TAG BASED RECOMMENDATION SYSTEM OF A SOCIAL NETWORKING SITE* 

Hatice BÜBER KAYA  
*Kırklareli University, Kayalı (Merkez) Yerlesshkesi, Kırklareli, Turkey 

Ozan KAYA, MBA  

© The Author(s) 2020 

ABSTRACT 

Social networking sites (SNSs) are web-based services that allow people to create a profile and communicate with each other via their personalized web pages (profiles). The history of SNSs begins in 1997 with the site named SixDegrees.com and from 2003 onwards too many SNSs have been launched and hence the term YASNS: “Yet Another Social Networking Service” came into our lives. In 2004 with the launch of MySpace, SNSs gained popularity remarkably and hit the mainstream. 

Eventur was a tag-based event recommendation system about local events in Pittsburgh. The system used an event recommendation system to make recommendations more personalized. Recommendations are based on two of the core elements of social network - people and tags. Relationship information among people, tags, and items, is collected and aggregated across different sources within the enterprise. Tags applied on the user by other people are found to be highly effective in representing that user’s topics of interest. This study was performed to understand the effectiveness of the Eventur system and this paper provides the methods, highlights the findings and suggests possible future research. 

KEYWORDS 

Social networking sites (SNSs), web-pages, profiles, recommendation system, Eventur system, core elements
**Introduction**

Eventur was a tag-based event recommendation system about local events in Pittsburgh. “Eventur completely relies on users to collect local cultural events and contribute to the system, serving as a content integration platform. Meanwhile, the system provides community features by giving the users online identity and provide social features, such as friending and recommending events to each other”. Guy and Carmel define Social Recommender Systems (SRS) as systems to reduce the level of data exposure on social media users through its algorithm that funnels the most relevant information for users. Another major task of SRSs is to increase the total time spent, as well as the amount and frequency of activity on a social media site (Guy & Carmel, 2011). The main difference between the Social Recommender Systems and the classic Recommender Systems is, by gathering data from users’ profiles and interaction between the users, social networking sites can generate a better system that can suggest more relevant recommendations (Al Falahi, Mavridis, & Atif, 2012).
Content-based filtering methods use algorithms that aim to suggest items to the users based on information about the previous and current likes of the users (Brown, Robertson, & Sullivan, 2012). User profile is made up of total data collated through the user’s online activities such as the items user has seen and liked/unliked and or rated in some form. During the recommendation process, the properties of unseen pages are compared with the information stored in the user profile and those resemble a higher match are recommended to the user (Mobasher, 2007). Nevertheless, we are faced with some systematic problems with the content-based filtering approaches. For example, user profiles are highly dependent on text that is made up of items which are previously visited and/or rated by the user. In this context the data misses the critical relationship between the items within the web pages (Fernández, Arias, Solla, Cabrer, & Nores, 2008). Moreover, item utilisation which is a common feature between objects in a particular context is not taken into account while formulating the user profile, which then create the base for recommendations of new items (Anand & Bamshad, 2003).

Collaborative-based (or in other words social-based) approach is an alternative to the previous approaches, targeting to overcome the restrictions of content-based approach. Rather than relying on the items seen and/or rated by the particular user, this approach collates the profiles of the users within an online community and recommends similar items to the users, based on the presumption that users share similar interests within a particular community (Amini, Ibrahim, & Othman, 2011). Collaborative filtering methods collect, process and analyse a huge chunk of information gathered about the users’ online behaviour, focusing on their activities, preferences in order to predict what users will like by comparing and contrasting their behaviour vis a vis other users. Collaborative-based filtering was also augmented to include social information in order to enhance the detecting of user neighbourhood and bring a solution to data sparsity (Groh & Ehmig, 2007; Ma, Yang, Lyu, & Ki, 2008).

Known systems of recommendation like the Content-Based systems were upgraded with data and functions that include friendship and interaction data into the function formula, while designing the hybrid systems (Liu & Aberer, 2013). There are a few methods to make use of hybrid approaches; for example making content-based and collaborative based assumptions and joining those predictions (Adomavicius & Tuzhilin, 2005), secondly adding content-based capabilities and collaborative-based approach into each other (in both orders) (Hoekstra, 2010), or thirdly joining two approaches and making a new model (Boyd & Ellison, 2007).
The Eventur system had some substructure features of a content-based recommendation system at the time of study. These features were being used for narrowing down the search results by some attributes like:

- Date based grouping (Upcoming events, grouped by daily and monthly events)
- Type based grouping (Dance, Exhibition, Festival, Film/Video Art, Lecture, Literary Program, Music Concert, Opera/Musical, Other, Play, Sports)
- Organization based grouping
- Venue based grouping
- Neighbourhood based grouping

Datasets

The system aimed to show different events. There were ten different types of events to be shown by the system. In the dataset that was used for this study, there were a total of 130 events such as:

| www.ijhsdr.com |
50 of them were chosen for the first phase of study. While choosing the first group of events, the number of events in each type calculated as weighted with the formula to select a diverse group:

\[
1\text{st Phase type } n \text{ event count} = \text{round} \left\{ \left[ \left( \text{total type } n \text{ event count} \right) \times \left( \text{1st Phase total event count} \right) \right] / \left( \text{total event count} \right) \right\}
\]

Ex: 1st Phase “Music Concert” type event count = round \{28\times \frac{50}{130}\} = \text{round} (10.7) = 1

The remaining 80 of the total number of events were used in the second phase of the study.
**Pre-Work**

The tagging process was needed in order to run the algorithm. More than one tag submitted by more than one person was needed to have more accurate results. The problem faced at that point was where to find those people (must be objective) as soon as possible for 130 events.

Amazon Mechanical Turk (MTurk) is defined as "a crowdsourcing website for businesses (known as Requesters) to hire remotely located "crowdworkers" to perform discrete on-demand tasks that computers are currently unable to do" (mTurk).

Tagcow (by MTurk) provides descriptive tagging service to add metadata to images that helps the requester finds and organizes their image libraries more effectively (TagCow).

To use Tagcow,

- Screenshots were taken and uploaded to an image hosting system for every single event.
- A .CSV (A comma-separated values (CSV) file consists of any number of records, separated by line breaks of some kind; each record consists of fields, separated by some other character or string, most commonly a literal comma or tab ) file created with the event url, event name and event screenshot for every single event.
- A design layout was created for HITs.
- After uploading the csv file, in a very short time; all 130 events were tagged by three people.

<table>
<thead>
<tr>
<th>Batch Results from MTurk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input.event</strong></td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td><a href="http://eventur.The">http://eventur.The</a> Addams F</td>
</tr>
<tr>
<td><a href="http://eventur.Amish">http://eventur.Amish</a> Monkey</td>
</tr>
<tr>
<td><a href="http://eventur.Amish">http://eventur.Amish</a> Monkey</td>
</tr>
<tr>
<td><a href="http://eventur.Amish">http://eventur.Amish</a> Monkey</td>
</tr>
</tbody>
</table>
Algorithm

The recommendation program was coded in python (most_corresponded_tags.py - see the screenshot) and the algorithm is as follows:

1. Takes the file that was downloaded from MTurk that includes the HIT results (results_from_mt.csv - see the screenshot).

2. Cleans (takes the irrelevant information out because there were some information about the time of submission and like that) the file and removes the duplicate & empty tags for the same event then parses the first file in a new file (all_events_and_their_tags.csv - see the screenshot).

3. Then takes the event ids that the subject picked as they like from the 1st Phase dataset list.

4. Writes the set of liked events' unique tags into an array.

5. Compares events with picked events' tags and gives reputation.

6. Compares the events that are going to be recommended to subject with 1st Phase dataset list in order not to recommend the events that were already seen and rated by the subject.

7. Sorts the events that are for recommendation by reputation.

8. Writes down the recommended event ids and reputations to a new file (reputations_of_events_which_are_the_most_corresponded_to_events_3099_4410_4445.csv - see the screenshot).
Demographics of the subjects

By Sex:

Male 8
Female 7

By Education Status:

Graduate 10
Undergraduate 5

Summary of 1st Phase User Study

Subjects were asked to tick yes if they like the event that was being showed. They were presented randomly in a mixed format to reduce the possibility of subjects’ rapid/inaccurate answering due to boredom. The e-mail addresses were required from the subjects at the beginning to follow up with the study. The email submission page and an example form are shown in the pictures below. A page break was placed after every question in order not to make subjects confused.
Picked events statics by gender

As it is shown below, most of the event types were picked with the same frequency by both males and females. However, while Literary Programs were more likely to be picked by male subjects, the Opera/Musical type events were liked by female subjects mostly.
Summary of 2nd Phase User Study

After getting the results of the 1st Phase study, the code was run with the results for every subject separately. The top ten recommendations were sent to the subject as follow up study, and the subjects were asked to rate the recommendations in the same way that they picked the events that they liked.
Results

Unique recommendation event IDs and their occurrence

There were 38 out of 80 unique events that have been recommended to subjects. The number of the recommended events (38) was lower than the un-recommended events (42). Un-recommended events were not forwarded to the subjects because they were not preferred enough based on the study done with the 50 subjects.
<table>
<thead>
<tr>
<th>Type</th>
<th>Times</th>
<th>Out of</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>1</td>
<td>5</td>
<td>0.20</td>
</tr>
<tr>
<td>Music Concert</td>
<td>14</td>
<td>18</td>
<td>14.18</td>
</tr>
<tr>
<td>Opera/Musical</td>
<td>5</td>
<td>11</td>
<td>0.45</td>
</tr>
<tr>
<td>Dance</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Festival</td>
<td>7</td>
<td>12</td>
<td>0.58</td>
</tr>
<tr>
<td>Play</td>
<td>6</td>
<td>11</td>
<td>0.55</td>
</tr>
<tr>
<td>Literary Program</td>
<td>2</td>
<td>10</td>
<td>0.20</td>
</tr>
<tr>
<td>Sports</td>
<td>2</td>
<td>6</td>
<td>8.00</td>
</tr>
<tr>
<td>Film/Video Art</td>
<td>0</td>
<td>2</td>
<td>0.00</td>
</tr>
<tr>
<td>Exhibition</td>
<td>1</td>
<td>5</td>
<td>0.20</td>
</tr>
</tbody>
</table>

As the table above represents, the frequency of the reputations and the number of events to be recommended were proportional. The most frequent one, “Music Concert”, was picked the most. Because the diversity of event types and the event counts were not even, it is wrong to make a conclusion about users’ preferences.

**Performance measures**

**Precision:**

Precision is the fraction of the documents retrieved that are relevant to the user's information need. For example, average precision @ 10 can be calculated:

\[
AP @ 10 = \frac{\sum_{i=1}^{10} (P[r] \times rel[r])}{\text{number of relevant items} \times \text{retrieval point}}
\]
Mean average precision

Mean average precision for a set of queries is the mean of the average precision scores for each query.

\[
\text{MAP} = \frac{\sum_{q=1}^{Q} \text{AveP}(q)}{Q}
\]

where \( Q \) is the number of queries.

Precisions are calculated for every step of the follow up (2nd Phase of study).

<table>
<thead>
<tr>
<th>Subjects</th>
<th>@1</th>
<th>@2</th>
<th>@3</th>
<th>@4</th>
<th>@5</th>
<th>@6</th>
<th>@7</th>
<th>@8</th>
<th>@9</th>
<th>@1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.87</td>
<td>0.88</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Sub 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>0.8</td>
<td>0.83</td>
<td>0.71</td>
<td>0.75</td>
<td>0.66</td>
<td>0.7</td>
</tr>
<tr>
<td>Sub 3</td>
<td>0</td>
<td>0.5</td>
<td>0.33</td>
<td>0.5</td>
<td>0.6</td>
<td>0.66</td>
<td>0.71</td>
<td>0.75</td>
<td>0.77</td>
<td>0.7</td>
</tr>
<tr>
<td>Sub 4</td>
<td>0</td>
<td>0</td>
<td>0.33</td>
<td>0.5</td>
<td>0.6</td>
<td>0.66</td>
<td>0.71</td>
<td>0.75</td>
<td>0.77</td>
<td>0.7</td>
</tr>
<tr>
<td>Sub 5</td>
<td>1</td>
<td>1</td>
<td>0.66</td>
<td>0.5</td>
<td>0.4</td>
<td>0.33</td>
<td>0.42</td>
<td>0.37</td>
<td>0.33</td>
<td>0.4</td>
</tr>
<tr>
<td>Sub 6</td>
<td>1</td>
<td>1</td>
<td>0.66</td>
<td>0.5</td>
<td>0.4</td>
<td>0.33</td>
<td>0.28</td>
<td>0.25</td>
<td>0.22</td>
<td>0.2</td>
</tr>
<tr>
<td>Sub 7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.85</td>
<td>0.87</td>
<td>0.77</td>
<td>0.8</td>
</tr>
<tr>
<td>Sub 8</td>
<td>1</td>
<td>0.5</td>
<td>0.66</td>
<td>0.75</td>
<td>0.8</td>
<td>0.66</td>
<td>0.57</td>
<td>0.62</td>
<td>0.55</td>
<td>0.5</td>
</tr>
<tr>
<td>Sub 9</td>
<td>1</td>
<td>1</td>
<td>0.66</td>
<td>0.5</td>
<td>0.6</td>
<td>0.66</td>
<td>0.57</td>
<td>0.5</td>
<td>0.55</td>
<td>0.5</td>
</tr>
<tr>
<td>Sub 10</td>
<td>1</td>
<td>0.5</td>
<td>0.66</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.42</td>
<td>0.5</td>
<td>0.55</td>
<td>0.5</td>
</tr>
<tr>
<td>Sub 11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sub 12</td>
<td>1</td>
<td>1</td>
<td>0.66</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.42</td>
<td>0.37</td>
<td>0.44</td>
<td>0.5</td>
</tr>
<tr>
<td>Sub 13</td>
<td>1</td>
<td>0.5</td>
<td>0.66</td>
<td>0.75</td>
<td>0.8</td>
<td>0.83</td>
<td>0.71</td>
<td>0.75</td>
<td>0.77</td>
<td>0.8</td>
</tr>
<tr>
<td>Sub 14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0.28</td>
<td>0.37</td>
<td>0.44</td>
<td>0.5</td>
</tr>
<tr>
<td>Sub 15</td>
<td>0</td>
<td>0.5</td>
<td>0.33</td>
<td>0.5</td>
<td>0.6</td>
<td>0.66</td>
<td>0.71</td>
<td>0.75</td>
<td>0.77</td>
<td>0.7</td>
</tr>
</tbody>
</table>

MAP 0.734 0.7 0.64 0.62 0.654 0.652 0.624 0.632 5.714 0.6

MAP 0.734 0.7 0.64 0.62 0.654 0.652 0.624 0.632 5.714 0.6
The highest mean average precisions (MAP) are at the first two recommendations, when we look at the 5th (in the middle) and the 10th (the last one) are around the same range with each other and also the average MAP, therefore we can suggest that the performance of this recommendation system is consistent.

**Limitations**

The most important limitation was not having enough count of events in every type. Since Eventur.us was a local web site for Pittsburgh, there was not enough chance to find many events in every classification of events. Since there was time limitation for the study, we could not wait until more events to be added to underrepresented classification of events.
Due to the nature of this study, human subjects were needed to test the algorithm. However, since the subjects are asked to rate 50 events in the 1st phase of the study, every subject indicated that they got bored during selection; therefore although we even tried to encourage them for being patient and not to rush the rating process without attention, we can never be sure about it.

**Conclusion**

Although the recommendation algorithm was a basic algorithm, the results showed that Average MAP was above 0.5 and so we can suggest that the algorithm was successful. Since several limitations were encountered during the research; with a better-structured study, a better algorithm and higher number of events to use for survey, this research could be improved. Therefore, a future study can be designed with improved conditions for further research.
References and notes:


