Composite Indicators Visualization: Exploration of Multivariate Temporal Changes using Radial Visualization

Yael Albo
Faculty of Management
University of Haifa

Joel Lanir
Information Systems department
University of Haifa

Peter Bak
IBM Research Haifa Lab, Israel

Sheizaf Rafaeli
Faculty of Management
University of Haifa

Abstract

We present a work in progress on visualization of a popular Measurement and Benchmarking method, namely Composite Indicators (CIs). Composite indicators are increasingly recognized as an important tool in policy analysis and public communication and are often used to reflect performances of different units (e.g., countries, regions or organizations). Construction of regional Israel ICT Index pointing out digital divide with real data is our research opportunity. Existing CI visualizations introduce a variety of solutions for the complex challenge of presenting multi-dimensional, time-oriented hierarchic data. Radial visualizations are popular among these solutions because of CI's multi-dimensionality, although different coding is in use. In our work, we aim to shed light on radial presentations for multivariate temporal changes, hoping to produce relevant guidelines for CI visualizations. Inspired by existing CI's visualization, in our first controlled experiment we compare the effectiveness of three static techniques: classic radars, flowers glyphs and colored Concentric Radial Space Filling (Circles). Two variables quantities were tested (3 and 7) using lookup and comparison tasks adjusted to digital divide questions. Results show a significant effect of visualization technique on trends and gaps seeking

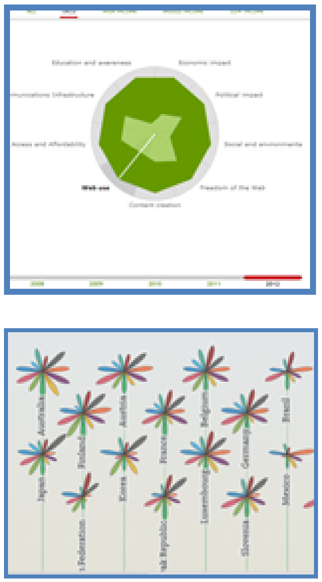

Figure 1: Examples of radial presentations in Composite Indicators' Visualizations. Up – former Webindex by WWW Foundation; Down –OECD BetterLife

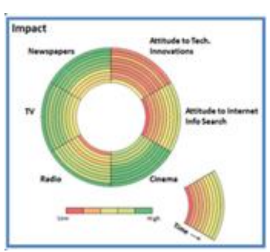

Figure 2: Concentric Radial Space Filling (RSF)

tasks. Subjective preferences show a general advantage of flowers glyphs, while Circles are preferred for trend seeking tasks. In future work, we plan to conduct a second controlled experiment comparing static and dynamic radars and flowers visualization, as well as implement and deploy selected designs in Israel ICT Index visualization.

Introduction

A famous quotation from Lord Kelvin says: "If you cannot measure it, you cannot improve it" [7]. Measuring and benchmarking (M&B) are necessary for understanding an entity's (e.g., a country, a region, a company) position and identifying growth opportunities. M&B is a continuous process of monitoring, benchmarking and improving [9].
A composite index (CI), as an M&B method, is a measure derived from a series of observed facts that can reveal relative positions on a regional basis in a given field (i.e. environment, economy, society or technological development). Multi-dimensional complex concepts are summarized into a single indicator. Sub-indicators construct a CI in a hierarchal way, as a weighted variables tree. When evaluated at regular temporal intervals, an indicator can point out the direction of change in different units and over time [10]. A Composite Indicator might be useful in setting policy priorities and in benchmarking or monitoring growing year after year [3]. Simplicity and easy interpretation is one of the advantages of CI's, as it produces a "bottom line" [12]. On the other hand, there is a need to support deeper insight acquisition (e.g., what is the best practice one should learn from), in order to improve performance [11]. Domain experts, as well as non-experts, might be interested in making deeper sense of CIs. A clear visualization of a composite indicator is highly needed and recommended [10]. Visualization of the benchmarking scores of a variety of measures and aspects should enable the evaluation of the effects of policy-makers' actions, understand where they stand and select improvement measures [9].

We specifically look at a CI called Information and Communication Technology (ICT) index, aiming to measure the "information society" and pointing out digital divides. The questions related to performance indicators might be linked to growth potential, as well as to gap risks: Which indicators and variables should be improved? Does a divide exist? What is the gap's size and trend - is it getting wider or narrower? What is the rate of change? Inspired by existing CI's visualizations (e.g. former WebIndex and OECD Better Life Index) we explore radial solutions for presenting CI's time-oriented data.

