition: how do graphical representations work? Int. J. Human – Computer Studies (1996) 45, 185 – 213.
19. Shneiderman, B. and Maes, P. (1997, Nov.). Direct manipulation vs. interface agents. Interactions. (ACM), 4(6), pp. 42-61.
20. Usery, E. L. (2003). Multidimensional representation of geographic features. Technical report, U.S Geological Survey (USGS).
21. Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. Decision Sciences 22: 219–240.
22. Zachry, M., & Thralls, C. (2004). An Interview with Edward R. Tufte. Technical Communication Quarterly, 13(4), 447-462.