Who, Where, When and with Whom? Evaluation of Group Meeting Visualizations

Conference Paper · August 2016

DOI: 10.1007/978-3-319-42333-3_19

4 authors:

Simone Kriglstein

Johanna Haider
TU Wien

Günter Wallner
University of Applied Arts Vienna

Margit Pohl
TU Wien

Who, Where, When and with Whom? Evaluation of Group Meeting Visualizations

Leave Authors Anonymous1, Leave Authors Anonymous1, Leave Authors Anonymous2, and Leave Authors Anonymous1

[1] Anonymous for Submission
City, Country
{anonymous,anonymous,anonymous}@anonymous.com
[2] Anonymous for Submission
City, Country
anonymous@anonymous.com

Abstract. Visualizing time-dependent and location-based data is a challenging problem but highly relevant for areas like intelligence analysis, traffic control, or social network analysis. In this context, we address the problem of visualizing meetings between persons, groups of persons, vehicles, or other entities. However, the temporal dimension inherent in such data makes traditional map representations less well suited for this kind of problem as they easily become cluttered. To overcome this issue we developed a modified map representation and three alternative representations (two matrix-based visualizations and one based on Gantt charts). An empirical evaluation comparing these four visualizations and assessing correctness, recognition rates of groups, and subjective preference indicates that the alternative visualizations perform significantly better than the map-based representation when meetings need to be identified. In addition, we identify specific strengths and weaknesses of the investigated visualizations and propose design considerations.

Keywords: Information Visualization; Map; Matrix; Gantt charts

1 Introduction

The analysis of spatio-temporal data to identify patterns of movements (e.g., movements of persons or groups of persons, vehicles, etc.) is of great interest in several domains such as location-based social networks analysis (e.g., [5,15]), crime analysis (e.g., [4]), or movement pattern analysis (e.g., [3]). However, analysts and decision-makers are often confronted with an enormous amount of multidimensional information that can be extracted from spatio-temporal data. In order to make these large amounts of data manageable there is a need to develop effective visualization methods. Especially in case of the analysis of group meetings not only questions relating to *when* and *where* are of interest but also more complex questions such as *who* meets with *whom*. For representing spatial data, map representations are very popular to show different types of geographic information in an intuitive way. Such maps can support analysts to deduce associations and connections from the spatio-temporal data (e.g., population density versus recreation areas) [17]. However, the temporal dimension is problematic with map representations as a single map usually only represents a single slice in time. Yet, this temporal aspect is vital for understanding meeting patterns of various people.

Addressing this challenge, we designed alternative visualizations in the course of two projects to facilitate the analysis of location data of a group of individuals. It was important for us to find a possibility to answer *where*, *when*, *who* and with *whom* questions with a map representation but we also investigated alternative representations: two matrix-based visualizations and one adapted from Gantt charts. Matrix visualizations simultaneously show the relationships between multiple variables while at the same time being free from occlusions [6] whereas Gantt charts are usually well suited to show activities displayed against time. Therefore, we believe that these representations could be suitable alternatives to the traditional map representation. To assess and compare these representations – henceforth referred to as *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* – with respect to correctness, recognition rates of groups, individual preference, and their utility for identifying and interpreting meetings of persons we conducted an empirical evaluation with 24 subjects. Our results show that the three alternative visualizations (*Augmented Matrix*, *Matrix*, *Gantt*) perform better than the *Map* representation and have also been preferred by the participants for identifying meetings. However, although *Augmented Matrix*, *Matrix*, and *Gantt* have their strengths, the results also indicate certain weaknesses. Based on the results, we derive design considerations for designing time-dependent location-based visualizations. The goal of this study is not to evaluate a specific system, but to provide general information on what kind of visualization is appropriate for identifying meetings between entities.

2 Related Work

The map is the predominant visualization for spatial data, as the timeline is for temporal data. Nevertheless, the visualization of spatio-temporal data is a challenging problem [8] since a map easily gets cluttered when several points in time are represented on a single map. To compare multiple time slices several maps can be displayed as small multiples, time can be represented by adding a third dimension, or animations can be used to visualize changes between the different slices in time. However, these solutions can be suboptimal. For example, while small multiples offer the analyst multiple time slices at once the number of maps which can be displayed simultaneously is limited by the available screen space. Moreover, small multiplies can make it difficult to understand how a map evolves over time [7]. Although these issues can be addressed by employing animation, evaluation studies have shown that animation can lead to confusion when too many data points move simultaneously which, in turn, can cause analysts to miss relevant information (cf. [13]). Many geo-visualization tools (e.g., [12]) use the concept of *Space Time Cubes* [2] to visualize time-dependent movement data inside a cube and where the height axis is used to represent time. However, in the visualization research community there is some discussion about the possible implications of 3D representations (see, e.g., [1]). Especially if spatio-temporal data needs to be analyzed with maps it seems that 2D visualizations are less error-prone for simple tasks (cf. [11,14]) compared to 3D representations.