1. Characterization of the Visualization Problem

Typical Composite Indicator data consists of 3 main properties. ***Time-oriented***: Measurement and benchmarking is a continuous process. Time-oriented data and tasks are the core of Composite Indicator visualization challenge. ***Multi-dimensionality***: Composite Indicators aim to represent a multi-dimensional reality. Since the phenomenon complexity is simplified by the CI's single value, unfolding dimensions is required for sense-making and taking actions as a result. ***Hierarchy***: "the Variables Tree" structure and the user's need for orientation in it must be considered. Since variables amount is varying for each hierarch level, we decided to explore two amounts: of three and seven variables.

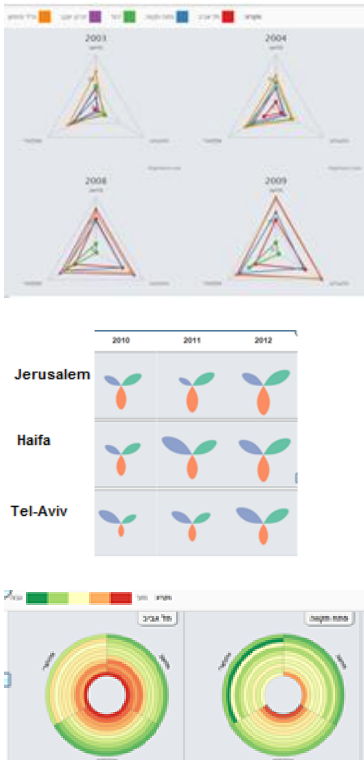

Figure 3: Radial visualizations tested. Radars on top, Flowers glyphs at center, Circles on bottom.

2. Alternative Radial Visualizations

We wish to examine radial time-oriented visualizations as a way to visualize CIs. Radial visualizations are in use in conveying multi-dimensional data [5,1], and are common in CIs' visualizations over the Web (See Figure 1). However, most Web solutions are incomplete for CI exploration purposes (e.g. lack of time dimension visualization). Still, they might be altered to fit our needs. However, radial presentation is in dispute. On one hand, its use can be related to its aesthetic appeal, compact layout, and ability to put selectable data within easy reach of the user [5]. On the other hand, criticism of some radial visualization emerged [6]. We decided to study radial visualization in CI's context, as more effort to study radial visualizations is needed [5]. We chose three comparable technique (see figure 3): classic radar, flowers and Circles. Classic Radar was chosen for its popularity in CI visualizations (e.g. WebIndex) and reports (e.g. [4]). Items (regions) are mapped to color, dimensions are mapped to angle, values are mapped to distance from center and years are represented by small multiples. Flowers glyphs are chosen after the OECD Better Life Index. In this option, dimensions are mapped to color and specific flower leaf; values are presented by the leaf's size. Items and years are presented by small multiples. Items might be moved by interaction. The third option (see also figure 2), inspired by Keim et al.'s Circle View metaphor [8], uses Concentric Radial Space Filling (RSF) presentation. We use the color coding traffic model (green for high values, red for low values) that was shown to be able to simplify M&B complexity [9]. This might provide a knowledge management tool for decision makers. In this RSF presentation, years are mapped to the concentric rings, dimension to angles, values to color and items (regions) to small multiples which might be moved by interaction.

3. Experimentation

We performed a study to test the effectiveness of the three radial visualizations. We used a 3 (Vis: Radar vs. Flowers vs. Circles) x 2 (Variables amount: 3 vs. 7) x 10 (tasks) mixed study design. The Variables amount was between subjects, while the Vis and Tasks were within-subject variables.

3.1 Dataset

We deliberately use real-world data. Our data source is taken from TGI (Target Group Index) marketing and media surveys conducted in Israel during the years 2003-2012. The data contains information from a representative sample of 30 regions and about 10,000 adult Hebrew speakers annually. Using this data we plan to construct in the future a ICT Index. Data was normalized to 1-100 time dependent scores. Each visualization presented data for five items (regions).

3.2 Tasks

For task analysis, we use the formal task model by Andrienko and Andrienko [2]. On the upper level, tasks are divided into elementary and synoptic tasks. Elementary tasks address individual and separated data elements (values or groups of data) and include: lookup, comparison and relation seeking tasks. Synoptic tasks involve a general view (sets of values or groups of data in their entirely), and are divided into descriptive (including: lookup, comparison and relation seeking tasks) and connectional tasks (homogeneous and heterogeneous behavior). We developed 10 tasks, which we found to be relevant to ICT stakeholders. See sample of tasks on Table 1.

Question	***Upper level task category***
What was the Internet users score in Jerusalem in 2010?	Elementary
On 2007, Cellular score in Naharia was greater / equal/ lower than in Haifa	Elementary
Is the divide between Ramat-Gan and Naharia in Internet scores getting wider / narrower?	Synoptic
Internet growth rate in Haifa during 2006-2009 was higher / slower than Cellular growth rate.	Synoptic

Table 1. Task Examples.

