Arrangement of Face-to-Face Meetings Using Social Media

Jaroslav MICHALCO, Pavol NAVRAT
Slovak University of Technology in Bratislava,
Faculty of Informatics and Information Technologies,
Ilkovicova 3, 842 16, Bratislava, Slovakia,
michalcojaroslav@gmail.com, navrat@fiit.stuba.sk

Abstract: Social networks, especially Facebook, currently as the largest one, have made the organization of events apparently simpler. Facebook offers the event service, which has greatly simplified the invitation process. Still, organizing an event is usually coupled with the risk of guest's no-show. We report an investigation designed to identify factors that might help predict a person's likelihood of attendance to an event s/he is invited to. Our research tries to combine information research with information technology tool design. The factors affecting the probability were determined by an analysis of data acquired in surveys among hundred and fifty or so Facebook users. We also developed a program that implements (some of) these findings. Simple quantitative and qualitative analyses were carried out on the data, sufficient to identify some of the key factors influencing invitees in their decision to attend a meeting. The factors as identified by the surveys are indeed relevant to attendance of meetings arranged using social media. Our application can help the event creator estimate how many people would attend his event and predict the likelihood of each invitee's attendance. Moreover, the application can also help the invited guest learn which of his friends are likely to attend the same event. Research that combines information side with information technology side can be fruitful as shown by this simple result. There is room for future work in this interdisciplinary space.

Keywords: social network, Facebook, attendance prediction

1. Introduction

Organizing meetings of more people has a difficulty every organizer knows about – one invites a few dozens of friends to an event and only a handful of them show up. Many of the invited guests even do not respond to the invitation by either confirming or declining their attendance. The issue becomes even more problematic when the event is a mass event, such as a graduation class meeting, in which the event’s attendance may reach hundreds or even tens of thousands.

Not a long time ago, to organize an event, such as a graduation class meeting, one had to call every classmate and tell him/her about the event. When an invited guest did not respond right away, the organizer had to either wait for the guest to call back or he could call him again some time later to remind the guest of the event. Another option was to send everyone a short text message with the invitation. However, it could be difficult to write everything important in a limited space.

Social networks, especially Facebook as currently the largest one, have become an important media for social communication (Young 2011). In particular, they made organizing of events much simpler. This is a very recent observation. Still in 2008, with the event function already available (Hei-man 2008), the use of Facebook to attend an event organized online was relatively infrequent type of use according to Joinson (2008). Nowadays, to organize a meeting, concert or other event, in which many people participate, one does not have to send a text message or an email to all guests. Letting people know about an event translates into just a few clicks on Facebook and all invited guests get notified about it. All invited guests can see who has responded to the invitation and how. However, even though the organization process has changed, it seems people have remained essentially the same and many of them still do not respond to some invitations at all, which is a problem for the organizer, because he often needs to know at least the approximate number of guests, so that he can, for example set up a reservation in a pub or change the time of the event if not many people intend to attend the event.

The motivation of this work is to help the organizers to facilitate face-to-face meetings. Predictions of attendance are not only useful for setting up a reservation, but can also be used as a tool to better reach the invited guests and adjust the event in such a way (by inviting more guests or changing the location of the event) that more guests would attend. Using attendance predictions in this way has a huge potential not only in organizing parties, but also in organizing mass events or for advertising purposes.
2. Related Work

For the information technology part of our research, we have no knowledge of a Facebook application that predicts the likelihood of event attendance or otherwise advises the event organizer. However, some works using different platform are noteworthy in relation to the topic this paper.

Works determining the factors affecting attendance

Mynatt and Tullio (2001) created a program that is perhaps most similar to our project. Their calendar extension Ambush could predict the probability of person's attendance at the events on his schedule according to various factors. According to the authors "the priority of an event is influenced by a number of factors, including the alarm status, recurrence status, the type of the event, and the user's role in the event" (Mynatt and Tullio 2001). These factors were collected from informal interviews with potential users. The same factors were used in another work by these authors in a shared personal calendar called Augur (Tullio et al. 2002). Some of these factors could be incorporated into our application; on the other hand, these factors had been collected for the purpose of a calendar application, not a Facebook application, which is slightly different. The agent in Ambush could also learn over time, so its predictions became more accurate. However, when the semester ended, the accuracy of predictions decreased due to change of persons' habits.

The above mentioned works are not the only ones that are aimed on determining factors that affect attendance. Works by Douvis (2007) and Hansen and Gauthier (1989) focus on determining factors that affect user's attendance at sporting events. In Hansen and Gauthier (1989) the authors determined 40 factors that affect the attendance. Interesting approach is shown in Tomlinson, et al. (1995), in which the authors asked the respondents a negative question about factors that discourage people from attending sporting games. Unfortunately, most of these factors that were identified in these works are domain (i.e. sport) specific and are not applicable to events in general.

Works helping to manage events

There are some other works that help users manage events. Older works by Mitchell et al. (1994) and Maes and Kozierok (1993) incorporate an agent that learns the user's habits and helps the user organize events in his calendar. A recent project by Masli, et al. (2011a, 2011b) deals with a social calendaring site called Timely that helps the users organize and share events. This application provides an interesting feature. When an organizer is undecided about a certain attribute (such as time or location of the event), he can leave it up to the guests to decide on these attributes using a simple voting mechanism supported by this website. Web-based service for collaborative organization of academic events called Takeplace has been developed by Skrabalek et al. (2010). None of these works, however, includes estimating user's attendance probability, which is our fundamental intention.

Organizing events in the age of Facebook

Barkhuus and Tashiro (2010) conducted a research about the use of social networks in organizing face-to-face meetings. From their meetings with 18 students they identified four different types of events:

1. Scheduled social gathering: These are mostly regular events created on Facebook.
2. Semi-scheduled social gathering: These are mostly after-school events, such as lunch usually organized by text messaging or phone.
3. Ad-hoc social gathering: Ad-hoc meetings are usually organized by Facebook statuses or by accident.
4. Special events: Events such as birthday parties or reunions are mostly based on invitations and are facilitated by Facebook Events.

Lastly, the researchers discovered that Facebook might affect student's decision about attending an event. As one of the students recollected: "If my friend put that she's going to an event that I wanted to check out as well ... like the seminars or like lectures ... sometimes I want to check them out ... but I hate going to those kind of things by myself" (Barkhuus and Tashiro 2010). This sentence is of serious importance for our research and we will explore the decision process about attending an event later in this paper.
3. Factors Affecting Attendance

The most challenging issue in creating an application that estimates the likelihood of attendance at events is to identify the factors that affect users' decision. The better is the identification of such factors; the more precise estimations are given.

There are several factors that can affect one's attendance at the event. Mynatt and Tullio (2001) created a Bayesian network to model event attendance and identified several factors. To create such a network, we have to find out the preferences of Facebook users. We chose a survey as a tool for determining the factors that affect one's likelihood to attend an event and to identify the strength of these factors.

First survey

Several factors can affect one's attendance at an event. We devised a survey to determine the factors that affect one's likelihood to attend an event and to identify the strength of these factors. This survey was given to 151 Facebook users from various age groups. The sizes of age groups are shown in Table 1:

<table>
<thead>
<tr>
<th>Age group</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>below 16</td>
<td>2</td>
</tr>
<tr>
<td>from 16 to 19</td>
<td>21</td>
</tr>
<tr>
<td>from 19 to 25</td>
<td>115</td>
</tr>
<tr>
<td>from 25 to 30</td>
<td>9</td>
</tr>
<tr>
<td>above 30</td>
<td>4</td>
</tr>
</tbody>
</table>

In this survey we asked the respondents to determine the importance of the following factors on a 7-degree scale: Time of the event, Inviter to the event, Other guests and Other factors. As other guests invited to the event might affect user's attendance in both positive and negative way, we asked the respondents to rate, whether the following statements are true – slightly true – slightly false – false in their consideration about attending an event. The statements in survey were:

- "My friend goes there so I'll go there"
- "All of my classmates/colleagues/... will be there so I'll be there"
- "If s/he goes there then I won't"
- "Only a few people will go there so I won't go there".

Table 1. Sizes of age groups in survey 1

<table>
<thead>
<tr>
<th>Age group</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>below 16</td>
<td>2</td>
</tr>
<tr>
<td>from 16 to 19</td>
<td>21</td>
</tr>
<tr>
<td>from 19 to 25</td>
<td>115</td>
</tr>
<tr>
<td>from 25 to 30</td>
<td>9</td>
</tr>
<tr>
<td>above 30</td>
<td>4</td>
</tr>
</tbody>
</table>

After collecting the answers we adjusted the 7-degree scale into a 4-degree scale representing the strength of these factors, where 1 means "no influence" and 4 means "strong influence". From these numbers we computed relative frequency for each factor to get the strengths of these factors. From these strengths we computed weights that will be used to compute the probability of user's attendance. The strengths of all factors enquired in this survey are shown in Table 2. The sum of all weights of subfactors is equal to the weight of corresponding factor and the sum of all factors is equal to one.

Table 2. Strength of factors according to survey 1

<table>
<thead>
<tr>
<th>Factor</th>
<th>Subfactor</th>
<th>Strength</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td></td>
<td>2.9272</td>
<td>0.3525</td>
</tr>
<tr>
<td>Inviter</td>
<td></td>
<td>2.5563</td>
<td>0.3078</td>
</tr>
<tr>
<td>Other people</td>
<td>My friend goes there, I’ll go there</td>
<td>2.8742</td>
<td>0.1061</td>
</tr>
<tr>
<td></td>
<td>All of my classmates/... will be there, I’ll be there</td>
<td>2.5894</td>
<td>0.0956</td>
</tr>
<tr>
<td></td>
<td>If s/he goes there, I won’t</td>
<td>1.6554</td>
<td>0.0612</td>
</tr>
<tr>
<td></td>
<td>Only a few people will go there, I won’t go there</td>
<td>2.0795</td>
<td>0.0768</td>
</tr>
</tbody>
</table>
Second survey

After finishing the survey as described above we decided to launch a second survey in order to find other factors and determine the relevance of factors mentioned above. We asked 152 respondents to rank the importance of these factors from most important to least important: Time of the event, Place of the event, Distance to the event, Type of the event, Other guests at the event, Inviter to the event, Weather, Finances, Mood and whether the person has set a reminder to remind him of an event.

Respondents were supposed to rank these factors on a scale from 10 to 1, where 10 was the most important. From the collected data we computed the strength of these factors as mean values of importance. The results show that the preferences of people vary according to sex, age (high school, undergraduate, graduate students, older) or type of job (for example full-time, part-time employees). As a result, we believe that it would be worth exploring these differences and also implementing the differences in weights of factors. Table 3 shows the strengths of the individual factors, the computed weights and also the strengths according to age groups. The three strongest factors for each age group are in boldface.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Strength</th>
<th>Weight</th>
<th>15 - 18</th>
<th>19 - 22</th>
<th>23-25</th>
<th>25+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other guests</td>
<td>8.00</td>
<td>0.1456</td>
<td>8.83</td>
<td>7.98</td>
<td>8.39</td>
<td>6.00</td>
</tr>
<tr>
<td>Type</td>
<td>7.61</td>
<td>0.1386</td>
<td>7.17</td>
<td>6.89</td>
<td>8.19</td>
<td>8.00</td>
</tr>
<tr>
<td>Time</td>
<td>7.21</td>
<td>0.1313</td>
<td>7.00</td>
<td>6.70</td>
<td>7.55</td>
<td>7.37</td>
</tr>
<tr>
<td>Mood</td>
<td>6.21</td>
<td>0.1131</td>
<td>8.17</td>
<td>6.13</td>
<td>6.10</td>
<td>5.50</td>
</tr>
<tr>
<td>Place</td>
<td>6.05</td>
<td>0.1102</td>
<td>7.00</td>
<td>5.89</td>
<td>5.74</td>
<td>6.50</td>
</tr>
<tr>
<td>Distance</td>
<td>5.32</td>
<td>0.0968</td>
<td>4.00</td>
<td>5.65</td>
<td>4.81</td>
<td>6.37</td>
</tr>
<tr>
<td>Inviter</td>
<td>5.08</td>
<td>0.0925</td>
<td>5.17</td>
<td>5.54</td>
<td>4.65</td>
<td>5.50</td>
</tr>
<tr>
<td>Finances</td>
<td>5.07</td>
<td>0.0923</td>
<td>4.33</td>
<td>5.35</td>
<td>5.10</td>
<td>4.75</td>
</tr>
<tr>
<td>Weather</td>
<td>3.20</td>
<td>0.0582</td>
<td>2.33</td>
<td>3.24</td>
<td>4.48</td>
<td>3.87</td>
</tr>
<tr>
<td>Reminder</td>
<td>1.18</td>
<td>0.0216</td>
<td>1.00</td>
<td>1.44</td>
<td>1.00</td>
<td>1.12</td>
</tr>
</tbody>
</table>

4. Design of a Tool

Our research continues to the information technology sector, since we wanted to operationalize the findings resulting from the analysis. Our aim was to design and implement a simple tool that supports the networked people when arranging a face-to-face meeting. It helps both the person who invites and those who are invited. We designed the tool as an application that uses Facebook's Social Graph http://developers.facebook.com/docs/reference/api/, in which all users are connected with everything they can interact with (such as events or pages). To get data from social graph we used both Facebook Query Language and Facebook Graph API, whichever was more suitable. However, each Facebook user can have different rules to access his data, which is a limiting factor for this application, because some important data can be unattainable with these APIs.

When a user creates an event, the application creates a regular Facebook event and invites all the selected guests to it. After that it obtains all necessary data from all invited guests to make a prediction. When the application has all necessary data, it computes the probability of attendance for each factor separately. If the
factor consists of more subfactors it computes the probability for each subfactor separately and then computes the overall probability for this factor. Finally, it computes the overall probability and saves it to database.

**Overall probability**

To compute the overall likelihood of user's attendance at an event, we use the equation (1):

$$P = \sum_{i=1}^{\text{factors}} \frac{P_i \omega_i}{\omega_i}$$

(1)

In this Equation $P_i$ represents the probability computed from $i^{th}$ factor and $\omega_i$ represents the weight of $i^{th}$ factor. For factors we use the weights from Table 3, because of higher relevance of these values and for subfactors we use the weights from Table 2.

As we can see in Equation 1, the first step in determining user's probability of attendance is to compute the probability for each factor separately and after that compute the overall probability. The same method can be used for a factor that consists of more subfactors (such as other guests according to first survey). For the purposes of this paper we will focus on the factors of time, inviter to the event and other guests. According to Table II these factors account for 37% of the overall influence.

**Time of the event**

Time is a very important factor affecting user's attendance. There are some days when a person has free time and can attend events, and there are some days when a person is so busy that he cannot attend any event at all. To find out, whether the day of the event suits the invited guest we use his replies to all previous events that took place on the same day of the week as the event he is invited to. As soon as we have the number of all invitations and responses to a certain day for a certain guest, we can compute the probability of this person with respect to the factor of time using the following formula, where $P_t$ represents the probability that a person comes to the event because of person $i$, $S$ is the number of same responses to invitation and $R_{both}$ is number of all invitations that both persons responded to.

$$P_{time} = 0.50 + 0.50 \frac{A}{R} + 0.50 \frac{0.50 \cdot M}{R} + \frac{N}{R}$$

(2)

Probability of time depends on these variables: ratio of "attending" responses ($A$), "maybe" responses ($M$) and "not attending" responses ($N$) to the number of all responses ($R$) for a certain day. If the guest has not responded to any invitation to a certain day (or he has not received any invitation), the probability of time for that day is set to 0.5 – which means we are completely uncertain about his attendance.

**Other guests**

When deciding whether to attend an event, persons often look at other guests that are invited to the same event and their decision is affected by the strength of the bindings with these people. The factor of other guests can be divided into four subfactors as mentioned earlier in this paper.

According to our preliminary research, if two people are good friends, it is very likely that if one of them decides to attend an event, the other one will attend as well. To implement this we need to estimate the strength of social bond between two persons by using all past events of these two persons and their responses to them. However, we are aware that past events cannot solely give us precise estimation of a social bond. It would be worth considering taking other factors, such as common photos or number of responses to statuses, into account.