There exist only a few approaches for the representation of movement data where the map is not the main visualization. Two such approaches are described below. The first, concerned with visualizing movement data of players of mobile outdoor games and proposed by Orellana et al. [18], shares some similarities with the Gantt charts we used in our study. They especially emphasize the interaction between different players such as meetings of two or more players. The second, proposed by Shen and Ma [20], augments adjacency matrices with a path visualization to facilitate path finding. In our study, we use Gantt charts as well as matrices as possibilities to represent meetings. Gantt charts are commonly used to support planning activities [1] and are especially suited for the representation of the interactions between different processes. This is a problem which can be compared to the meeting of several people at the same point in time. In general, Gantt charts are appropriate and usable for the representation of events in time [16]. Matrix-based representations are also a possibility to show meetings of several persons. Ghoniem et al. [6] compared matrix-based visualizations to a node-link visualization and found out that matrices are preferable when graphs become bigger than 20 nodes. On the other hand, matrices make it more difficult to detect paths, an issue addressed by Shen and Ma [20] (see above) whose approach influenced our *Augmented Matrix* visualization. Kessell and Tversky [10] conducted a study which has some similarities to our approach. They also address the issue of the combination of space, time and persons/objects. They argue that matrices and line visualizations are highly appropriate for this kind of tasks. Their results are that in general matrices are preferable to line visualizations. However, there are some significant differences between their study and our approach in that our tasks are much more exploratory in nature and to support that we used more complex data sets. Nevertheless, there are some similarities between the results of the two studies and the visualizations used.

To sum up, existing literature on visualizing time-oriented data indicates that 3D, animation, and small multiples can be error-prone and difficult to analyze. Possible alternatives are 2D visualizations based on a timeline metaphor as described by Kessell and Tversky [10].

3 Time-Dependent Location-Based Visualization

For our study we have chosen four visualizations – *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* (see Figure 1) – that use different characteristics to encode two major attributes of time-dependent location data: a) the position of a person and b) the point in time associated with this position. In the following we provide a short description of these visualizations:

Map. The presence of people at a certain location is marked by a circle located at the $x-$ and $y-$coordinates associated with the location on a two-dimensional map. Each person is represented with a unique color. If a location is visited multiple times or by different persons then the circle is split into evenly sized sectors, with each sector representing one visit of a person and being colored based on the color associated with that person. In other words, if a person visits the same location multiple times then that person also occupies as many sectors. The number of visits to a location is also reflected by the size

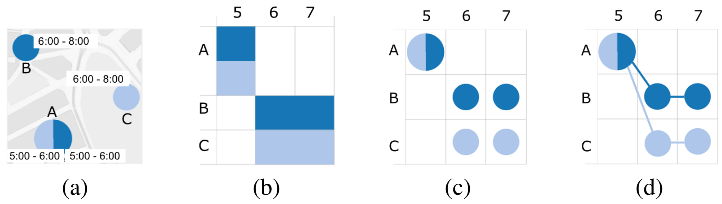

Fig. 1. Comparison between the tested visualizations *Map* (a), *Gantt* (b), *Matrix* (c), and *Augmented Matrix* (d). *A*, *B*, and *C* denote locations and 5, 6, and 7 denote hourly time intervals. Different persons are represented by different colors: ■ Person 1 ■ Person 2

of the circle. Next to each sector the time period of the stay is depicted. In addition, the name of the location is displayed next to each circle. For example, in the map shown in Figure 1(a) Person 1 was at location *B* from 6:00 to 8:00 and at location *A* from 5:00 to 6:00, at the same time frame when Person 2 was there.

Gantt. Gantt charts [22] help to convey schedules by illustrating the start and finish dates of tasks of a project. We adapted this method for the representation of time-dependent location data. The x-axis represents one-hour intervals and locations are listed alphabetically along the y-axis. Each person is represented by a colored bar and the position and length of the bar reflects the time of arrival and departure, i.e., the duration of stay of a person at a location. For example, in Figure 1(b) Person 1 is at location *B* between 6:00 and 8:00 and Person 2 is at location *C* between 6:00 and 8:00. Persons in the same location are drawn beneath each other (cf. location *A* from 5:00 to 6:00 in Figure 1(b)).

Matrix. Matrix visualizations [23] are a further promising visualization to depict the presence and strengths of relationships between different variables in a compact way. In our case each row corresponds to a single location and each column represents a one-hour interval. The presence of a person at a specific location i during a specific hour j is marked by placing a circle at cell (i, j). If multiple persons are present at the same location during the same interval then the circle is divided into equally sized parts. For example, in Figure 1(c) Person 2 is at location *C* from 6:00 until 8:00 whereas both, Person 1 and Person 2, have been at location *A* from 5:00 to 6:00.

Augmented Matrix. Matrix visualizations have the advantage to be free of visual clutter but are not well suited for path finding (cf. [6]). To overcome this limitation Shen and Ma [20] augmented adjacency matrices with a path visualization to facilitate the easy tracing of paths. Influenced by their approach we extended the above described matrix representation with lines to explicitly show the movements of persons between different locations (see Figure 1(d)). These lines are color-coded to show which person has changed the location.

4 Empirical Evaluation

The goal of our study was (a) to clarify the advantages and disadvantages of *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* for the purpose of identifying meetings of persons and (b) to assess correctness and subjective preference. For this purpose, the study aimed to address the following research questions:

RQ1 - Correctness: Which visualization is interpreted more correctly by the participants?
RQ2 - Recognition Rate: Have the properties of the groups (specifically, the number of people in a group, number of meetings, and the total amount of time a group spent together) an influence on the recognition rate?
RQ3 - Preference: Do participants prefer the *Map*, *Gantt*, *Matrix*, or *Augmented Matrix*?
RQ4 - Saliency: Which groups are perceived as salient by the participants? Are there differences between *Map*, *Gantt*, *Matrix*, and *Augmented Matrix*?

At this point we should also note that the emphasis of this study is to test possible visualizations independently of interactions. Testing visualizations and interactions in conjunction would confuse the relationship between independent and dependent variables. It would not be possible to clearly assess whether it is the visualization or the interaction which produces positive or negative effects. We intend to study interactions influencing the perception and interpretations of meetings in the future. Another possibility to improve such visualizations would be to compute meetings algorithmically and provide a list of these meetings to the users. A problem in this context is that if there are many meetings the list will be quite confusing. Nevertheless, it could be a valuable addition to such visualizations to provide such a list. Again, we did not add this feature because we did not want to confound the results.

Tasks. For our study, different types of tasks were designed to identify strengths and weaknesses of the different visualizations. These tasks should reflect common questions that can occur during the analysis of groups and their movement patterns (e.g., how long does a group stay at the same location; do groups repeatedly change locations to meet). We generalized these tasks in a way that they are also applicable in and relevant to other domains. In summary, the following three tasks were developed to investigate our research questions:

Task 1 - Duration: Participants were asked to identify groups that met in the same constellation[3] for more than one hour at one location.
Task 2 - Location Changes: Participants had to list groups that met in the same constellation[3] at different locations for at least one hour each.
Task 3 - Salient Groups: Participants were instructed to list up to three groups which they perceived as striking in some sort of way and to provide an explanation why they thought these groups were special.

[3] That is, groups that consist of exactly the same persons. If a person joins or leaves a group then it is considered a different group.

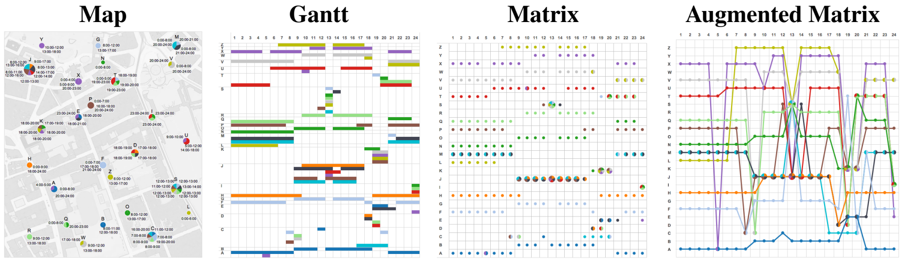

Fig. 2. One of the four test cases which were used in the study (Test Case 1).

Test Cases. The dataset we used consisted of four days and contained visits to 26 different locations (labeled *A-Z*) from 12 people (color-coded). In addition, we ensured that each person is at least once at the same time at the same location as another person. The colors for encoding the different persons were selected based on a color-scheme for qualitative data from ColorBrewer [9]. For each day a *Map*, a *Gantt*, a *Matrix*, and an *Augmented Matrix* visualization was created by using Tableau[4] (see Figure 2). Based on these four days we built four different test cases: Test Case 1 and Test Case 2 consisted of the visualizations of a single day, whereas Test Case 3 and Test Case 4 were composed of the visualizations of two consecutive days displayed next to each other. The number of correct groups to identify varied for the different test cases and between Task 1 and Task 2, resulting in a total of 60 groups. In addition, the number of people (from 2 to 5 people, $M = 2.5$, $SD = 0.74$), the number of meetings (from 1 to 7, $M = 1.83$, $SD = 0.94$), and the duration of the meetings (ranging from 1 to 8 hours, $M = 2.2$, $SD = 1.55$) varied from group to group.

Procedure. To address our research questions, we decided to use an online questionnaire which was created using LimeSurvey[5]. We decided on a survey in order to reach a broader audience since this way the participants could partake in the study from their home anytime without time pressure. The questionnaire included a) closed questions using, for example, checkboxes, b) open-ended questions to offer the participants the possibility to explain their decisions, and c) rank-ordered questions to rank their preferences. The questionnaire consisted of general questions, questions concerning the presented visualizations and questions about the participants' preferences. We also conducted a pre-test with five participants to ensure that questions were understandable and tasks could be completed within a reasonable time. As participants reported loss of concentration and focus during the pre-test we split the survey into two parts of one hour each (consisting of Test Case 1 and 2 and Test Case 3 and 4, respectively), with the second part being conducted one week after the first part.

Sample. The invitation to participate in the online survey was sent via email to Computer Science students who were considered to have at least basic knowledge with visualizations. In total, 24 participants between 23 and 40 years ($M = 27.6, SD = 3.6$)

[4] `http://www.tableau.com/` (Accessed: January, 2016)
[5] `http://www.limesurvey.org` (Accessed: January, 2016)

responded to the survey. One third of the participants was female and 16 participants were male. More than half of the respondents considered themselves to be highly ($n = 3$) or very familiar ($n = 11$) with visualizations, seven claimed to be moderately familiar, and only three said to be slightly or not at all familiar. None of the participants reported color blindness. All 24 participants completed the first part of the survey and 18 of them also completed the second part.

5 Results

The following results are based on quantitative analysis and a qualitative content analysis [19] which was applied to the responses to the open-ended questions.

5.1 RQ1 - Correctness

Correctness was measured by the number of correctly identified group meetings. To analyze the effect of the factors VISUALIZATION (V), TEST CASE (TC), and TASK (T) on the number of correctly identified groups in Task 1 and Task 2, a $4 \times 4 \times 2$ ($V \times TC \times T$) repeated measures ANOVA was conducted. Participants who did not fill out the second part of the survey were excluded from this analysis, yielding a sample size of 18 participants. Furthermore, as the number of correct groups differed from test case to test case and from task to task the number of correctly identified groups was normalized to the range $[0..1]$ by dividing the individual values by the total number of groups for the respective test case and task, yielding the fraction of correctly identified groups. Please note, that the values reported in the following are thus fractions and not absolute values.

The results show a significant main effect of VISUALIZATION ($F(3,51) = 44.64, p < .001$). Post hoc tests using Bonferroni correction revealed that the participants performed significantly poorer with the *Map* ($M = 0.303, SE = 0.038$) representation than with the *Gantt* ($M = 0.540, SE = 0.026$), *Matrix* ($M = 0.585, SE = 0.034$), and *Augmented Matrix* ($M = 0.605, SE = 0.031$) representations ($p < .001$ in each case). The main effect of TEST CASE was also significant ($F(1.99,33.79) = 25.74, p < .001$, Greenhouse-Geisser adjusted). Participants performed best on *TC 1* ($M = 0.589, SE = 0.026$), followed by *TC 2* ($M = 0.54, SE = 0.027$) and *TC 3* ($M = 0.491, SE = 0.032$) and lastly *TC 4* ($M = 0.411, SE = 0.035$), that is participants generally performed better on the smaller datasets than on the larger datasets. The analysis also revealed a significant main effect of TASK ($F(1,17) = 116.54, p < .001$), indicating that participants performed considerably better on Task 1 ($M = 0.691, SE = 0.027$) than on Task 2 ($M = 0.325, SE = 0.036$).

Among the interactions effects only the interaction between VISUALIZATION and TASK was significant ($F(3,51) = 6.114, p = .001$). In case of Task 1 participants performed the best with the *Augmented Matrix* ($M = 0.788, SE = 0.032$) and the *Matrix* ($M = 0.783, SE = 0.035$) representation, followed by the *Gantt* ($M = 0.762, SE = 0.029$) chart and lastly the *Map* ($M = 0.431, SE = 0.049$) visualization. Examining the fraction of correctly identified groups by visualization for Task 2 shows an identical order with *Augmented Matrix* ($M = 0.422, SE = 0.044$) being best, followed by *Matrix*

Table 1. Summary table of the repeated measure ANOVA results (correctness).

Source	SS	df	MS	F	p
V	8.396	3	2.799	44.640	$<$ **.001**
TC[†]	2.480	1.988	1.248	25.739	$<$ **.001**
T	19.322	1	19.322	116.536	$<$ **.001**
TC $\times$ T	0.116	3	0.039	1.527	.219
TC $\times$ V	0.229	9	0.025	1.071	.388
T $\times$ V	0.691	3	0.230	6.114	**.001**
TC $\times$ T $\times$ V	0.154	9	0.017	0.894	.533

V (Visualization), TC (Test Case), T (Task), SS (Sum of Squares), MS (Mean Square)
significant p-values are highlighted in bold, [†]Greenhouse-Geisser adjusted

($M = 0.386, SE = 0.044$), *Gantt* ($M = 0.317, SE = 0.042$), and *Map* ($M = 0.174, SE = 0.037$). However, the differences between the representations are slightly more pronounced and the average number of correctly identified groups is generally lower for each visualization than in Task 1.

Decomposing the interaction effect by VISUALIZATION with a simple effects test confirmed that the differences between Task 1 and Task 2 are statistically significant for each visualization ($p < .001$ in each case). Decomposing the interaction effect by TASK showed that participants performed significantly worse with the *Map* representation than with the other three visualization for both tasks ($p < .001$ for each comparison). However, in case of Task 2 the difference between the *Augmented Matrix* and *Gantt* visualization was also significant ($p = .02$), indicating an advantage of the *Augmented Matrix* for identifying repeated meetings of the same group compared to the *Gantt* chart. All other two-way interactions and the three-way interaction were not significant. Table 1 summarizes the results of the repeated measures ANOVA.

5.2 RQ2 - Recognition Rate

Next, we assessed if the properties of the groups, that is, number of people in a group, number of meetings per group, and whether the total amount of time a group spent together had an influence on the recognition rate – the proportion of people who correctly reported the group – of the group. To evaluate the relationships between the above mentioned properties and the recognition rate – and due to non-normality of the data – a Spearman rank correlation has been conducted separately for each visualization and task (see Table 2).

In case of Task 1 the results revealed moderate to strong, statistically significant, correlations between the total amount of time a group spent together and the recognition rate for each of the four visualizations. In case of the *Matrix*, *Augmented Matrix*, and *Gantt* visualizations this can be explained by the fact that groups that meet longer are visually more evident (e.g., pie charts appear consecutively, bars are longer) and thus easier to spot. In addition, the number of meetings correlated positively with groups being recognized for all but the *Gantt* representation. Given that Task 1 asked for identifying groups who meet at least for two hours at any location, multiple meetings

Table 2. Spearman rank correlations between recognition rate and total amount of time a group spent together (Δt), number of people in a group (N_p), and number of meetings (N_m) for Task 1 and Task 2.

	Task 1			Task 2		
Recognition rate	Δt	N_p	N_m	Δt	N_p	N_m
Matrix	$.663^{**}$	$-.123$	$.657^{**}$	$.447^{*}$	$-.276$	$-.076$
Augmented Matrix	$.792^{**}$	$-.022$	$.733^{**}$	$.663^{**}$	$-.350$	$.082$
Gantt	$.497^{**}$	$-.137$	$.292$	$.329$	$-.262$	$-.047$
Map	$.631^{**}$	$-.185$	$.503^{**}$	$.402^{*}$	$-.054$	$-.012$

* significant at $p < .05$, ** significant at $p < .001$

increased the chance for recognizing a group. Task 2 showed similar positive correlations among duration and recognition rate for all representations, except for the *Gantt* chart. This was, in some way, unexpected as longer bars should also be beneficial for detecting groups who meet at different locations. In addition, the number of meetings did not positively influence the recognition of groups anymore. This, however, had to be expected since almost all of the groups contained in the datasets for Task 2 met at maximum two times. The number of people a group consists of, on the other hand, had no influence in neither of the visualizations and tasks, most likely because the number of persons was quite limited in most of the cases (in 53 out of 60 cases, the groups consisted of a maximum of three people).

To further investigate the correlation among duration and recognition rate which was not significant in the case of the *Gantt* chart we examined the recognition of individual groups more closely and could observe that groups with a white space between people of the group were seemingly more difficult to recognize than groups without white space. This observation was confirmed by a paired sampled t-test ($N = 18$) comparing the average recognition rate of groups with and without white space, $t(17) = -8.792, p < .001$. Groups without white space were on average easier to identify ($M = 0.61, SD = 0.11$) than those containing white space in at least a single meeting ($M = 0.35, SD = 0.16$). This is in line with the Gestalt law of proximity that *things that are close together are perceptually grouped together* [21, p.189]. In case of the *Gantt* chart the white space sometimes suggested that people belong to different groups although actually belonging to the same (see Figure 3, left).

Comparing the recognition rate of individual groups between the *Matrix* and *Augmented Matrix* visualization, the auxiliary lines of the *Augmented Matrix* improved the recognition of 50.0% of the groups but also impeded the identification in 38.3% of the cases (no change in the remaining 11.7%). Although these differences were not statistically significant at the .05 level except in a single case (assessed via a McNemar test separately for each group) there is a tendency that meetings are missed if there are too many lines in its proximity, or to put it differently, if the lines increase the visual clutter in the area surrounding the meeting. However, lines also improved the discovery of groups, usually in cases where the lines were more distinguishable or where the convergence of lines was more readily apparent. By way of example, Figure 3 (right) shows

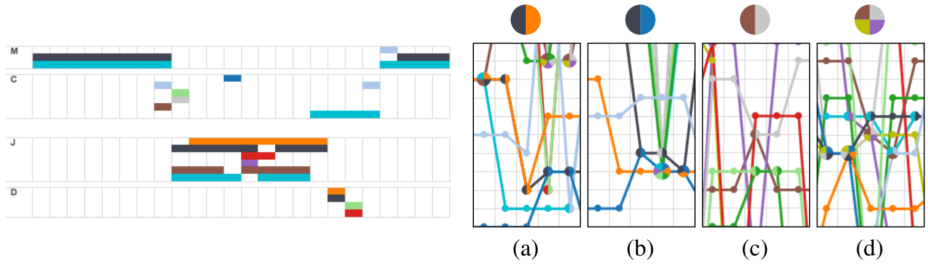

Fig. 3. Left: White space between people of the same group hindered group identification in the *Gantt* representation. For example, group meeting at *M* and *C* (top) was only recognized by 18% of the participants in Task 2 as opposed to group meeting at *J* and *D* (bottom) which was identified by 47%. Both groups met twice for one hour each at different locations. Right: Examples of cases where auxiliary lines improved (a, b) or impaired recognition of groups (c, d). The group under consideration is depicted above each example.

some groups where the auxiliary lines improved or reduced the recognition rate by at least 0.2.