3.3 Procedure

One participant was tested at a time, using a full HD monitor. Four tutorial tasks were provided for each visualization (to prevent learning effects), followed by survey questions specific to that visualization. A general survey questions was provided at the end of the session. Each session lasted 30-45 minutes. There were 36 non color-blind subjects (22 male, 14 female) with an average age of 26 (STD 3.75). Participants were asked to work as quickly and accurately as possible. Visualization order was counterbalanced. The order of the tasks was fixed, most difficult tasks at the end. This allowed participants to build their skills as they proceeded.

3.4 Study Results

We present study results in three parts: completion time, accuracy and subjective preferences. A 3 (Vis: Radar vs. Flowers vs. Circles) x 10 (Tasks) x 2 (Vars: 3 vs. 7) Mixed Analysis of Variance (ANOVA) was conducted with Vars being a between-subject variable and Tasks and Vis being within-subject variables.

3.4.1 Task completion time

We first wanted to rule out a learning effect by examining whether the three sessions differ (i.e., whether participants would perform faster in their second and third session). A repeated-measures Analysis of Variance on task completion time did not find a significant difference between the first, second and third session ($p=0.104$).

We then compared results of the three Visualizations. In terms of task completion time we observed a highly significant main effect for Vis, $F(2,33)=23.97$, $p<.001$. The average time when using the radar visualization (*M* = 43.1 sec., *SD* = 2,24) was slower than when using Circles (*M* = 35.2, *SD* = 1.81) and Flowers (*M* = 35.9, *SD* = 1.65). Post-hoc analyses using the Bonferroni adjustment showed that the difference between the radar and the other two visualizations was significant ($p<0.001$). Average completion times per visualizations are presented in Figure 4.

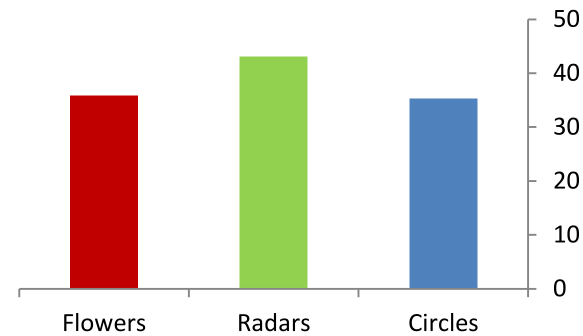

Figure 4: Average time per visualization

Main effect of variables amount was marginally significant, $F(1,34)=3.559$, $p=0.68$. This result is surprising as we expected a significant difference between 3 and 7 variables stemming from the more complex data. Practically this finding might be considered useful since it implies that the same visualization could be used independently on the amount of variables used.

Interaction effect between Vis and Task was also significant, $F(2,18)=4.02$, $p=.003$, indicating that different visualizations were performed faster with different tasks. Each task was tested separately to compare performance of each visualizations per task. Significant differences were found in nine tasks. While a specific per-task analysis is out of the scope of this document, we mention that out of these nine tasks, Radar visualization was slowest and responsible for the difference in 6 out of the nine tasks. Results per task are presented in Figure 5.

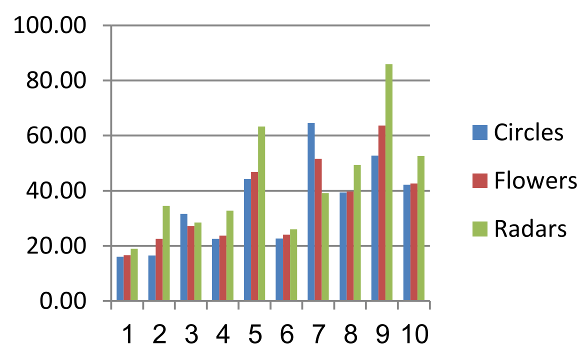

Figure 5: Average time of tasks per visualization

3.4.2 Accuracy

The average percentage of correct answers for a participant is 86.7% for Flowers visualization, 86.1% for Radars visualization and 85.6% for Circles visualization. While the overall differences in error rates were neglectable, an analysis of errors per tasks showed that synoptic tasks showed a higher error rate and that Radar was inferior in 3 specific tasks.

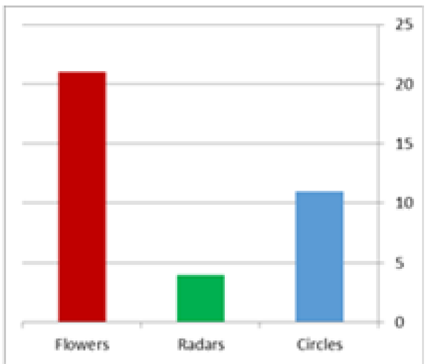

Figure 6: General subjective preferences

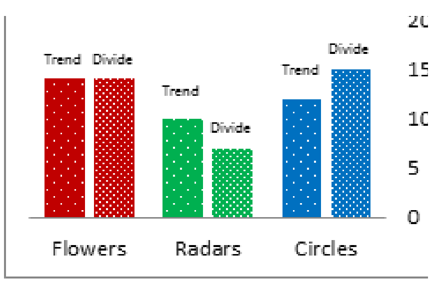

Figure 7: Subjective preferences for trend and divide search tasks

Subjective preferences

After a user completed the set of tasks for each visualization type, one of the three visualization types had to be selected as the preferred one. As can be seen in figure 6, the flowers visualization was selected 21 times out of 36, showing an overall preference for the Flower visualization. Circles was chosen 11 times and Radars visualization only 4 times. Preference by task shows an advantage to the Circles view for specific trend tasks. Subjective preferences changes in favor of circles and radars for specific task preferences, as illustrated in figure 7. These tasks were trend and divide search tasks which are tasks that are specifically important in the CI domain.

Discussion

Three static radial visualization techniques for the display of multivariate time-series were examined by a series of user tests with 36 subjects: Radars, Flowers glyphs and Concentric Radial Space Filling Circles. Performance was measured by task completion time and accuracy. Partial results suggest a significant main effect of visualization on task completion time, as well as significant interaction effect of visualization and task. Surprisingly, variables amount was not proved to be a major effective factor, which practically might mean no visualization change is necessary while presenting different levels of CI hierarchic variables tree. Flowers and Circles are fast for certain tasks, radars are slowest for all tasks. Flowers are generally preferred. Circles are preferred for divide search task, while radars found to be least preferred, mainly because of cluttering problems. Radars cluttering could be solved by a suitable interaction.

Future Plans

Analyzing process of this experiment's results will be continued. We plan to analyze factors' effects separately by task and by task aggregation. Next we plan to conduct a second study, aiming to compare Dynamic vs. Static conditions, both by interactive Radar and Flowers visualization. Based on these results, our ultimate goal is to develop an interactive visualization component for the Israel ICT Index and implement it among "real-world" stakeholders.

Acknowledgements

This work is done as part of the National Israel ICT project by The Center of Internet Research supported by the Israel Internet Association-ISOC-IK; Bezeq; Appleseeds Academy; LINKS Learning in a Networked Society.

References

[1] ANDRIENKO G., ANDRIENKO N., BAK P., KEIM D., WROBEL S.: Visual analytics of movement. Springer Publishing Company, Incorporated, 2013.

[2] ANDRIENKO G., ANDRIENKO N., *Exploratory analysis of spatial and temporal data*. Berlin, Germany: Springer, 2006.

[3] BANDURA R.: A survey of composite indices measuring country performance: 2008 update. Office of Development Studies, New York: United Nations Development Program (2008).

[4] BILBAO-OSORIO B., DUTTA S., LANVIN B.: The global information technology report 2013. Growth and Jobs in a Hyperconnected World (2013).

[5] DRAPER G., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. Visualization and Computer Graphics, IEEE Transactions on 15, 5 (2009), 759–776.

[6] Feldman, Roger. "Filled Radar Charts Should not be Used to Compare Social Indicators." *Social indicators research* 111.3 (2013): 709-712.

[7] KELVIN W. T. B.: Electrical units of measurement. Institution, 1883.

[8] KEIM D. A., SCHNEIDEWIND J., SIPS M.: Circleview: a new approach for visualizing time-related multidimensional data sets. In Proceedings of the working conference on Advanced visual interfaces (2004), ACM, pp. 179–182.

[9] MAHESHWARI D., JANSSEN M.: Measurement and benchmarking foundations: Providing support to organizations in their development and growth using dashboards. Government Information Quarterly 30 (2013), S83–S93.

[10] NARDO M., SAISANA M., SALTELLI A., TARANTOLA S., HOFFMAN A., GIOVANNINI E.: Handbook on constructing composite indicators: methodology and user guide. Tech. rep., OECD publishing, 2005.

[11] PETROVI´C M., BOJKOVI´C N., ANI´C I., PETROVI´C D.: Benchmarking the digital divide using a multi-level outranking framework: Evidence from ebrd countries of operation. Government Information Quarterly 29, 4 (2012), 597–607.

[12] SAISANA M., TARANTOLA S.: State-of-the-art report on current methodologies and practices for composite indicator development. European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen.