Master’s Thesis in Graduate School of Library, Information and Media Studies

Impact of Event Recommendation Systems in User's Decision Making

March 2017
201521654

Jesse Chiula Samacuva Ulundo

Impact of Event Recommendation Systems in User's Decision Making

Jesse Chiula Samacuva Ulundo

Graduate School of Library, Information and Media Studies
University of Tsukuba

March 2017

Impact of Event Recommendation Systems in User's Decision Making

201521654
Jesse Chiula Samacuva Ulundo

Recommendation systems are part of many computer applications today. The importance of such systems are vital due to the fact that the information shared on internet is fast growing. Many of the web users today depend solidly on information shared with them. In this sense, recommendation systems facilitate such users by automatically recommending their largest preferences.
Simultaneously, user’s needs and daily information seeking behavior is growing fast. So there is a huge importance of current recommendation systems adopt their algorithms to the vast user’s behaviors and their influential factors to help them with their final decisions.
Well known researches have been done to improve recommendation systems. Starting with more traditional systems such as user based recommendation, content based or collaborative filtering systems have been the pioneers to solve such problem. Hybrid methods where later on introduced but also looking more on joining the two previous mentioned methods. Most of this methods have not looked on introducing into their variables user’s social network activities and various influential factors that can be taken from social network from their exchange activities.
The purpose of this research is to bind the technical aspects of recommendation systems to most social aspects. So the overall purpose of this research is to evaluate the user’s behavior interacting with event based recommendation systems. By evaluating their behavior, I want to clarify the possible key factors that would influence the user’s decision making by the usage of the recommendation systems.
My research looks therefore to introduce such social network influential factors in the known hybrid recommendation method. I conducted a research to study and discover the influential factors to be introduced as an extra variable for the precision of recommendations based on the user’s behavior on social networks. I study such influential variables and evaluate them in a utility function to compare how this variable will perform on different recommendation algorithms. I conducted this experiments using an online event based recommendation system data collected from meetup.com.
Using the meetup API, I collected information from the Japanese most popular meetup cities Tokyo and Osaka. Therefore, to minimize and work with more data accuracy I worked with three categories that have similarities in activities thus a good way to measure impact and influence among users. The three decided categories where music, sports and camping. Each of this categories in this cities have average of more than 10 groups and each group having an average of 300 members. One of the key factors of influence is the usage of RSVP which stands for “Répondez s’il vous plaît” meaning please answer. This makes it possible for us to evaluate some infiltration among users from different groups and cities.
How results proof that due to social influential factors such as location of the user, social awareness of information being shared serves as an important role to influence an individual to accept the recommendation, thus, is safe to say that based on the interaction between the different recommendation methodologies there is a high impact and utility of recommendations in user’s decision making process.

Main Advisor: Prof. Masao Takaku
Sub Advisor: Taro Tezuka

Acknowledgments

I would like at first to thank God, for being my shield and support in all of the life specially during my academic currier, without him nothing would have been done and I would not have been at this position now.
Secondly my gratitude goes to my parents Abel Ulundo and Laurentina Ulundo for always being present and support me with wise councils and push me in hard moments. I can’t forget my sisters Henriqueta Ulundo da Cruz, Any Ulundo Sessoko, Analdina Ulundo Bravo and my Brother Henrique Ulundo for all their help and advises to believe on myself.
I would like to thank my dearest Professor Dr. Masao Takaku for all his advice and guides throughout this masters program. I would also like to thank all Professors that have helped me directly or indirectly with this work mentioning directly Prof. Hiroyuki Kitagawa, Prof. Taro Tezuka and Prof. Shigeo Sugimoto.
This research has been done according to the policies given by the Graduate School of Library Information and Media Studies of the University of Tsukuba.

Contents:

1. Introduction

Recommendation Systems are mostly seen in web applications. These systems that are becoming well known in most of web application, conduct predictions that respond to many of web user's needs. So it's possible and safe to say that Recommendation systems play a very important role today in solving many problems that users in previous years had, in achieving their information needs.

In our world today, information is growing exponentially. With such growth of information, is difficult for many times users to get exactly what they need. However, information and technology tools facilitate users to achieve to their daily and most common needs. Information retrieval tools for example, are well known for this matter, where the user will undertake a query to receive his information need. As good as information retrieval techniques are, they sometimes are time consuming and hard for users to get what they need. In usage of web services, it's important to focus on time as therefore recommender systems play their important role where users can get their needs satisfied by not doing anything.

Recommendation systems are known for their variety of technologies, most applied technologies or methods are the content based systems and the collaborative filtering systems. Both systems proved their importance, however, there are some problems which needs to be improved.

Many studies have been undertaken to come up with solutions to solve the problems that both systems present, one of this solutions are seen in the hybrid recommendation technique where it incorporates both content based and collaborative filtering systems. This method has shown great achievements and improvements concerning the effectiveness of recommendations. But still the problem to recommend new items to users are seen. Likewise, the lack of effective recommendation to new users is another reality.

For this reason, I am trying to incorporate the reality of social influence as a variable in recommendation systems to solve the problem of both new users and new items recommendations. But to start with such improvement I need to understand the real situation of impact that recommendations have on user's decision making process. So my first target in this research is to evaluate such impact by measuring the utility function that each recommendation method presents.

To take this into action I have to conduct studies based on influence and impact theories on a particular data to understand the behavior of user's when making their decisions. I will then evaluate such decisions and present their results to come up with a conclusion if by applying my methodology brings any improvement in the utility function. Then I measure the utility function in the presence of different recommendation methods to analyze if such factor can bring any improvement in the different recommendation processes. So let's look at the influential theory factor that will be the key variable during this research in analyzing impact recommendation systems have on user's decision making process.

1.1. Background

Many researches today do not look solidly into a single aspect of science. To improve the outcome of research results, researchers tend to find some solutions in other fields. Engagement of knowledge can well serve as a key factor to improve many questions being asked today. This same technique is being seen also in the computer science researches, whereby, many questions being asked before are not solidly being found in mathematical solutions or algorithms, but the tendency to improve algorithms answers have been coming from other sources. Recommendation systems being part of the great achievements of artificial intelligence today improvements of algorithms are not solidly found in the different prediction and ranking theories like presented in the past. The large need of contextual data been used today in recommender systems brings also the need to understand how humans interact with such systems. And nothing better than to look at theories that would explain given human behaviors and decisions to use such systems then to look at social influences that come directly with the daily human activities.

1.1.1. Social Influence Theory

Humans are social beings, looking at this perspective of life's relationship, it's vital to understand that great part of influence comes from interactions among users. Now, knowing that we all belong to a given group or society in general, we come to understand that most of our decisions are based on the activities we engage in this groups or society in general. Thus, is safe to say that most of our information needs, or needs in general are also somehow a product of our social interaction. For this reason, I came to understand that it is important to look at what influences us to come with given decisions. Taken this into consideration I will be able to bring forward key aspects of the social influence theory into the recommendation systems to be able to evaluate user's behaviors in their social activities to serve as a comprehensive information matrix in the context of collaborative filtering method.

Every day we all have our needs such as what type of news we want to hear, or what type of product we want to consume, the type of music we want to hear or places we want to visit. All of this information needs sometimes are hard for an individual to process and even harder for automatic algorithms to satisfy the individual needs. But all of us somehow are influenced by another person or a particular commercial we have seen. This factor is what we take as important in this research. According to many behavioral studies scholars such as Paul Lazarsfeld and E.M. Rogers, social influence has many branches. These two scholars have many researches bringing forth on how humans make their decisions and what are the key points of what influences or affects their decisions in their daily activities. Goldsmith in one of his books quotes:

"Theory is as an organized system of ideas or beliefs that can be measured; it is a system of assumptions or principles. The word 'theory' comes from the Greek verb theorein (to behold or contemplate). We form theories, for example, when we wonder why a couple has decided to marry or divorce. We look for clues to the outcome—why is the couple compatible or not compatible? A theory summarizes what is known about a phenomenon and permits the formation of hypotheses, or predictions about future occurrences (Goldsmith 2013, p. 45).".[1]

This could be understood that someone's influential rate is the combination of his interests and the interest of others around him. So for this case we can investigate and understand if an individual is mostly commonly found when making a decision to look at himself or at the other. With an answer to this we can utilize such combination to the recommendation process as social influence playing a vital role in the utility function.

1.1.2. Social Networks

Social networks are part of most of internet users today. In this social networks platforms, is very common for users to live their emotions, state of mind, their preferences etc. So, understanding that one way to measure social influence is by communicating and in communication we understand the level of intimacy among users, and this can be one of the key factors why social networks can be the pathway to understand some user's behaviors.

The benefit of introducing social network into my research is that they provide a large amount of collaborative data that can be used for recommendation procedures. In many social networks it's even common to see how recommendation systems are playing a big role in connecting people. For example, on Facebook collaborative methods are used to recommend new friends to a particular user. On Twitter, a fid given on a particular news story is key in content based recommendations for user to receive the ideal news stories. This are some examples that social networks are useful platforms to retrieve latent information to be incorporated into the recommendation procedure.

Incorporating social networks into the social influence theory will be vital to structure an individual's thoughts and knowledge about his/hers behaviors. The fact that recommendation systems utilize a lot of past information to predict future decisions is true. Problem is that to such methods users that are new to systems don't present enough past information to make an effective recommendation. If social network data is incorporated into a given recommendation process however a lot of factors can stablish or provide information for future predictions. Even if the user is new to the social network a lot of information from his friends can be used to provide future predictions. And for that reason aspects of behavioral theories such as the social influence theories can be applied to such aspects to understand user's decisions even when they are new to systems. According to social theories users get used to a given pattern of life. For example, they buy the same brand of shoes, they eat at the same locations or attend events that familiar people to him attend. It's very hard to be able to changed one's mind unless of very reasonable circumstance. This patterns can be studied and analyzed and understand why users make these decisions. There are no better place then social networks to investigate such online behaviors and incorporate them to hybrid recommendation methods.

1.2. Impact

According to many dictionaries definition, impact is the action of a specific object that uses force upon another by contact. This contact can be seen by it effect or influence. For the case of this research, I am trying to study and get the findings of how recommendation cause a given change in user's decision. This causes could be seen by the way an individual will react to a recommendation. This will be measured according to behavioral changes seen in the perspective of individuals belonging to a given group, and by the cause of the recommendation factors they make a change in their behaviors to attend events from other groups.

We will look the effectiveness of recommendations influencing the user's behavior to attend or not to attend a given event. Thus, such influence we define as the impact that the recommendations had on the user's decision making process.

1.3. Research Objective

Looking at the two theories one mostly inclined into human science and the other into computer science, my research incorporates both sciences to present the final results. Thus, the reason to investigate the key aspects that would influence a user's decision. And looking at this key factor I want to make a comparison study and analysis on how useful a recommendation can be such as that it can impact the user's final decision.

I therefore conducted a quantitative research studies on the meetup event recommendation system platform to understand such key factors and come up with hypothesis of the impact the recommendations have. I utilized the meetup data set to come at the results presented in this work.

1.3.1. Research Objectives

My particular research objective is to look at issues related to the impact that recommendation systems have on user's decisions. So we need to understand:

1. How do new users to recommender system respond to recommendations that are being attended by old users belonging to their group. My objective here is to explore the potential that previous RSVPs presented by old users or group creators can serve as an influential factor. In this case I consider that a given preference has rating value of a RSVPs as being positive and will work as a prior rating value;
2. Distance issue is another key factor that I would like to investigate to see how users tend to respond to events near them and how they respond to events far away from them.
3. How does contextual information play in the recommendation influence? In this area I want to explore how latent information such recommendation feeds, comments, ratings serve as an influential factor in user's decision.

This are the three main objectives that I will be focusing in this study analyzing it on an online data sample provided by the API of meetup.com. We will investigate in different procedures to look both at response from old and new users to the system. From this we can come up with a solution with the problem of new users to the system when there is not enough content based information to be explored thus not having an effective recommendation process.

1.3.2. Research question

With this possible hypothesis I bring the following research questions to be able to evaluate the above hypothesis:

RQ. 1. Are RSVPs enough to influence user's decisions?

RQ. 2. What are the main factors that would create a big impact to influence a given individual?

The above questions bring me to the central research question which is:

- Are recommender systems capable to influence user's decisions?

With this possibility been evidenced I take this study by using online data supplied by the event based recommendation web system meetup (www.meetup.com) [12] API. With the data been supplied by this API we want to achieve the above mentioned objectives.

1.3.3. Decision making process

All decisions are made based on a given need. For that reason, it is believed that the user's needs can be the initial value to understand his final decision. This initial value can be adjusted or molded to arrive to the final judgement. To come to this final judgement, the user has to pass by three other steps. The first step is believed that users judge their decision based on their preferences. The second step is believed that user takes more of psychological efforts to his preferences. At this stage the user looks at his surrounding to see who are the people making same decisions as his. Playing the psychological efforts this is one of the parts where the influence is more noticeable as his decisions are commonly made on others positive experience. This theory is similar to the great anchoring effects introduced by Jacowitz and Kahneman 1995[13]. This is believed, that causes an extend among social groups preferences where the tendency of making a decision is mostly suitable to the individuals well-being. This can be then measured as the suggestions provided by the contextual information provided among this social groups.

The third step before final judgment looks at past experiences. Here an individual is believed that makes his decisions summing up all the above procedures to his past experience and then makes the final judgement. At this level it is believed that the user will have a high hypothesis of making a positive decision if his past experiences where positive. The past experiences of the other individuals sharing the same environment is also put into stack as it gives an influential factor specially if previous experience between them was positive.

This decision making steps or procedures can be vital to judge what are the exact events to recommend to users based on all of this contextual information. In my research I will view also such content to come up with the conclusions on why certain decisions where taken by the users in meetup.com in context of recommender systems.

This work consists of the following chapters: Chapter two are the related works where I have looked at works in both domains to illustrate my objective, in this same chapter we look at the different recommendation theories to present how useful a given recommendation was, we will look at works related to social influence theories and how could such theories be seen in social networks. In chapter three I will introduce the methodology used in this research and the data collection, this includes the research procedures such as data collection methodology, data analysis. Chapter four presents the results and discussions and future works and in chapter five I will present the conclusion to this work.

2. Related Work

Recommender system research have very solid ground, that comes from the very first research done in the 90s by W. Hill et al.1995[2] using collaborative filters to evaluate user's choices in a virtual community when they receive a given recommendation. Thou there has been a large advance done in recommendation system researches, there is still much of work to be done, as there are many interests in this area for much practical applications that would be able to organize more personalized recommendations that best suits users interest. Today we see many web applications utilizing different recommendation methods to help the users with a rapid content search and decisions on what item to get.

With all this great advances that the industry has made there is still room for improvement as recommendation systems tend to adopt a more effective and applicable services that best adjusts user's needs.

As presented by G. Adomavicius[3] et al., 2005 on his research towards the next generation of recommender systems, he presents the central problem of recommender system in a utility function u that is defined by all the space of users C and items S. So he defines the utility function as being the rating that a set of users C make on the set of items S. this can be better understood in the utility function that measures the usefulness of item s to user c where R is the raking given by a nonnegative integers or real numbers with a given range. So this can be resumed in the function that will maximize the user's utility function where each user $c \in C$ will choose a specific item $s' \in S$. They present this function in more formal way as:

$$\boxed{\forall \mathbf{c} \in \mathbf{C}, \quad \mathbf{s'_c} = \mathbf{arg\,max_{s \in S}\, u(c, s)}}$$

According to them this will be represented by a rating which indicates how a particular user liked a given item. This function can be utilized in the different recommendation systems approach. For example, in content-based recommendation systems the utility function is measured according to the item s assigned by user c to similar $s_i \in S$. This will estimate the effective recommendation to items that are similar to those items that have been already ranked by the user. However, this method presents a problem well-known as the new user problem whereby if the user is new to the system there are no much options to recommend due to the fact that users have to rate a number of items to have an effective recommendation.

In a very different approach collaborative recommendation systems try to predict useful items based on items that were previously rated by other users sharing preferences with a given user.

This can be better understood in the utility function by those $u(c_j, s)$ assigned to items s by $c_j \in C$ who are similar to user c. This is introduced in algorithms that will aggregate all ratings of other users to the same item as expressed in the following formula:

$$\boxed{\boldsymbol{r_{c,s} = aggr_{c' \in C} r_{c',s}}} \qquad [4]$$

Like the in the previous method this method also requires that in order to make more effective recommendations the system has to learn more about the user's preferences from the rating that user often makes. Adding to the problem is that collaborative methods tries to recommend items that were rated by users sharing similarities, but new items that have not yet been rated at all might cause a problem since there is no much latent information attached to it. Several recommendation systems use their utility functions based on a hybrid approach making the combination of the two above methods as introduced by M. Balabanovic and Y. Shoham et al 1997.[5] This method tend to solve this problem by introducing the utility function in a multi dimension matrix combining both methods by adding content based characteristics to collaborative Models. This however, brings to a new style of organizing the recommending process as many researches tend to understand users behaviors and incorporate them into the recommendation pattern.

Shani et al. 2005[6] describes the employment that many e-commerce websites use recommendation methods to improve their revenue. In more advanced approaches that combine both methods, utility functions can be defined in various ways. They however, describe utility as the value that either the user or the system gains from a recommendation. Evaluating recommender utility functions can help the optimization of recommendations to users. For this reason, it's important to evaluate how current systems or methods are impacting the user's decision. This needs not only a technical approach but also a more intensive study in user's behaviors towards recommendation systems. Thus the reason to a better understand in social influence theories.

Paul Lazarsfeld and Elihu Katz in 1955 introduced their research on personal influence[7]. They discovered in their research that certain people were more central and influential than others in a group. This research carried by Lazarsfeld and Katz was later brought into evidence as studies on social networking activity found that social influence is not evenly distributed among cybercitizens, but there a group of people known as opinion leaders that have high influence on the internet just as they are online. This study was carried by Kozinets et al. 2010[8]. They proved by conducting interview to a group of people during the American presidential in an Ohio city and came up with the conclusion that people interact with one another to share information. This sharing of information can influence one another to come up with a final decision.

This is the reason why many scholars believe that by incorporating behavioral studies in the recommendation model or even study user's behaviors in recommendation systems can add a key value to recommendation accuracy. Tomoharu Iwata et al. 2007[9] proposes a model for user purchase behavior in online stores that provide recommendation services. In their research they simple features to model user's interests by combining no effect model, uniform effect model and individual effect model. Out of their research the results showed that estimating the individual recommendation effect is important to predict purchase behavior. However further improvements can be made in their user behavior model by using other features such as demographic information and content information. This type o information can be seen in social network where my research is mostly focused on.

G. Adomavicius et al. 2013[10] explore the impact of recommendations in consumer's preferences. They conduct a laboratory test experiments to explore the effects of system recommendations on preferences. The result of their research provide strong evidence that biased output from recommender systems can bring a significantly influence. In this research they did not cover a key feature which is to understand at which situation and what are the points that present high influence.

Similar to the other works is the research conducted by Augusto et al. 2015[11] conduct a study on event based social networks to use data from such social networks to be incorporated in the collaborative information procedure. They use meetup data to collect and investigate user's decisions based on different types of analysis provided by the event based social network. However, they don't fully use contextual data provided by the system to evaluate seatrain influences and decisions made by the users. In my research I will detail such occurrences to determine the impact recommendation systems have on user's decision making process.

3. Data Collection and methodology

Judging from the behavioral model conducted by many recommendation system researches, I come to conduct my research and studies in events recommendation systems to analyze the influential factors of recommendations on user's decisions. In my studies we will focus on the number of RSVPs ("répondez s'il vous plait" which stands for please answer in English) that each user makes in a given event recommended to them. Like in the other recommendation systems users tend to rate their recommendation by giving a score, I will use the RSVPs as a possible rating that users give to events recommended to him.

With this I then create a hypothesis that when users receive a recommendation prior rated with others RSVP's will have a significantly influence in the user's decision and preference. Such prior rating are given by the RSVP's of mainly Group creators or event creators that serve as the influential leader of the group. Therefore, I come to the following hypothesis:

1. Group Creators RSVP will lead to influence positively other users in the same group to attend a particular event;
2. Group Creators RSVP can lead to influence positively other users in the same group to attend an event created in a different group;
3. Users sharing same groups can use their latent information to provide a more accurate recommendation to users that are new to the group.

Thus is safe to say that extra information supplied by this Event Recommendation systems based on a social networking can improve effectiveness and influence users to make a given decision specially when they are in doubt. In normal recommendation systems its common to have the dataflow presented in the following figure 1:

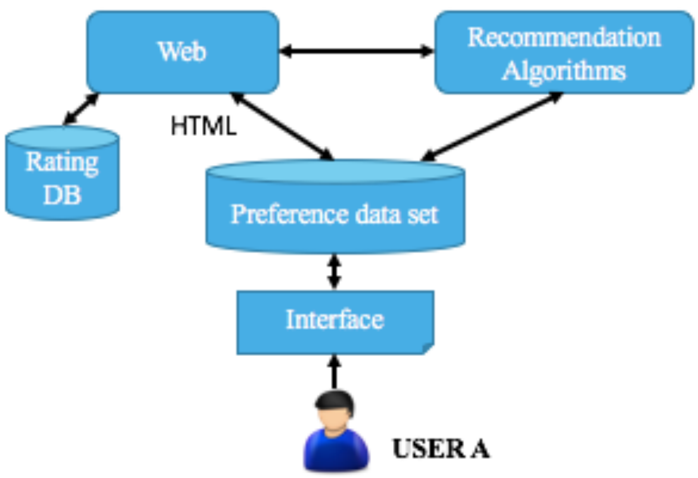


Figure 1: illustration of the recommender system dataflow

Like it was explained before such methodology proved to have problems to new users when they have not rated anything yet. Or another problem comes up if the item is new and no users sharing the same preferences have rated it. This sometimes causes recommendations to have a low effectiveness. Introducing social network contextual information however, might bring a slightly little change. Instead of just relying on the direct user interaction with the system, I believe that extracting contextual information from a user's social network will bring some improvement into the utility function. Our method looks to illustrate exactly that like shown in figure 2.

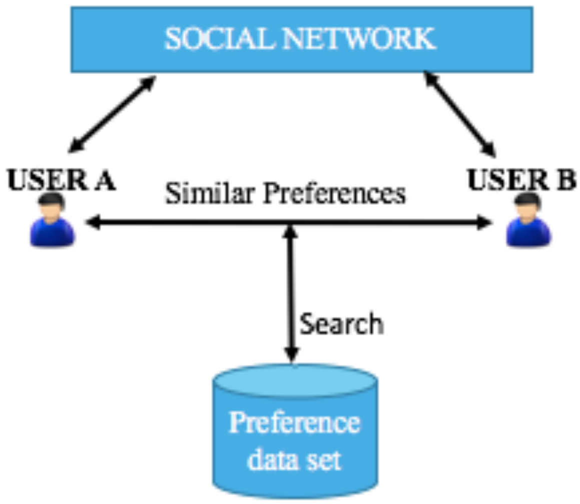


Figure 2: social network based recommendation system data flow.

Social networks tend to give users liberty to express themselves. Is in social network that a user expresses in contextual information like comments, photos, videos how his emotions are or the experience with a given location, event or service. So, social networks are a product of a large amount of data that could play as important to evaluate future predictions to recommender system users. For this reason, I am conducting my research more related to an event based social network to be able to judge the capabilities of such information given by the social networks.

In this chapter I will be explaining every detail of the data collection process and resume the main characteristics and reasons that made me choose the data that I present. My study focuses on event recommendation systems. There are many systems based on events such as the events recommendations offered by Facebook as illustrated on figure 3, or the event based social network www.plancast.com and booking agencies like booking.com, I have decided to use the meetup.com event recommendation system. This platform offers a very large portion of event data through their API which is not complex to work with or understand their methodologies to provide such data. For sure the volume of data presented by APIs like the Facebook API is far larger than the meetup data, but there are a lot of limitations in retrieving the data from such source.

Figure 3: Facebook event recommendation page

However, Meetup platform is a very popular event recommendation social network that contains also a significant data sample to be analyzed on my work and quantify it to better understand the social interaction shown on this platform. Another reason why I chose the Meetup platform is that it solidifies the groups and attendance among different groups in different regions of the world. With this there is a big advantage in our studies since we can easily examine the sample data in different ways such:

1. Group leadership influence;
2. Distance influence;
3. Social activities influence;
4. Responses to attendance;
5. And contextual information;

This are the key factors which we solidify our basis to understand the social influence that such recommendations systems can have on the user's decision making. With Meetup we can understand and have a wide view on the decision process of an individual as it was explained in 1.2.3.

3.1. Meetup platform

As it was explained on the previous section of this chapter, meetup.com is an event recommendation social network where millions of users share their preferences and build groups organizing number of events. The Meetup application gives a very good and solid API that allows programmers to develop their own applications within their own service to better control data of their groups.

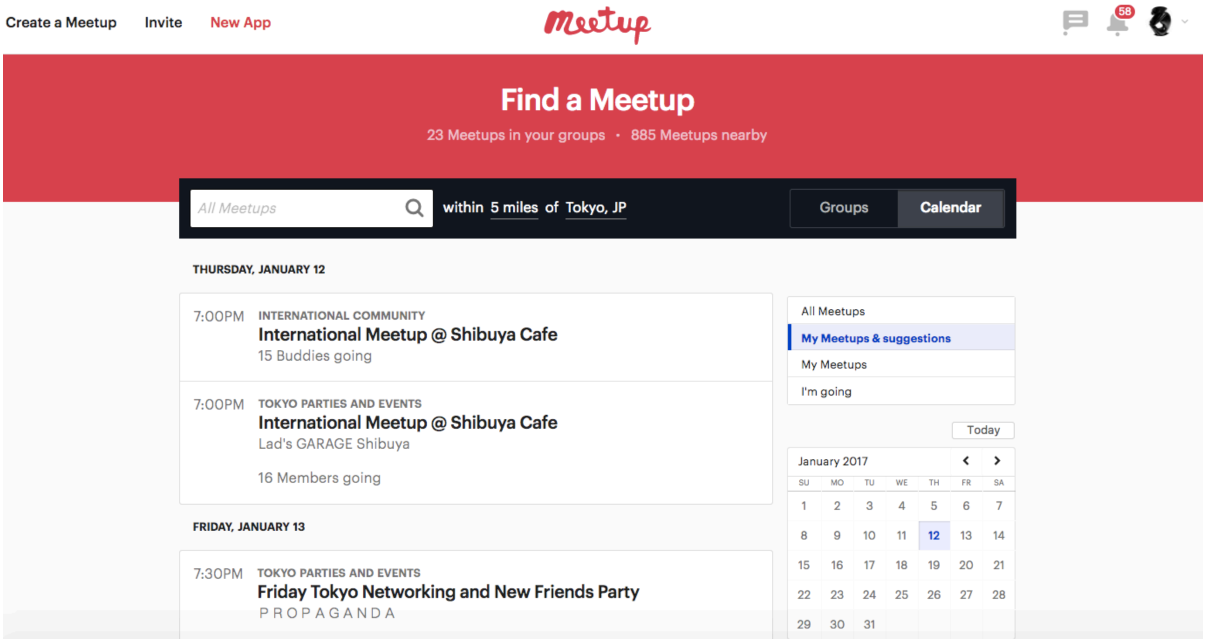


Figure 4: meetup.com event based recommendation system.

As part of this research I utilized the meetup.com API methods to achieve the data analysis. In figure 4 shows the meetup recommendation API that was used to retrieve the sample data for my studies. Meetup contains a number of methods that allows to retrieve information in JSON format and later on be transformed to any data type.

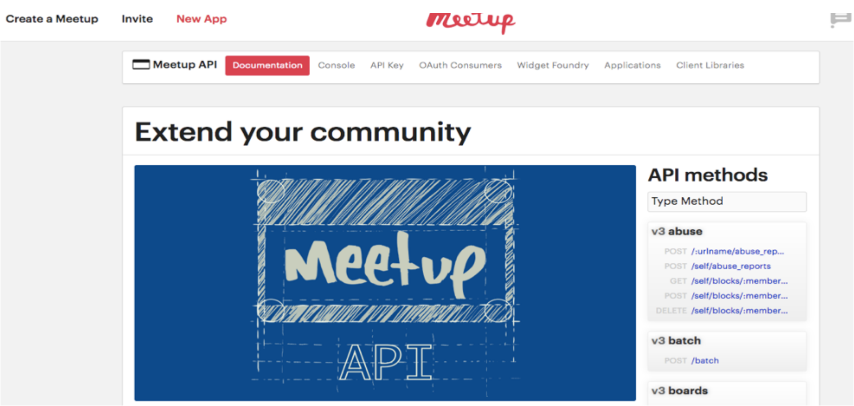


Figure 5: meetup.com API and its different methods;

For my research procedure I wanted to carry my studies on user's decisions making in Japan. So, looking at the various methods provided by the meetup API, my first goal was to understand the demographic distribution of meetups Japanese population.

By applying the meetup GET Cities methods I have conducted the following approach to retrieve the number of population in each city in Japan:

GetCities(c*ountry*, *distance*, *lat*, *lon*, *query*, *radius*, *size*, *smart*, *state*)

This method would return the meetup cities in the respectful query conducted on the meetup API. So for it to return the cities in Japan with their respectful population I used the following query:

GetCities(jp, , , ,zip , *200*, , , *all*)

The same query can be made as illustrated in figure 5, where it basically allows us to work directly on the meetup API console:

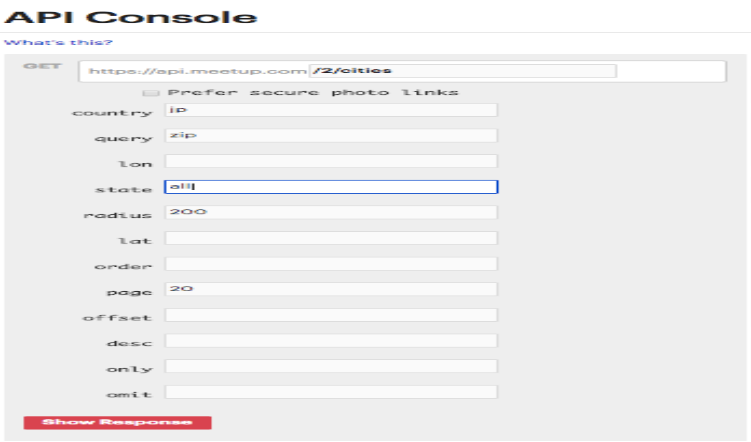


Figure 6 meetup API console

Period of Collection:

The period of collection for the data set in meet started in mid-October 2016 to January 2017. During this period of that collection I was constantly doing the extraction upgrade just to verify if any of the sets had changed from time to time. So many of the data being described here is from November of 2016, thou there is no much difference with last checkup done early January 2017.

The query conducted by the above method then returns the results in JSON format including the query URL that can be used in any programming language to convert the code into raw data.


```
Request URL
https://api.meetup.com/2/cities?&sign=true&photo-host=public&country=jp&page=20
Signed URL
What's this?
https://api.meetup.com/2/cities?country=jp&offset=0&format=json&photo-host=public&page=20&radius=50&order=size&desc=false&sig_id=217366220&sig=83b62683aa54127921788e92dba775f67bd90ec7

HTTP/1.1 200 success {
"results": [
{
        "zip": "meetup1",
        "country": "jp",
        "localized_country_name": "Japan",
        "distance": 34.56030266931076,
        "city": "Tokyo",
        "lon": 139.77000427246094,
        "ranking": 0,
        "id": 1023444,
        "member_count": 7706,
        "lat": 35.66999816894531
},
        {
        "zip": "meetup3",
        "country": "jp",
        "localized_country_name": "Japan",
        "distance": 279.69382990458075,
        "city": "Osaka",
        "lon": 135.5,
        "ranking": 1,
        "id": 1023446
```


The large JSON code is then converted to a csv file where all the raw data requested on the query is given and thus coming with the following results illustrated in table 1. With the results illustrated bellow it's clear that meetup popularity is not as big as in the USA showed on the research conducted by Augusto Q. de Macedo et al.[11] clearly showing the difference between America's top meet population having 719,011 users in Chicago compare to Japans most popular city being Tokyo. However, this population is sufficient to be represented in behavioral studies. So, we conduct the experiments in the two Japanese top cities Tokyo and Osaka.

Table 1: Japan's meetup top cities Demographic;

zip	Country	City	Lon.	Lat.	Ranking	Members
meetup1	Japan	Tokyo	139	35	0	7706
meetup3	Japan	Osaka	135	34	1	451
meetup184	Japan	Okinawa	127	26	2	342
meetup2	Japan	Yokohama	139	35	3	302
meetup4	Japan	Nagoya	136	35	4	295
meetup39	Japan	Yokosuka	139	35	5	123
meetup10	Japan	Kawasaki	139	35	6	111
meetup14	Japan	Chiba	140	35	7	107
meetup516	Japan	Abashiri	144	44	8	77
meetup8	Japan	Kyoto	135	35	9	67

From the results retrieved its clear to see that the cities with big population are close to each other and one of the factors might be the tourism attraction in this cities. Cities with big population is mostly common to find the largest amount of events. And where there are many events there is a high probability that influence among users can be high.

This study would be much complicated if the approach was to be conducted in all top 10 cities of japan. For this reason, I have minimized the data collection into the two main cities Tokyo and Osaka. Within these two cities there are thousands of groups under different categories, so working with all the groups would increase the sparsity of an efficient result in the end. So I decided to conduct my research on simply three meetup group categories in both of the cities:

1. Music;
2. Sports;
3. Camping;

I will present the methods used to collect the data given by the API. For each method I am looking to work with the groups having higher members and events with the highest RSVP. We will see the inflation of some members in groups different from their preferences and investigate the possible reasons that they attended events from groups completely different from their preferences. With the experiment, we can conduct our quantitative data analysis to present (1) Group leadership influence; (2) Distance influence; (3) Social activities influence; (4) Responses to attendance; (5) And contextual information;

3.2. Data collection methods

As mentioned above, the meetup API contains many methods to retrieve the desired dataset. In this section I describe the different methods that I used to collect the data needed for our final evaluation.

3.2.1. Get group method

After having the demographic data retrieved from the API we use the get group methodology also offered by the meetup API to be able to extract the group information. The goal of this method is to retrieve the groups in their different categories and be able to compare the density population among the groups and evaluate the popularity among the topics introduced. Bellow we present the method used by the API as well as part of its JSON code:

```
GetGroups (category_id, city, country, domain, fields, group_id, group_urlname, groupnum, id, lat, location, lon, member_id, members, name, organizer_id, radius, state, topic, zip)
```

The above method will bring forth all the information presented on a given group under a specific category. For our groups case study, we are looking at three categories music, sports and camping. The same method was used to retrieve all three categories on the two main cities. After conducting the query, the meetup proceeds by giving us JSON code where it supplies us with all the group classes and data on these categories. The example bellow show, the JSON code after requesting for sports category in the city of Tokyo. This query was conducted on a radius of 50 miles from the center of Tokyo thus presenting the following result:

Resuming this we have the following resumed information in table two comparing the results of the three categories in both cities:

Table 2: Statistics related to the groups belonging to the tree different categories in both Tokyo and Osaka

City	Category	Number of groups	Number of members
Tokyo	Music	20	11078
	Sports	20	19956
	Camping	14	5250
Osaka	Music	20	5721
	Sports	8	3134
	Camping	4	776

The goal of this first methodology is to perform a primary test on the location influence. In the next chapter where we argue the results we will look at the difference between the number, total number of users in this two cities in relationship to the total amount of users in the groups belonging to this two cities, bring the hypothesis that there are users belonging to more than one category thus having a high hypothesis of user's infiltration due to some possible influential factor caused by the recommendation system.

Other than that, there is the possible proof of the social influential theory presented in chapter 1 when it was stated that possible group leaders might convince the lower class of a given group. The decision process steps have also illustrated that during an individual making a process to make up his mind, there is a high factor of his social influence, that the table briefly illustrates the possibility of members of one group with a total different preference being influenced by recommendations done by the system, fruit of various factors that are used as latent information.

3.2.2. Get members method

From the get members method, I was able to retrieve all the information needed about the members belonging to a specific group. At first I had the choosing criteria from the previous method. By executing the get group method, I was able to look at the details presented by each group such as total number of members in a group, the ratings given to the group, the organizing members etc.

Based on all of this information the criteria to select the group to be carried the experiments on, I chose the groups with the highest number of members and with the highest ratings. Its frequent that groups with high density of population have more activities and the more activities are created the more latent information is produced. And is by seeking more latent information that we can analyze the factors based on user's decision to attend or not a given event recommend by him. Should the recommendation process carry enough information the possibilities of helping the user make his decisions are high.

Once again we use the meetup API method to be able retrieve or extract the sufficient information about the users belonging to the groups. This method however, has many similar methods. So using the API documentation [14] I chose the method that would best represent the data that was needed. Instead of selecting the Get profile method of conducting the query I choose the Get members because it would time consuming to analyze the profile of each member belonging to a given group, when what we need is the basic interactions between the members preferences supplied by the above method and engage with the RSVP method to look at which members have attended which type of events.

So the best way to query the information of the members belonging to a given meetup group is presented below:

GetMembers(***fields***, ***group_id***, ***group_urlname***, ***groupnum***, ***joined***, ***member_id***, ***name***, ***service***, ***topic***, ***visited***)

Thou this method presents many parameters to carry the query the meetup API console only requires to fill in one of the key parameters (group_id, group_urlname, groupnum). We then insert into the query the topic or category and group id with the highest members and ratings. This return as response the JSON code with the users belonging to the group as shown below.


```
HTTP/1.1 200 success {
"results": [
{
"country": "jp",
"city": "Osaka",
"topics": [
{
"urlkey": "baseball",
"name": "Baseball",
"id": 80
},
{
"urlkey": "salsa",
"name": "Salsa",
"id": 1122
},
{
"urlkey": "art",
"name": "Art",
"id": 1502
},
{
"urlkey": "travel",
"name": "Travel",
"id": 1998
}],
"joined": 1450913464000,
"link": "http://www.meetup.com/members/185809784",
"photo": {
"highres_link":
"http://photos1.meetupstatic.com/photos/member/4/3/3/4/highres_245057204.jpeg",
"photo_id": 245057204,
"base_url": "http://photos1.meetupstatic.com", …
```

The method extracts successfully all the information containing in the user's profile. Thou there is a vast content on the users profile I resumed the key contents in the user's table to be analyzed on his events preferences, and with it come to the conclusion of the decisions made. Below is table 3 containing the basic information of the user from the music category in Osaka needed for analysis.

Table 3: illustration of user's information extracted from the meetup recommendation system

country	hometown	city	joined	lon	id	lat	status
Jp	USA	Osaka	1405040000000	135.5	48471002	34.68	active
Jp	Dubai, AE	Ashikaga	1370010000000	139.45	94877362	36.34	active
jp		Kobe	1409150000000	135.17	128774452	34.68	active
jp		Osaka	1447320000000	135.5	185809784	34.68	active
jp		Kyoto	1429880000000	135.75	186190185	35.01	active
jp		Osaka	1396670000000	135.5	140642772	34.64	active
jp		Nara	1382360000000	135.83	109955932	34.69	active
jp	Malaysia	Osaka	1439710000000	135.5	36394942	34.68	active
jp		Osaka	1471670000000	135.5	208032767	34.68	active
us	San Diego	San Diego	1475490000000	-117.22	60693832	32.95	active
jp		Osaka	1443840000000	135.5	185200953	34.68	active
jp	San Jose	Osaka	1444630000000	135.5	177678532	34.68	active
ca		MontrÌ©al	1378780000000	-73.57	110429672	45.49	active
jp		Kobe	1430010000000	135.17	184381635	34.68	active
jp		Kyotanabe	1451970000000	135.39	197617144	33.73	active
jp	Osaka	Osaka	1436340000000	135.5	135215422	34.68	active
jp		Nishinomiya	1446730000000	135.34	190817297	34.73	active
jp		Osaka	1432310000000	135.5	182919135	34.68	active

3.2.3. Get Event method

The recommendation process that we are looking ate involves events. So it is important to filter the events that were recommended to the group categories selected above. It is known that user's belonging to a group when they receive a recommendation they interact with system by presenting possible RSVPs being it yes, no or maybe. Conducting this research, I want to analyze specifically the positive RSVPs to be able to know who are the members that are possibly attending an event. The Get Event method, is one of the methods with most sub-methods. I therefore have select the Get/2/events where we can get the description of the events attended by the group elements and other elements from other groups.
With this information we will be able to look at the possible reason why people attended the events recommended to them. After extracting the data, I used the same requirements to select the best events. I went and selected the events with higher number of participants and with a high rating. This events where selected in the timeline from September 2016 to January 2017. The objective of selecting this time was to use the best contextual and recent data made available by the meetup API. After conducting the query, the following results were presented as illustrated in the table 4, showing the total number of events in the same category. The query was conducted to present the first 200 events during the stipulated time:

Table 4: Number of events recommended by each category

Cities	Category	Total Events
Tokyo	Music	200
	Sports	200
	Camping	200
Osaka	Music	53
	Sports	200
	Camping	17

3.2.4. Get RSVP method

The final method utilized to retrieve or extract information from the meetup API and come up with my assumptions, is the Get RSVP method. Like the one before, there are many Get RSVP methods. Now, I want to use the Get RSVP method that would present me with the information related to the selected event. According to the explanations presented on the previous section I would like to present the resume of the dataset collected with the event method presenting the comparison between the number of people belonging to the group in relation to the total number of RSVP. Table 5 presents all this data explaining the data extracted using this method:

Table 5 RSVP data per group

Cities	Category	Total Events	Total members in the group	Total Positive RSVP
Tokyo	Music	200	1006	70
	Sports	200	1871	15
	Camping	200	2158	11
Osaka	Music	53	885	14
	Sports	200	2205	79
	Camping	17	477	18

After retrieving and structuring all this data I then started analyzing them according to the different influential factors and then present the results of such analysis in relation to my hypothesis and giving an answer to the research question.

3.3. Data Analysis

After the long process of extract retrieving, converting and interpreting the data, it's time to present what exactly is this data showing to us. I start by analyzing the demographic distribution between the actual population presented in the city and the population distributed in the groups. It's obvious that there is a big misrepresentation represented in figure 4 about the population.

In this specific details the recommendations given to a new user could result in a very poor recommendation influence. Supposing that the user is new to the system there is a little information about him and his/her preferences are related to group members far away from him. The other users might be linked to a given group in a different city and that influence the recommendation to send notification about an event that eventually is far away from him.

The usage of social network therefore comes and incorporates more contextual information as seen in the case of the meetup event base social network. In this case even if the user being new to the system the old users have already interacted with different events thus providing more illustrations or options to the recommender process due to its richness in latent information.

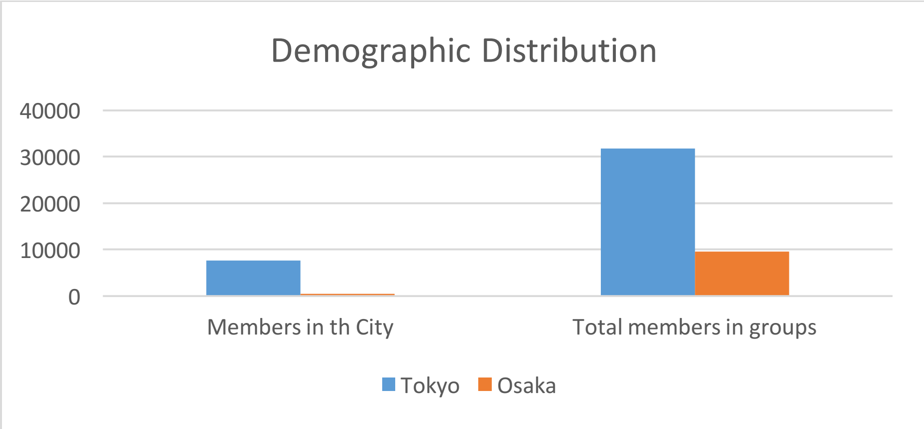


Figure 7 demographic distribution of the members in the two cities.

With figure 6, I am also cable to look that there are various hypothesis which could come to agree with my assumptions.

1. Knowing that Tokyo has the total demographic of 7706 meetup members, all the 31810 members represented in the sum of the groups represented on the different categories in my research can as well be:

1.1. Some members belonging to other nearby cities;

1.2. Some members belonging to other groups within Tokyo;

1.3. This are members from different group preferences influenced by the recommendations sent to them;

2. Tokyo receives many tourists and has a very high foreign demographic so this could possibly:

2.1. This could members that once lived in Japan and are no longer in Japan;

2.2. This could members that are linked with Japanese meetup groups to attend highly recommended events;

The above assumptions are not only applied to the city of Tokyo but could also be applied to the city of Osaka or other major cities in general. To come up with an answer to the above assumptions I performed several experiments to be able to filter this information and clarify the impact that these recommendations have on the users.

For this specific reason I made a graphic representation of the amount of meetup members belonging to the group categories chosen on this research are actually from Tokyo and Osaka and how many are from other cities or rest of the world. Luckily, the meetup data provides us with a hometown variable presented by the user when he is stating his origin.

	A	B	C	
1	country	hometown	city	jc
2	jp	USA	Osaka	
3	jp	Dubai, AE	Ashikaga	
4	jp		Kobe	
5	jp		Osaka	
6	jp		Kyoto	
7	jp		Osaka	
8	jp		Nara	
9	jp	Malaysia	Osaka	
10	jp		Osaka	
11	us	San Diego	San Diego	
12	jp		Osaka	

Figure 8: sample data from meetup extracted in cvs files showing details of users location and origin.

With this variable I will be able to examine the infiltrated members by applying the following column comparison to search if they belong to other cities. The cities name or their id data can be used for this matter. By implementing this I came to the following analysis using a sample population of less than 200 people from the selected group:

Table 6 shows the number of outsiders in different categories on a range of 200 people

City	Category	Number of outsiders	Number of citizens
Tokyo	Music	61	139
	Sports	48	152
	Camping	60	140
Osaka	Music	106	94
	Sports	84	116
	Camping	105	95

According to this statistic it’s safe to say that in Tokyo over 28% of the population taking part of their groups are being influenced by this possible factors:

1. The events taking place above attracts people due to their high rating and information spread;
2. Most of the surrounding cities have very limited events or groups;
3. The recommendation sent to these users have similar preferences thou it’s far away;

Looking at the Osaka demographic it’s safe to say that 49% of the population taking part of their groups are being influenced by the following factors:

2. Osaka is the biggest city among less popular meetup cities in Japan, therefore many of its groups attract people from near smaller cities;
3. Looking at the latent data about locations where events are normally held distance is not a big discouragement to users wanting to experience a high rated event;

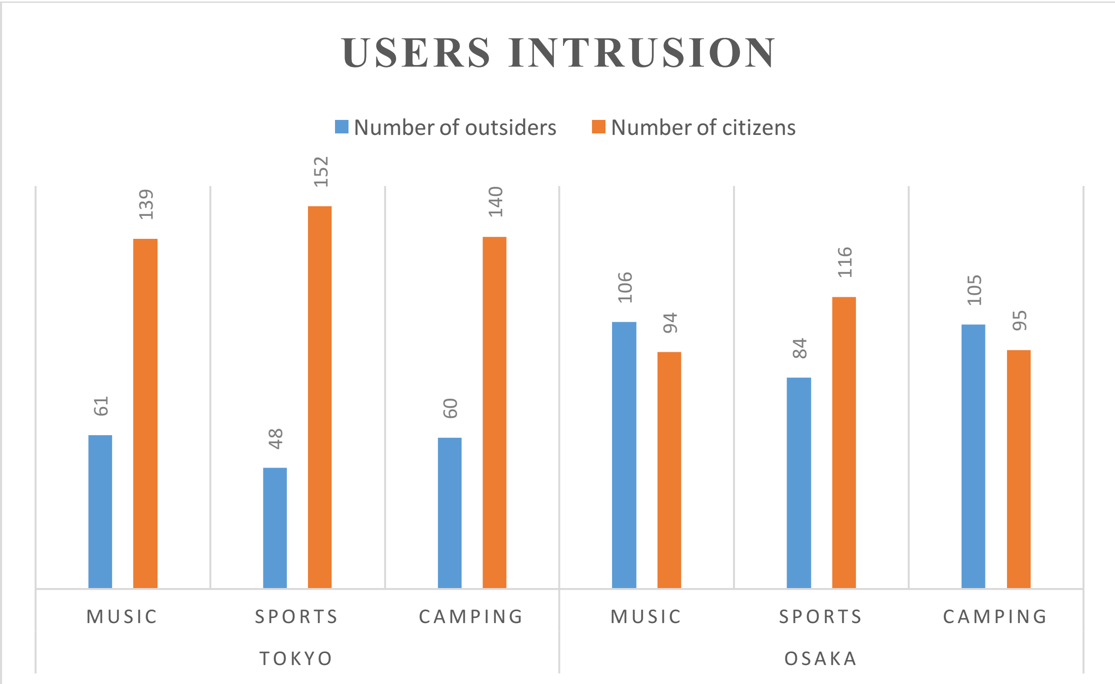


Figure 9: users intrusion demographic in relation to groups

We now have the proof that there is a significant number of users from other cities being influenced in making their decisions to attend events in bigger cities with a high population density.

4. Results and discussions

After having all the analyzed it's time to look at the hypothesis presented in the beginning of my work. At first we will look at the results of user distribution in relation to their preferences. Chosen before where the three categories in the meetup recommendation system. I decided to go with the three in the chosen cities. We saw the illustration in figure 7 describing the number of intrusion in the various groups. This has made us to retrieve the information in the major events organized by each group to be able to measure intrusion in line with the groups.

In order to achieve this results knowing that each group have a large number of people, it would take a long time to analyze user per user the events they attend and their RSVP. RSVP data are important because it brings a possible rating to event that might influence others to attend a given result. Meetup however provides descriptions and ratings to events so it's an extra latent information that could be used as well. But in this research RSVP analysis was used to be able to follow and track such feeds given by the users if they will attend the event or not. If the RSVP is positive and user ranked the event we then classified as an event attend by the user. Such information can be used to store his background experience for similar upcoming recommendations and such that can be used both in the collaborative method to recommend similar upcoming events to users that are part of his meetup groups or share similarities in their preferences.

To achieve this analysis however on which users are not part of a given group organizing the event I had to do a user ID comparison match between the data of those attending the event in comparison to the regular users of the group. I then performed a query on the list of members attending the event but their user IDs don't belong to list of regular group members. The query makes a search a comparison and returns NA (Not Available). With this query I was able to retrieve the number of users belonging to the group and those that are just visiting or where interested in the event organization. Achieving this data extraction we could come to understand better the reasons why the RSVP data was provided by a user not belonging to the group.

4.1. RSVP Results on music category

We present the results bellow taken from the experiments carried to see the number of users belonging to other groups but have attend the events in other groups. Bellow we present the RSVP statistics used to check the number of intrusion in each category and groups in particular.

Music			
RSVP Statistics Tokyo			
Total Participation	Group Members	Non Members	Leader Participation
78	63	15	yes
RSVP Statistics Osaka			
Total Participation	Group Members	Non Members	Leader Participation
18	9	9	yes

Table 7: representation of infiltration in the music group

The results above represent a significantly high intrusion rate of 19% for the Tokyo music group and the 50% in the Osaka group. The results presented are qualified to be high because according to behavioral science scholars like E.M Rogers stated that it takes a great effort to influence a user out of his usual social well-being. So having the possibility of having at most one user from other groups accepting your recommendation can be rated has high. From the results presented we can clearly see that the participation of the group leader in this event might have also been a influential factor for both users from this group and from other groups who have joined in. When looking at the data itself, I could relation the user's location as well as most of the participants in this specific event are near the events location. The query was done for a radius of 50 miles so there is also a very high chance that location serves as an important key factor to influence a particular group of people.

So analyzing this results we could say that leadership and location have an important role in users decision making process. However, the data was not clear enough to match the results in relation to their time in using the recommendation system. However, one of the two factors was probably used to recommend the event to users. Later in this section we will evaluate this factors in utility function to present possible key features that recommendation systems could use from social networks to improve their recommendation scheme.

4.2. RSVP Results on Sports Category

Next are the results of the sports groups also in relation to RSVPs given to the event. Like the analysis done before the same is applied on this group category.

Sports			
RSVP Statistics Tokyo			
Total Participation	Group Members	Non Members	Leader Participation
15	15	0	yes
RSVP Statistics Osaka			
Total Participation	Group Members	Non Members	Leader Participation
54	22	32	yes

Table 8: Representation of infiltrated members in the sports group;

The results presented by the sports event was a little different compared to the previous groups. The RSVP data from the sports group from Tokyo showed no infiltration in this event. Demographically also it's a smaller event compare to event of Osaka. This could be one the reasons this event has a low infiltration rate thus is safe to say that there was a low influential factor attached by eat. When investigating the data supplied by the meetup API I discovered that there was little latent information such as comments or past similar events related to it. So this enlightens me by discovering that the higher the information is shared the higher the influence can be. As seen in the social influence theory in previous sections, it was described that one of the factor that impacts users for a influenced decision is the high awareness of social communication so sharing information in the such types of systems could as well elevate the impact future recommendation can have on a particular user. Osaka however show a completely different infiltration rate from Tokyo as far as the sports group is concern by having the rate of 59% which shows the high influence it had on both members and non-members of the group. Analyzing the characteristics of the group there was more social awareness and exchange of information proving again that such information is a key factor for future recommendations. Lastly we look at the results related to the camping group.

4.3. RSVP results camping category

The results for Tokyo showed a result of 11% of the infiltration rate which once again according to the theory of social influence it can be considered as positive. Thou it's not very it can still be considered as a good value.

Camping			
RSVP Statistics Tokyo			
Total Participation	Group Members	Non Members	Leader participation
18	16	2	yes
RSVP statistics Osaka			
Total Participation	Group Members	Non Members	Leader participation
22	9	13	no

In Osaka showed a very rare value where by the group leader did not attend the event but the infiltration rate was very high with the rate of 65%. This could probably of the high ratings of past events since the group has very high ratings. In order for us to understand well how the impact of this recommendation systems will work we need to understand the different recommendation methods or approaches and I have applied them to my data set.

This results however could be argued if they indeed are of relevant significance. We have seen the amount of users this groups present. But in regards to the events the attendance is low compared to the high number of group members. One could argue that the numbers presented are of high significance since is not easy to gather thousands of people in a particular event. Thus is safe to say that recommendation systems do play their importance in impacting users decisions. But what do this numbers mean for the utility function? One of my research objectives is to prove that there will be an improvement in the utility function of the different recommendation systems methods if they incorporate social network information.

Based on the results presented I have concluded that recommendation systems do have an impact in user's decision making. But there is a need to understand what were the key factors which brought such impact in their decision making process. Based on the results I pointed the leadership role in influencing those sharing preferences with him. But there is a need to transport this a more practical and technical field to evaluate the utility of the recommendation.

4.4. Utility function in recommendation methods

One of the objectives of this research is to find answers if the recommendations have impact or influence in the user's decision making process by evaluating how useful are the recommendations based on the data set we have from meetup. I will start evaluating such impact by looking at how recommendations are done the behavior of their utility functions in the different recommendation methods.

4.4.1. Content based recommendation methods

The main idea behind the content based methodology is to recommend an event to a given user x similar to previous events rated highly by the same user. In this case specifically for meetup we could evaluate such rating by the RSVP given by the users or the ratings itself that some users do.

This can be done to events hosted in the same location, same organizer or having the same conditions. To apply this experiment, I have looked at a set of events a specific user has liked, expressing it either by RSVPs or rating which could be described either by implicit or explicit data. By doing this there was a need to look at the description of a set of events that a given user has given a RSVP and has rated. This process will help me know that the user has gone to this particular event and for each of the events I have built an item profile which is very common on content based recommendation systems. The item profile (event profile) will contain the description of the event.

For this case I will look at what are the characteristics of the user's preferences for example, what are the group categories the user is or what are normal locations of events he attends etc. Once the profile is created a match can be done to see what are the events that are more suitable for him. The events put in the catalogue (event profile) that best matches user's preferences ate then recommended as illustrated in figure 10 below:

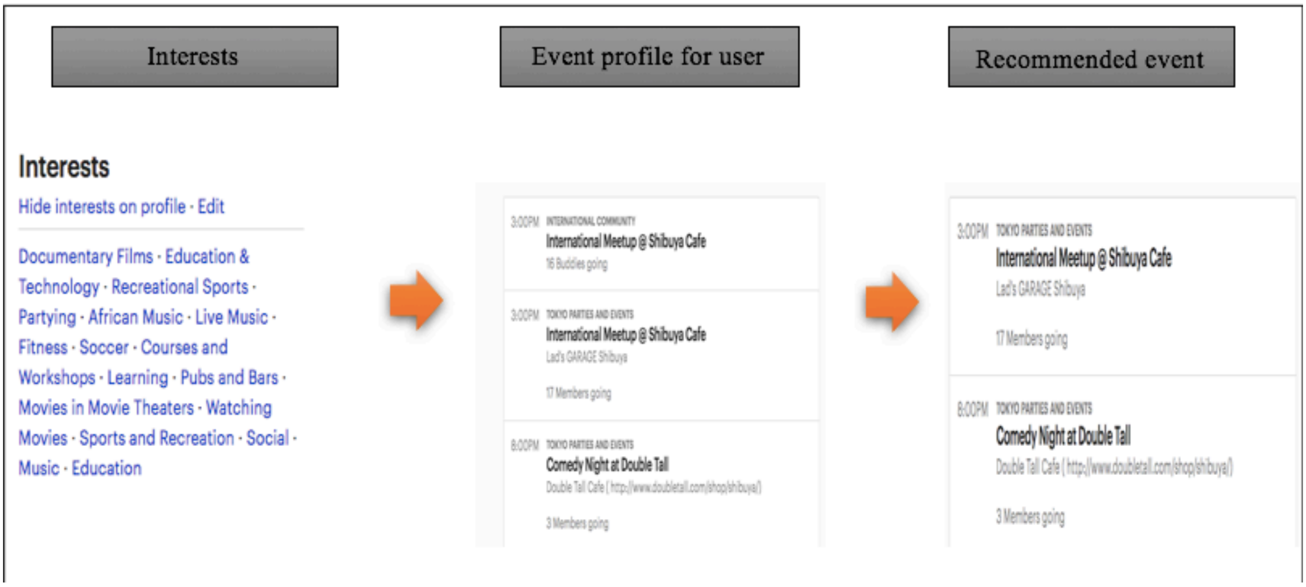


Figure 10: Content based recommendation process

4.4.1.1. Building event profile

For each of the user the content base approach builds an event profile so that it can be used to create the user profile. The following was then created on the experimental process:

- Knowing that a profile is a set of features, it can be presented as a vector. This vector can be Boolean or real value vectors, containing one entry for the feature of the event. Taking from the experiments we are looking at event description such as category that this event belongs or the location the event will be held or even the details that the group organizer inserts. So the event profile will be equal to a set of words in the event description. To pick this important words for words for the event profile I used what is common which is the heuristic from the text mining TF-IDF formally known as Term Frequency by Inverse Document Frequency. In this order the following was done:

To

f_{ij} = frequency of terms *i in* Description *j*

$$TF_{ij} = \frac{f_{ij}}{max_k f_{kj}}$$

n_i = number of description that mention term *i*

N= total number of events

$IDF_i = Log_{n_i}^{N}$

TF-IDF score: $w_{ij} = TF_{ij} \times IDF_i$

Event Profile = Set of words with highest TF-IDF scores

illustrate all this formulas, I show it in detail the results at the table below after all the computation process when searching for the word music in the music category events in the city of Osaka:

Table 9: Term Frequency and Inverse Document Frequency

Event ID	Event Description	Terms appearance	Term Frequency	IDF	TF-IDF
101625322	Hi everyone, Happy New Year! I hope all of you are doing good and looking forward to meetup in 2013. …	3	$\frac{3}{190}$	1.23	0.018
100955042	Hi Everyone, Spring is around the corner, and I just can't wait the Ohanami season! …	1	$\frac{1}{169}$	1.71	0.010
79742462	Would you like to stop by a free concert of classic music after the office hour? For those of you who would not be able to come on Thursday, I also recommend you to check out the schedule of other performances as there are a lot of another chances to hear good music! …	5	$\frac{5}{137}$	1	0.036

Taking into account that there are 52 events in this category in the city of Osaka, we are calculating the term frequency in each of the events description html page given by the meetup

API. So this I consider every description set as being a document. The results show that the term music appears in some of the descriptions, taking it into account that in the description there is a limitation of 500 characters group organizers tend limit themselves in describing the events which clearly affects the TF-IDF which might also cause a bad result in impacting the event profile which will be matching with the user profile. So if there is a poor description of the event there might be a problem in the recommendation algorithm. I noticed also that there are some events with 0 TF-IDF but have very high number of RSVP. In this case it is safe to say that if recommendation methodology simply depends on TF-IDF like in the case of content base recommendation systems the utility function will perform very poor. This is because the organizer is limited to set his description which will affect the rating vector. A way to solve this problem is later described with the hybrid method. After finding the TF-IDF its possible to create the user profile and eventually recommend the event.

4.4.1.2. Building user profile

After Creating the event profile there is a need of making a user profile to match both of them and find the right event that is mostly useful to the user in regarding to his preferences. So if we have a user that has given his RSVP or has rated the event profiles $i_{1\ldots}i_{2\ldots}i_n$ where this are part of the vector with n entries in a high dimensional space. The simple way of constructing the user profile from a set of number of event profiles, is just to average the event profile where N is the total amount of event profiles. The variant to this is to normalize weights of terms using average rating of the users.

So by creating a Boolean utility matrix we will look at a set of events that the user has attended has illustrated in figure 11, where all the values seen as 1 are the events users the users gave a positive RSVP and 0 (zero) are the events represented with a negative RSVP in music event chosen by its popularity.

Member ID	RSVP EVENT 1	RSVP EVENT 2	RSVP EVENT 3	RSVP EVENT 4	RSVP EVENT 5	RSVP EVENT 6	RSVP EVENT 7	RSVP EVENT 8	RSVP EVENT 9	RSVP EVENT 10
131080882	1	1	0	1	1	0	0	1	0	1
76189152	1	0	1	1	0	1	0	1	1	1
32399582	1	0	0	1	0	1	0	1	1	1
127412922	1	0	0	1	0	0	1	0	0	0
130735922	0	0	0	1	1	0	0	1	0	0
130243682	0	0	0	1	0	0	0	1	0	0
58824902	1	0	0	1	1	0	0	1	0	0
123884272	1	0	0	1	0	0	0	0	0	1
104795912	0	0	0	1	1	0	1	0	0	1
98077962	1	1	0	1	1	0	0	1	0	1
122956472	0	0	1	1	1	0	0	1	0	1
101130362	1	0	0	1	0	1	0	1	0	0
86838342	0	0	0	1	1	0	0	1	0	0
5869105	1	0	0	1	1	0	0	1	1	1
51256402	0	1	0	1	1	0	0	1	1	0
38463592	1	0	0	1	0	1	0	1	1	0
24753382	1	1	0	1	1	0	1	1	0	0
95705542	1	1	0	1	1	0	0	1	0	0

Figure 11 rating vector produced by RSVPs of users

So suppose we are creating the user profile of the first user *131080882* and has attend 6 events. Within this events 2 where located in a distance 50 miles from the center of Osaka which I will name it *location_zone_1*. The other 4 events where located simultaneously in a distance of 80 miles from the center of Osaka which I will name it *location_zone_2.* The simplest way to get this user profile is by finding the mean of the profiles.

- The weight of *location_zone_1* $=\frac{2}{7} = 0.28$
- The weight of *location_zone_2*$= \frac{4}{7} = 0.57$

Just for the better analogy I decided to look also of making the same with the ratings of events in case the RSVP data is having a high sparsity has seen in previous section. Figure 12 shows in case I work with rating of events since meetup allows users to give RSVP and rate the events.

Member ID	RSVP EVENT 1	RSVP EVENT 2	RSVP EVENT 3	RSVP EVENT 4	RSVP EVENT 5	RSVP EVENT 6	RSVP EVENT 7	RSVP EVENT 8	RSVP EVENT 9	RSVP EVENT 10
131080882	3	5	0	4	2	1	0	5	0	2
76189152	4	0	4	5	0	4	0	4	3	3
32399582	3	0	0	3	0	1	0	4	2	3
127412922	4	0	0	3	0	0	3	0	0	0
130735922	0	0	0	4	4	0	0	5	0	0
130243682	0	0	0	3	0	0	0	4	0	0
58824902	4	0	0	3	3	0	0	5	0	0
123884272	4	0	0	3	0	0	0	0	0	2
104795912	0	0	0	5	4	0	3	0	0	3
98077962	4	3	0	3.2	2	0	0	2.5	0	4
122956472	0	0	3	4	3	0	0	4	0	3
101130362	4	0	0	3	0	3	0	3	0	0
86838342	0	0	0	3	3	0	0	4	0	0
5869105	3	0	0	3	3	0	0	4	3	3
51256402	0	2	0	2	2	0	0	3	2	0
38463592	4	0	0	3	0	3	0	3	2	0
24753382	3	3	0	3	3	0	3	4	0	0
95705542	5	5	0	4	4	0	0	5	0	0

Figure 12: Rating vector produced by the ratings given by users to different events

In figure 12 we have the users giving by the users that have attend the set of events having the minimum rating 0 and the highest 5. We measure the same weight for each of the location as done in RSVPs vectors.

- Events in *location_zone_1={3;5} ratings*
- Events in *location_zone_2={4;2;1;5;2} ratings*

Making the average of all his rated events I found the average rating of *3,14*. This value is subtracted to the ratings to normalize the values so the values change to the following:

- Events in *location_zone_1= {3-3.14;5-3.14}= {-0.14;1.86}*
- Events in *location_zone_2= {4-3.14;2-3.14;1-3.14;5-3.14;2-3.14}={0.86;-1.14;-2.14;1.86;-1.14}*

After normalizing the weights are as follows

- The weight of *location_zone_1* $=\frac{1.72}{2} = 0.86$
- The weight of *location_zone_2*$= \frac{-1.7}{5} = -0.34$

From the results we can clearly understand that a rate above 2.5 is a positive rate while below it is a negative rating and by normalizing these ratings we can see that the user tends to prefer more events located in location 1 even thou he has only attend two of them. Matching both results from the event profile and the user profile it can be predicted the results for the event to be recommended to user.

4.4.1.3. Recommending events using content base method

In this research I am looking for the impact that recommendations have on users decisions. So based on the user's past experience we have looked at possible items that could be useful to the user by making the event profile and the user profile. Now the key point to recommend possible useful events to user is to pair *user profile(x), event profile(i)* and find out what is the rating of that user and item pair is likely to be. Remembering that both the user profile and item profile are vectors in a high dimensional space. In this case I will show how I find the results in a two dimensional space when in the reality they are embedded in a higher dimensional space. Having vectors in higher dimensional space a good distance metric between the pair of vectors is the angle θ between the pair of vectors as exemplified in figure 13:

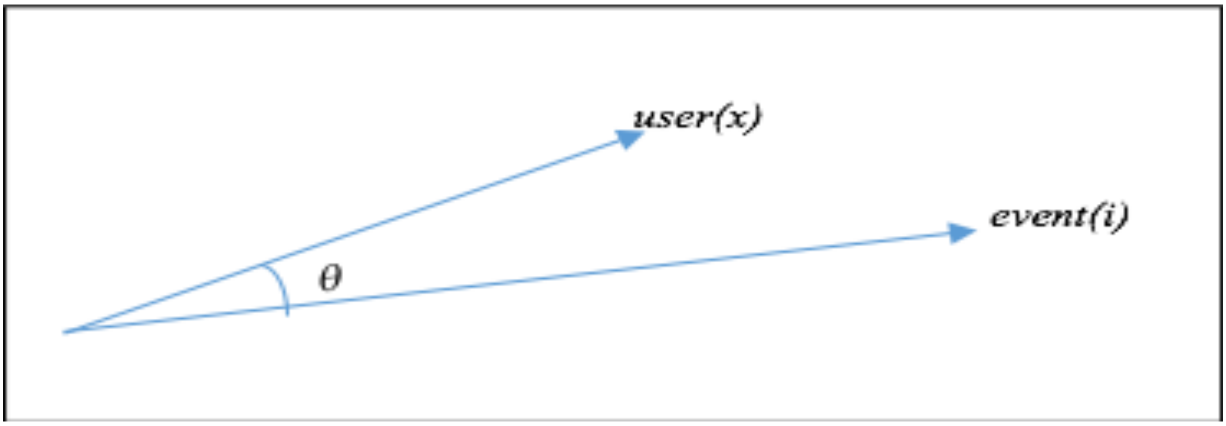


Figure 13: distance of vector in a two dimensional space

By using the cosine function is possible to estimate the angle in the following manner:

$$\boldsymbol{U}(\boldsymbol{x},\boldsymbol{i}) = \boldsymbol{cos}(\theta)\,\frac{(\boldsymbol{x}\cdot\boldsymbol{i})}{(|\boldsymbol{x}|\cdot|\boldsymbol{i}|},$$

the above formula is also known as the cosine similarity between user *x* and the event *i,* the distance in here is the angel θ and not its cosine, rather than cosine similarity is the angle 180 minus θ. In this case the smaller the angle the more similar event *i* and user *x* are, so the subtraction of the cosine similarity and theta (θ) are normally going to larger. In this case as the angle θ becomes smaller the *cos(θ)* becomes larger and vice versa.

So the prediction is made by having user *x,* I calculate the cosine similarity between that user and all the events in the event description, and from there the event with the highest cosine similarity are recommended to the user.

4.4.1.4. Advantages and disadvantages of content based methods

Here I will describe some of the positive points and negative points of this methodology and discuss how they can affect the influence of each recommendation given to an individual.

Advantages are:

- There is no need to have data about other users. This is good for the reason that events can have impact on users that are new to recommendation system.
- This method is able to recommend to users with unique preferences. In this case for users who do not have much preferences similarities from other users the recommendation can still impact the user based on the singular user and item profiles that is generated by this approach.
- This method also brings the possibility of recommending new events that are not so popular. This is because when new items appear there is no need rating to build the event profile. The event profile depends completely on the features and description of events.
- Recommendations can be explained based on the features of the event. When a given event is recommended it comes with an explanation on why the specific event was recommended to the user. This brings a good impact since user is reminded of his actions. This could also benefit the user to make better choices in the future actions to be taken by him.

Disadvantages are:

- Finding features in this approach is very hard. This approach has very solid basis in information retrieval, so there are problems generated in this area of research such as good description of items such as images, videos etc.
- Overspecialization is another problem in this methods. The user profile is created by the ratings that the user has made in the past. If this is not given its impossible to recommend since this only recommends events inside user's profiles. for the case of the data set that I have used RSVPs can be very sparse, creating an even cold start, because user can present his RSVP but not attend eventually the event. This prevents efficient quality in the judgement of the other users.

The cold start problem could generate difficulties to create a user profile for a new user. This is also known as the new user problem, where those that are new to the system might experience very few impact compared to old users who are familiar to the system.

4.4.2. Collaborative filtering method

After looking at the impact that content based methods can do to users, it's now time to have a look at the impact that collaborative filtering methods can have on the on user's decision making process.

The idea of collaborative filtering is suppose we have a user *x* to whom recommendation is sent, we find a group of other users who have preferences similar as that of user *x*. If user *x* shares his preferences with the recommendation system, the system will find people having similar preferences and then look at the ideal event for the user *x.* This set of similar users are known as the neighborhood of user *x.*

Once we can find the N set of users to user *x* then we find the events that are liked by a lot of users in the set N as illustrated by figure 14.

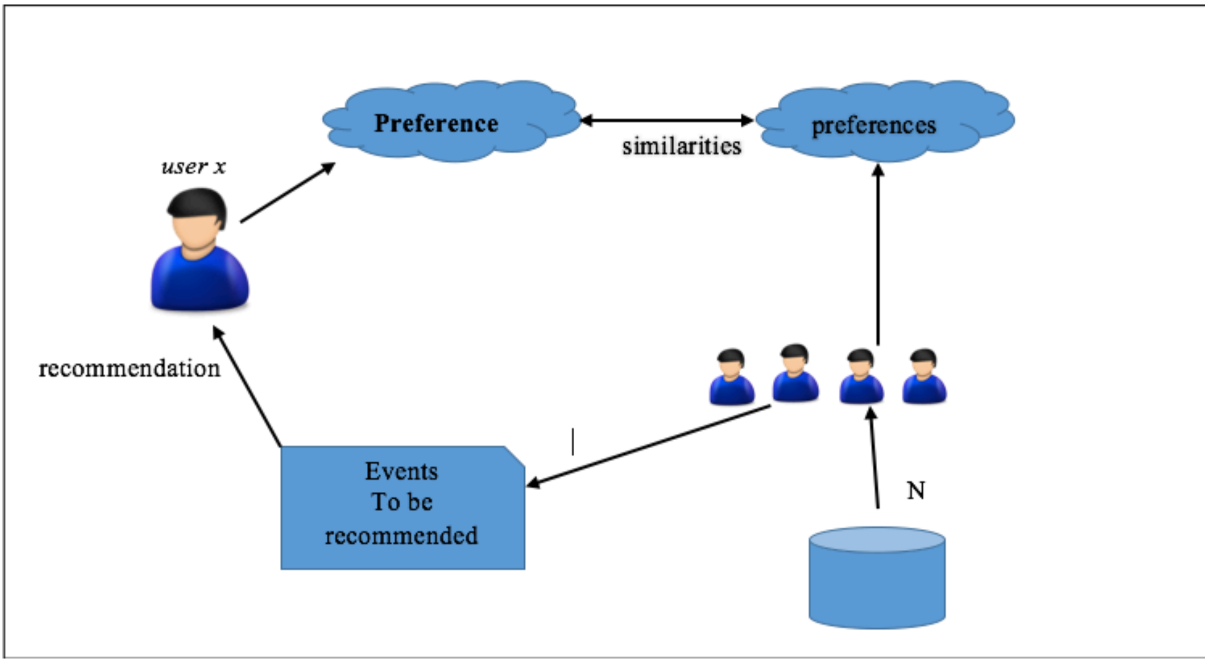


Figure 14: collaborative filtering method of recommendation

So the key in this approach is to find the set of users that are similar to user *x.* Therefore, there is a need to find a notion of similarities between users. For my research I looking at users that specifically belong to the same group since a group has many events. Suppose there are 4 users and a given amount of events. Defining two user *x* and *y* with rating vectors r_x and r_y , there is a need of finding the similarity metric *sim(x,y)* that looks at the rating of vectors r_x and r_y. so there is a high chance that there are events that neither of the users have given their RSVP or rates. In this case we need to know how to deal with the unknown values in the utility matrix. I have defined the intuition that captures users with similar preferences have higher similarities then users with different preferences. The formula bellow illustrates the Jaccard Similarity which I am using in the data set to find similarities in users ratings:

$$sim(A,B) = \frac{|r_A \cap r_B|}{|r_A \cup r_B|}$$

expressing the Jaccard similarity of A,B is known to be the rating vector of user A(r_A) intersecting with the rating vector of user B(r_B). All of this is then divided by the union of rating vectors A and B($r_{A\ \&}\ r_B$) as illustrated in the above formula. Here we will look at the similar events, that both users have rated and present the results in the table 10.

Table 10: Ratings users gave to events in the city of Tokyo in Sports Category

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
A		5	5		3	
B		5		4		4
C	4		2		4	3
D	5	4		2		5
E		2	2		1	
F	3		5		2	3

Looking at this results I have realized some similarities among the users from the ratings they give to events. Implementing the Jaccard similarity between the users it shows the following results illustrated in table 11:

Table 11: Jaccard similarity among users

Similarity	Results	Similarity	Results
$sim(A,B)$	$\frac{1}{5}$	$sim(B,F)$	$\frac{1}{6}$
$sim(A,C)$	$\frac{2}{5}$	$sim(C,D)$	$\frac{2}{6}$
$sim(A,D)$	$\frac{1}{6}$	$sim(C,E)$	$\frac{2}{5}$
$sim(A,E)$	$\frac{3}{3}$	$sim(C,F)$	$\frac{4}{4}$
$sim(A,F)$	$\frac{2}{5}$	$sim(D,E)$	$\frac{1}{6}$
$sim(B,C)$	$\frac{1}{6}$	$sim(D,F)$	$\frac{2}{6}$
$sim(B,D)$	$\frac{3}{4}$	$sim(E,F)$	$\frac{2}{5}$
$sim(B,E)$	$\frac{1}{5}$		

However, using the Jaccard similarity might have a low impact in the users decision, because, it relates the similarities without looking at the value of the rakings. For example in this results we have, the user similarities between A,B and A,E, I have realized that users A and E have many similarities but the rating values of A are higher than E. On the other hand users A and

B have few similarities, but the one event that both users have rate have a higher rating. This clearly that the judging method among users may vary and Jaccard similarity does not observe this. So I ran another experiment using the cosine distance method to compute similarities among the users.