From these data we estimate the strength of a social bond by using following formula, where $P_i$ represents the probability that a person comes to the event because of person $i$, $S$ is the number of same responses to invitation and $R_{both}$ is number of all invitations that both persons responded to.

$$P_{\text{guest}} = \frac{S}{R_{both}}$$

(3)

To compute the overall probability of other guests, we will use the Equation 1 for all subfactors.

1. Positive influence of other guests: Positive influence of other guests invited to the same event means that if a friend of person A decides to go to an event, it is more likely that person A will attend as well. If the strength of social bond is greater than 50%, we consider this relationship as positive. The overall positive influence for person A will be computed as a mean value of all the positive influences.

2. Negative influence of other guests: Negative influence of other guests means that if person B attends an event, person A is more likely not to attend. We assume
that when the strength of social bond is lower than 50%, this relationship is negative. The overall negative influence for person A will be computed as a mean value of all the negative influences.

Influence of small attendance: There are people, whose decision about attending an event depends on the amount of persons that attend that event. If the person sees that there are not many people interested in the event, he tends to hit the "Not attending" button. Therefore if the confirmed attendance is lower than some predefined threshold 25% of all invited guests, we count the probability for this subfactor as 0%, otherwise it will be 100%. (25% is only an estimated value and should be rectified by empirical observations.)

Inviter to the event

For some people it is important who invites them to the event. If their best friend organizes an event, they are more likely to attend than as if someone whom they don't know invited them. To compute the probability we use Equation 2, as we used for the factor of time, where R in this case will be number of requests a user received from organizer.

5. Verification of Results

After finishing the implementation, we tested our application on the first author's Facebook friends. Testing set had 251 persons currently living in 35 different cities and 15 countries. Firstly, we collected some general statistics about events and invitations to events. Secondly, we tested our application on past events and compared the predictions for them with the response of invited users. Lastly, we tested our application on past events and compared the estimations with real values, whether the user really attended the event, or not.

General statistics about events

Firstly, we collected some general statistics about events and tried to compare them with results from our first survey. On April 1, 2012 we downloaded all events from 251 participants for the past year, which accounted for 16,736 events. From these events we were able to extract some valuable statistical data.

Invitations per month

According to our first survey, 88% of persons get at most 10 invitations to events per month. To justify this statement, we counted all invitations these persons. Then we computed an average value for one person per month. The average number of invitations per month and person was 12, which is not consistent with results of the first survey.

Responses to invitations

According to our survey, 64% of persons respond to most of their invitations. To determine the truthfulness of this statement we counted all invitations in last year and all responses to these invitations. We discovered that 69% of all invitations were not responded, which also conflicts with the result from our survey. A pie chart showing the ratio of responses to invitations is shown in Figure 1.

![Figure 1. Replies to invitation](image)

Events per day of week

One of the statistics we extracted is the amount of events per day of week. From all 16,765 events in past year we counted all events for each day of week. As we can see in Figure 2, most events take place on Friday and Saturday and lowest amount of events is happening on Monday.

Comparison with responses on Facebook

After collecting general data about users and events, we tested the accuracy of our application. From the events downloaded earlier we selected 75 events that the first
author of this work was invited to, in which there were less than 300 guests. We did not select the events the author was not invited to, because in those events there were more people that were not friends with the author, and in that situation the results were incorrect due to Facebook's privacy policy and inability to access information for those people, whom we do not know. The reason why we only chose events with less than 300 guests is that for more users the execution time was extremely high, and in large events there are also many guests whom we do not know. For each of these 75 events we selected all guests, who were friends with the first author, and made an estimation of the attendance for them. All in all, there were 1,673 invitations to these 75 events. From these invitations we separated 728 responses that were either "attending" or "declined", so that we could compare them with our predictions. We let our application estimate the likelihood for these invitations and compared the results with responses. If the prediction was higher than 50%, we counted the prediction as "attending" and if it was lower that 50% we counted as "not attending". As a result, 538 predictions out of 728 were correct and therefore the accuracy of our estimations was 73.9%, which is a satisfying number.

Comparison with actual attendance

Lastly, we used 15 past events, in which the first author remembered the actual attendance – whether the invited guest came to the event or not, regardless on what he claimed on Facebook. We compared only those guests that are friends with the author. There were 515 invitations to these 15 events, for which our application estimated the attendance and after that we compared the estimations with real data. Out of these 515 invitations there were 304 correct estimations, which gives us the ratio of 59%. In some events the ratio was 40%, but on the other hand, in one case there were 88.9% correct estimations. These ratios can be seen in Figure 3. When we compared these 15 events with responses on Facebook as in previous section of this paper, the accuracy was 69.1%.

5. Conclusion and Future Work

In this paper we tried to investigate factors that influence invitees who are on Facebook and moreover who were invited to an event by using the Facebook event facility. Having identified a set of factors, we created a computer application that estimates the probability that a person will attend an event he is invited to. As we have found out, almost 70% of all invitations are not responded, and therefore the estimations of attendance have a
lot of potential for organizing parties, estimating the number of demonstrators or for business purposes. If the estimations were mostly correct, a company that creates an event where they want many attendees can use this application as a tool to adjust the event information in such way, that most people would attend.

The main purpose of this paper was to identify factors that can affect a person's decision. We identified ten factors that affect person's attendance at an event by using two separate questionnaires. We incorporated just three of these ten factors into our experimental prototype tool and tested the correctness of predictions. Our application gave correct results in about 60% of cases. As the three implemented factors account for only 36.94% of all the identified influences, this number is satisfying and promises a better result, when more factors are incorporated.

**Limitations of this work**

One of the possible problems with the application is that the predictions work only for friends of a person that allowed this application to access his data on the Facebook. The reason for this is that we cannot access most of the necessary data for persons that have not allowed our application. This could be solved by persuading more and more Facebook users who would like to benefit from receiving our assistance in arranging events to allow our application to access their data. Alternatively, this application might be implemented directly into Facebook, so that it becomes a feature of Face- book Events. In latter case, it would no longer be necessary to download all data to local database and the privacy issues would be overcome, as these data are already in Facebook.

**Future work**

To improve estimations we following work is needed:

- Further sociological research
- Implementing more factors

**Further sociological research**

Correct estimation of weights of factors is essential for correct estimations. To get more relevant values we would need to use regression analysis, which would require a longer research and collecting real empiric data. It should be considered whether it would be sufficient to ask the users retrospectively after each event how they were satisfied with various factors and whether they attended the event, or whether a completely different approach is needed. Another approach how to estimate the weights correctly is to use a learning agent that would adjust the weights for each person individually, which would provide us more accurate predictions than general weights. We can also set the weights according to demographic data, as our research suggested that factors like sex, age or occupation affect the importance of factors.

**Implementing more factors**

The three factors mentioned above do not account for whole influence on person and there are more factors affecting persons' decision, as mentioned in Table 3. Some of these factors, such as type of the event or user's attitude toward the place of the event can be obtained through Facebook and by parsing obtained data. Other factors, such as weather or distance to the event could be obtained from applications outside of Facebook, such as Google Maps or Foursquare. Some of the factors, such as mood or finances, are very difficult to determine.

Another problem with an application like our one are privacy settings of Facebook users. We suppose that the most accurate results would be when used among friends, because a user cannot approach data from a person, who is not his friend. However, if a person allows our application to use his information, his data then can be reached. The more users will use our application, the more accurate are the results. In conclusion, we believe that the best performance of our tool would be reached when it is implemented directly into Facebook feature, not as a third-party application. Nevertheless, each obstacle can be overcome in some way. For example, even though Facebook is not a calendar and does not have a reminder, Barkhuus and Tashiro (2010) found out in their research that some users use status messages as a reminder of an event. The possibilities to improve this application in future are countless.

From a more general point of view, it should be borne in mind that social networks treated as collaboration networks have a great potential for discovery of knowledge (Tutoky 2011). Arrangement of events is just one of many...
possible features that social networks participants will increasingly enjoy. Social networks open new possibilities for exploratory search (Tvarozek 2011). Other important concepts, e.g. homophily (Vojtek 2010) help study relations between humans within their social connections.

Acknowledgements

This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV- 0208-10.

REFERENCES