5.3 RQ3 - Preference

Preference refers to the subjective preference ratings of the participants. Analysis of the preference ranking ($1 = $ best, $4 = $ worst) of the visualizations after the small datasets (that is, Test Case 1 and 2) using a Friedman test ($N = 24$) showed a significant overall difference among the rankings ($\chi^2(3) = 41.15, p < .001$). The *Matrix* visualization ranked best with a mean rank of 1.5, followed by the *Augmented Matrix* (2.08), *Gantt* (2.63), and lastly the *Map* (3.79) representation. Post hoc analysis with Wilcoxon signed-rank tests with a Bonferroni correction applied showed significant differences between each pair of visualizations except between the *Augmented Matrix* and *Gantt* representation ($Z = -1.626, p = .104$). Analysis of the ranking after the large datasets (Test Case 3 and 4) gives almost identical results. Again, the Friedman test ($N = 18$) was significant ($\chi^2(3) = 43.13, p < .001$) with mean ranks of 1.22 (*Matrix*), 2.06 (*Augmented Matrix*), 2.78 (*Gantt*), and 3.94 (*Map*). However, in contrast to above, the difference between *Augmented Matrix* and *Gantt* representation was also significant ($Z = -2.422, p = .015$).

From the explanations provided by the participants it was apparent that the *Matrix* visualization was mostly appreciated for providing a structured way to solve Task 1 and Task 2, with participants describing the visualization as clear (6/6)[6], well readable (4/3), and mentioning that it offers a good overview (4/5). However, four participants noted that location changes are hard to see (4/0). We got contradictory responses to the *Augmented Matrix* representation: some participants noted that the lines are distracting (2/2) and clutter the visualization (7/7) while other participants found the lines useful to see relations and track changes (9/5). This ambivalence is also reflected in the impact of the auxiliary lines on the recognition rate of groups (see RQ2 above). With respect

[6] number of statements after the small/large test cases

Table 3. Categories developed from the qualitative content analysis along with the number of statements falling within each category.

Visualization	Duration	Group Size	Repetition	Color	Position	Proximity	Σ
Augm. Matrix	65	42	41	35	24	5	212
Matrix	49	28	27	33	19	5	161
Gantt	40	28	20	19	18	2	127
Map	11	39	17	16	11	3	97
Σ	165	137	105	103	72	15	597

to the *Gantt* representation most participants indicated that the locations were hard to discern due to the thin separating lines between locations (4/8) and that the irregular heights of the rows were irritating (6/3). On the other hand, some participants liked the *Gantt* chart because of its readability (5/0), although this quality was not mentioned explicitly after the large test cases anymore. A minority of the participants also ranked the *Gantt* higher because they did not like the pie charts in the other visualizations (2/1).

The poor result of the *Map* can be mostly attributed to the fact that participants found it difficult to infer actual meetings of groups as the meeting times were not readily apparent but rather had to be calculated from the labels (13/10) which, subsequently, also complicated the identification of recurring group meetings. For example, one participant criticized that *"you have to read out the times, as opposed to the other types where you can derive the times directly from the axis"* while another found that *"the need to compare all the labels makes it hard to figure out if people really meet or if they are just at the same location at different times"*. However, a few participants, though admitting that the huge amount of labels made it difficult to solve Task 1 and Task 2, liked the geospatial representation (1/3).

5.4 RQ4 - Saliency

In Task 3 participants were asked to identify any conspicuous groups (saliency). A qualitative content analysis [19] of the participants' responses to Task 3 was conducted by three researchers in an iterative process. In a first round keywords were extracted from the comments which were then analyzed to form categories. Based on the keywords the statements were then assigned to the corresponding categories. Please note, that this means that a comment could be assigned to more than one category. Finally, the comments within each category were counted. In the end, the comments provided by the participants were categorized into six categories. Table 3 lists the number of statements – summarized across all four test cases – falling within each category grouped by visualization. In the following we will discuss each category in more detail.

Duration. Most descriptions of salient groups with the *Augmented Matrix*, *Matrix*, and *Gantt* representation contained statements about the time a group spent together (approx. 1/3 of all statements made with each visualization). Longer meetings were more often highlighted with these three visualizations in comparison to short meetings, more specifically (from in total 165 statements) participants made 135 statements highlighting the long duration of a meeting as opposed to just 23 statements being concerned with short meetings. This, however, may well be because longer meetings were easier to recognize than shorter ones (see RQ2). Worth noting, but not surprising given that participants found it very troublesome to calculate the time-spans of meetings from the labels (cf. RQ3), is the comparable low number of statements about the duration of group meetings with the *Map* representation.

Group Size. In total terms, group size was the second most mentioned feature for considering a group as salient. Most noticeable is the large fraction of statements (approx. 38.6%) concerning group size with the *Map* representation: as the exact durations of meetings were hard to infer participants were seemingly more focused on this aspect. Summarized across all visualizations, larger groups were reported much more often (119 statements) than smaller groups (6 statements) whereas participants usually referred to groups with more than three people as large and groups consisting of two people as small.

Repetition. Most statements (approx. 39%) referring to repeated meetings of a group were made with the *Augmented Matrix* visualization, most likely because the auxiliary lines made it easier to trace the route of the group's members – an impression which is also reflected in some of the comments made by the participants who found the lines helpful to track changes (cf. RQ2). In 32 cases participants stressed that the repetitive meetings of a group are interesting because they met at different locations, while groups meeting at the same place have only been emphasized in 15 statements. In the other cases, participants did not specify why they considered the repetition as salient. Again, most statements highlighting different locations (15 out of 32) have been made with the *Augmented Matrix* visualization.

Color. Around 14% to 20% of the statements per visualization contained some sort of reference to the colors used to depict the individual persons. Participants, for example, emphasized groups because of the color combination in general (28 statements) or because of the contrast between the colors (29 statements).

Position. The position of the groups in the visualization has also been a contributing factor if groups were perceived as salient or not, with around 10.9% to 13.6% of statements per visualization explicitly referring to the position as reason for being salient. This was particularly an issue for the *Augmented Matrix*, *Matrix*, and *Gantt* representation since especially groups which were located in the center of the visualization (29 statements) or on the left side (11 statements) were considered to stand out.

Proximity. Proximity to other groups played a minor role for participants to consider a group as salient. In case of the *Augmented Matrix*, *Matrix*, and *Gantt* representation this is in line with our expectations as the order of the locations in these visualizations do not reflect the geographical distance. However, in case of the *Map* visualization this low number is a bit surprising but is most likely attributable to the nearly equidistant spacing between the locations.

6 Discussion

In the following we discuss our results according to the four visualizations, tasks, and test cases:

Visualizations. The results show that, in general, the three alternative visualizations (*Augmented Matrix*, *Matrix*, *Gantt*) are better suited for identifying meetings than the *Map* representation (Task 1 and Task 2). Furthermore, the *Map* was clearly the least preferred visualization. Based on the literature review we expected this result, since maps seem not to be an appropriate visualization to present several points in time. More detailed analysis indicates that there are some differences between the *Augmented Matrix*, *Matrix*, and *Gantt* representation. In case of Task 2 (identifying groups meeting at different locations) there is a significant difference between the *Augmented Matrix* and the *Gantt* chart. This corresponds to the preference rankings where the *Gantt* representation was rated third. One of the reasons why the *Gantt* chart did not perform as well as the other two alternative visualizations is probably that the white space between the bars of the *Gantt* chart were sometimes misleading. Participants preferred the *Matrix* visualization most, followed by the *Augmented Matrix*. Reasons given for the high ranking of the *Matrix* were that it enabled the participants to solve the tasks in a structured way. The *Augmented Matrix* was criticized because of the use of lines which sometimes led to clutter. However, some participants appreciated the auxiliary lines as they made it easier to see relations and changes of locations. This is probably related to the result that the lines sometimes made the identification of meetings easier and sometimes more difficult. Although the impact of these lines on overall correctness was not significant as shown by the ANOVA analysis it would make sense to investigate this issue in more detail based on the comments of the participants and on the recognition rate of groups.

Tasks. The results show that the visualizations are better suited for the duration detection task (Task 1) than for detecting location changes of groups (Task 2). A reason for that can be that Task 1 which asks for the duration of meetings is simpler than the task of detecting two or more meetings of the same group of persons (Task 2). In both tasks, groups with longer meetings were easier to identify than shorter ones most likely because they were visually more evident.

Test Cases. The study also shows that the participants performed better on the smaller datasets (Test Case 1 and 2) than on the larger datasets (Test Case 3 and 4). Although we expected this result, it was interesting to see that there was no significant interaction between the test cases and tasks nor between test cases and visualizations. Furthermore, the test cases did not influence the participants' subjective preference.

One issue to keep in mind is that we conducted the study with students and not with experts. As mentioned above the investigation of effectiveness and utility for group detection was the primary goal of this study. This depends more on the mechanisms of human cognition and less on domain-specific knowledge. We thus decided to choose students as sample since it is difficult to reach a large number of experts which also have the time to participate in a study that takes around two hours. Finally, when interpreting the results one should consider that the colors may have influenced the recognition rates of groups as the results of the qualitative content analysis suggest a certain effect of the colors on which groups were perceived as salient. However, by using an established color scheme and using different color combinations for different groups we tried to mitigate such effects.

7 Design Considerations

Based on the above discussed results we propose the following design considerations:

Use map in combination with other visualizations: It may be useful to combine a map with one of the other visualizations, for example, in a coordinated multiple view setting.

Use auxiliary lines carefully: We suggest to use auxiliary lines with caution, for example, only on a certain subset of entities.

Avoid white space between entities of the same group: It may be beneficial to provide possibilities to rearrange the order of entities in the visualization to ensure that entities belonging to the same group are displayed next to each other.

Show duration of meetings explicitly: Showing the time explicitly improved the identification of (recurring) group meetings.

Consider the influence of visual properties: The qualitative content analysis revealed an influence of certain properties of the visualization on which groups were perceived as salient or not.

8 Conclusion

We conducted an empirical evaluation to identify appropriate visualizations to support users in the activity of identifying meetings of persons, vehicles, or other entities. As expected, map visualizations did not perform very well. The other visualizations which we used performed significantly better. Based on the research results we also developed recommendations for the design of these visualizations. There are several issues which were not addressed in this study. Among them is the question which kinds of interactivity should be adopted to support the users. It is, for example, likely that the possibility to select interesting cases could assist users in solving their tasks. In addition, studies with domain experts are necessary to clarify in which areas the visualization we suggested are most beneficial.

Acknowledgments. removed for blind review

References

1. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: Visualization of Time-Oriented Data. Springer (2011)
2. Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., Wrobel, S.: Scalable analysis of movement data for extracting and exploring significant places. IEEE Trans. Vis. Comput. Graphics 19(7), 1078–1094 (2013)

3. Andrienko, G., Andrienko, N., Mladenov, M., Mock, M., Pölitz, C.: Discovering bits of place histories from people's activity traces. In: Proc. VAST'10. pp. 59–66. IEEE (2010)
4. Boba Santos, R.: Crime Analysis with Crime Mapping. SAGE Publications, 3rd edn. (2013)
5. Chorley, M.J., Whitaker, R.M., Allen, S.M.: Personality and location-based social networks. Computers in Human Behavior 46, 45–56 (2015)
6. Ghoniem, M., Fekete, J.D., Castagliola, P.: A comparison of the readability of graphs using node-link and matrix-based representations. In: Proc. INFOVIS'04. pp. 17–24. IEEE (2004)
7. Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C.D., Roberts, J.C.: Visual comparison for information visualization. Information Visualization 10(4), 289–309 (2011)
8. Guo, D., Chen, J., MacEachren, A.M., Liao, K.: A visualization system for space-time and multivariate patterns (vis-stamp). IEEE Trans. Vis. Comput. Graphics 12(6), 1461–1474 (2006)
9. Harrower, M., Brewer, C.A.: Colorbrewer.org: An online tool for selecting colour schemes for maps. In: Dodge, M., Kitchin, R., Perkins, C. (eds.) The Map Reader, pp. 261–268. John Wiley & Sons, Ltd (2011)
10. Kessell, A.M., Tversky, B.: Visualizing space, time, and agents: production, performance, and preference. Cognitive Processing 12(1), 43–52 (2011)
11. Kjellin, A., Pettersson, L.W., Seipel, S., Lind, M.: Evaluating 2d and 3d visualizations of spatiotemporal information. ACM Trans. Appl. Percept. 7(3) (2008)
12. Kraak, M.J.: The space-time cube revisited from a geovisualization perspective. In: Proc. ICC'03. The International Cartographic (2003)
13. Kriglstein, S., Pohl, M., Smuc, M.: Pep up your time machine: Recommendations for the design of information visualizations of time-dependent data. In: Huang, W. (ed.) Handbook of Human Centric Visualization, pp. 203–225. Springer (2014)
14. Kristensson, P.O., Dahlback, N., Anundi, D., Bjornstad, M., Gillberg, H., Haraldsson, J., Martensson, I., Nordvall, M., Stahl, J.: An evaluation of space time cube representation of spatiotemporal patterns. IEEE Trans. Vis. Comput. Graphics 15(4), 696–702 (2009)
15. Li, N., Chen, G.: Analysis of a location-based social network. In: Proc. CSE'04. pp. 263–270. IEEE (2009)
16. Luz, S., Masoodian, M.: Comparing static gantt and mosaic charts for visualization of task schedules. In: Proc. IV'11. pp. 182–187. IEEE (2011)
17. MacEachren, A.M.: How Maps Work - Representation, Visualization, and Design. Guilford Press (2004)
18. Orellana, D., Wachowicz, M., Andrienko, N., Andrienko, G.: Uncovering interaction patterns in mobile outdoor gaming. In: Proc. GEOWS'09. pp. 177–182. IEEE (2009)
19. Schreier, M.: Qualitative Content Analysis in Practice. SAGE Publications (2012)
20. Shen, Z., Ma, K.L.: Path visualization for adjacency matrices. In: Proc. EUROVIS'07. pp. 83–90. Eurographics Association (2007)
21. Ware, C.: Information Visualization: Perception for Design. Morgan Kaufmann Publishers Inc. (2004)
22. Wilson, J.M.: Gantt charts: A centenary appreciation. European Journal of Operational Research 149(2), 430–437 (2003)
23. Wu, H.M., Tzeng, S., Chen, C.h.: Matrix visualization. In: Handbook of Data Visualization, pp. 681–708. Springer (2008)