Similar to content base were I have used the cosine distance the same will be applied with rating vectors in collaborative filtering. In this case by defining the *sim(A,B)= cos(*r_A*,* r_B*)*. Since I am computing the sim here, there is a need to zero the unknown rating has illustrated in table 12 below:

Table 12: user ratings before normalization

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
A	0	5	5	0	3	0
B	0	5	0	4	0	4
C	4	0	2	0	4	3
D	5	4	0	2	0	5
E	0	2	2	0	1	0
F	3	0	5	0	2	3

The problem with the cosine distance is that it treats all missing ratings as negative ratings. To fix this problem I have applied the centered cosine method. This is done by normalizing the ratings of a given user by subtracting the average of the rating. Table 13 shows the results of this same ratings after normalization

Table 13: Ratings after normalization

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
A		$\frac{2}{3}$	$\frac{2}{3}$		$\frac{-4}{3}$	
B		$\frac{2}{3}$		$\frac{-1}{3}$		$\frac{-1}{3}$
C	$\frac{3}{4}$		$\frac{-5}{4}$		$\frac{3}{4}$	$\frac{-1}{4}$
D	$\frac{4}{4}$	0		$\frac{-8}{4}$		$\frac{4}{4}$
E		$\frac{1}{3}$	$\frac{1}{3}$		$\frac{-2}{3}$	
F	$\frac{-1}{4}$		$\frac{7}{4}$		$\frac{-5}{4}$	$\frac{-1}{4}$

After normalizing the rating of the users it's possible to zero (0) out the matrix if we add all the rows, meaning that all the ratings have been centered around 0, and 0 becomes the average rating of all the users. By doing this all positive ratings will show that the user has more interest in the event then average which is zero (0). Once the centering has been done it's possible to compute the cosine distance again.

After looking at how to find similarities among users to be able to have greater precision in the preferences, and have a good recommendation, I will briefly explain how this method predicts the user's ratings for future recommendations.

So making user A form the previous table has user x he makes ratings r_x. By using the notion of centered cosine similarity to find the group of N users which are the neighborhood of x and the neighborhood consists of K users who are most similar to user x and have rated same event i, once that is done there can be a prediction for user x and event i. This can be done by applying the following formula that takes the average ratings from the neighborhood:

$$\mathrm{r_{xi}} = \frac{1}{\sum_{\mathrm{y \in N}}^{\mathrm{k}} \mathrm{r_{yi}}}$$

The problem with this formula is that it ignores the similarity of the rating values between users. So it's important to weight the average ratings by the similarity values. So for this to be done the formula for the weighted average will be as follows:

$$r_{xi} = \frac{\sum_{y \in N} s_{xy} r_{yi}}{\sum_{y \in N} s_{xy}}$$

where $s_{xy} = sim(x, y)$. This specific technique is called user to user collaborative filtering so there is a dual approach to user to user collaborative filtering which is item to item collaborative filtering. For this method, instead of starting with the user the starting point could be the event and we can apply the same calculations or algorithm to find similar events. The rating of this events is then estimated based on the ratings for similar events using the same metrics like the one before. But to be able to do and event based collaborative filtering using RSVPs will not be a suitable way to go. Because, creating a rating vector of RSVPs may only be consisting of 1 and 0 so the weight of the average ratings is not going to be significant. This can cause a problem on ineffectiveness of recommendations causing a very low impact.

Thou these two methods in theory are doable, in practice item to item approach outperforms user to user approach in most cases. This is because items are simpler then users to be described. Items contain features that belong to a small set of types or genres or categories. While users have varied and specific tastes. Therefore, is safe to say that when applying item to item approach the impact is going to be higher since it's easier to find and have many features which are related to that user.

Advantages of collaborative filtering is that it works with any kind of item. There is no need of features selection needed. The disadvantages however are:

- This type of methods or approach is known for its cold start because it needs enough users in the system to make possible match. In this particular case if a user is new he might receive very bad recommendations which will cause little influence or impact in his decisions.
- Normally recommendation systems have many users and items. A rating matrix with a million of items and users tend to have many items unrated as we have seen in the utility matrix of the simple example of the data set. So this unrated items causes problems of sparsity. So its somehow hard to find users that have rated the same items.

This was one big problem in my research when trying to only with RSVPs. So given the dataset I will present the sparsity and cold start. Given a set of recommendable events, this events are set in partitions set by level {1-5;5-10; 10-20; +20}, where each level represents the number of positive RSVPs received.

Figure 15 shows this representation where the x-axis represents the partition of positive RSVPs and Y represents number of events in this music category. This clearly show that there are many events with no RSVPs or very little. Showing that event based recommendation systems and many other need more variables to have a higher effect, hence, causing a high impact.

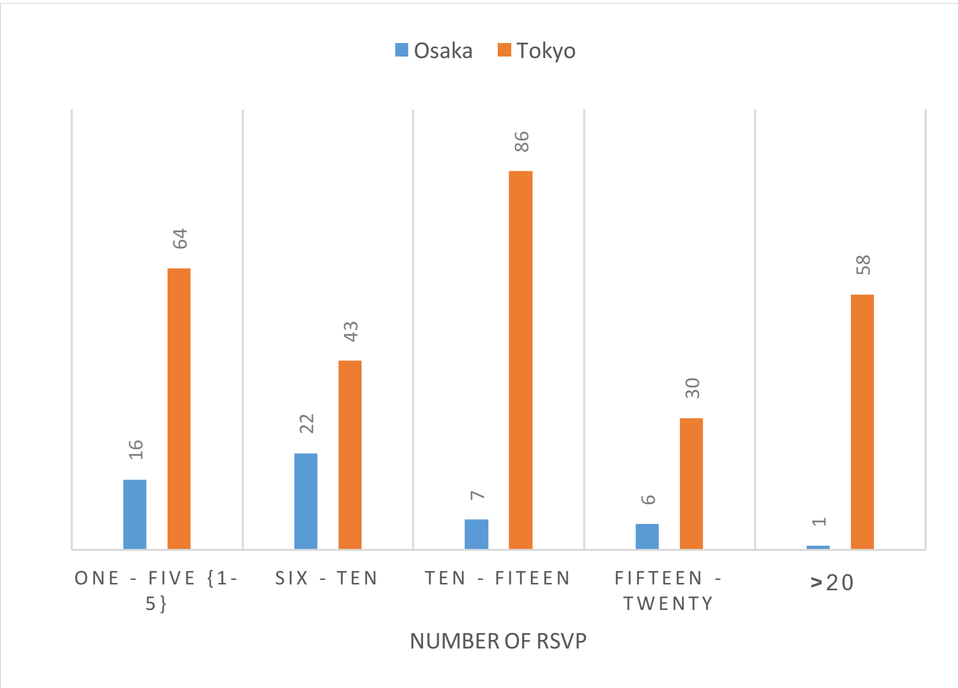


Figure 15: Music category RSVP sparsity

Another problem we see from this is that events that have not been rated get the so called new item problem. The unrated events or events with few positive RSVPs will most slightly not be recommended. This clearly identifies that using RSVPs or ratings will not bring improvement in the precision of the recommendation. It's evident that although there is a high number of people in groups still there is a few number of attendance. The reason under this could be the fact that meetup recommendation system does not use a hybrid method looking at the potential social network features that their system presents.

The figures bellow shows that the same behavior when using RSVP data in other categories is the same. Even thou events a popular the sparsity of RSVP data is still available due to the lack of evidence on what was the real factor that users have attended such event under this category. We look at figures from both Sports and Camping from both cities and look at the sparsity of the RSVP data.

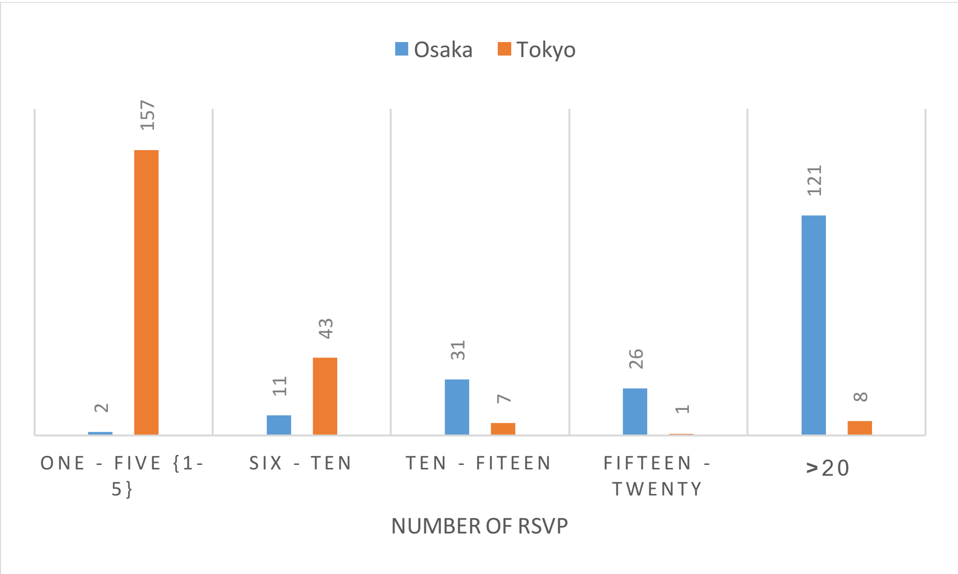
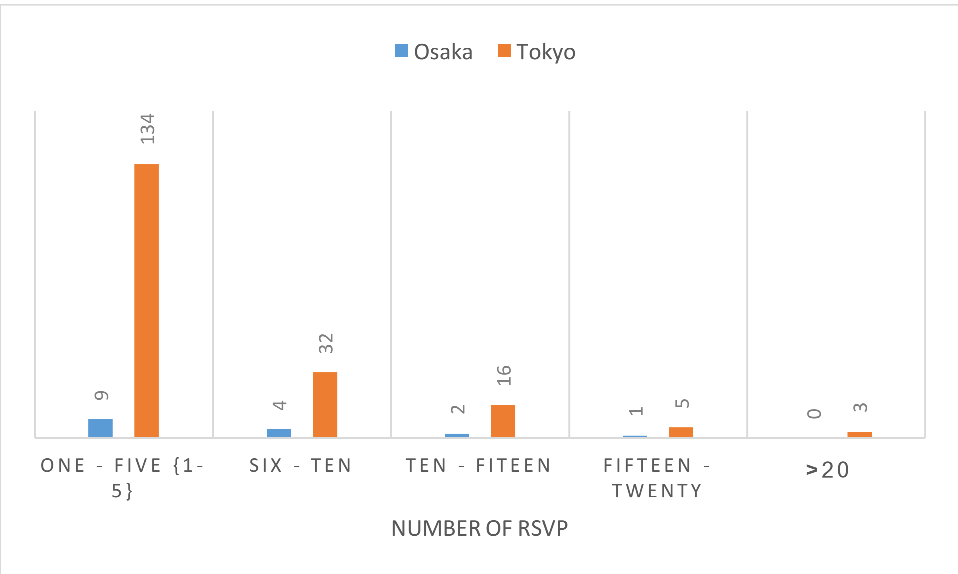


Figure 16: Sports category RSVP sparsity



4.4.3. Hybrid methods

Both of the methods we have seen before have a significant impact in user's decisions. But both present problems that could be overcome by using the hybrid methods. The hybrid methods are a combination of both collaborative filtering and content based methods. Normally the content based method is added to collaborative filtering then by so doing the item profile could solve the new item problem experienced in collaborative filtering.

In meetup and many other event based recommendation systems have a very good solid demographic data. This could also play a good role in solving the new user problem being experienced in the content based method, where set of users sharing similar demographic data or more technically the N neighborhood of users based on his demographic data can sometimes help build a very solid recommendation system.

Another combination that could solve the RSVP problem is by applying the global baseline approach. This will estimate the mean rating of the event and will be added by the above level of average of that event. Both the average rating and the level above will be subtracted from the rates average of the user we are trying to recommend this particular event. This is known to capture events popularity and users bias. The formula for this type of approach is shown below:

$$r_{xi} = b_{xi} + \frac{\sum_{j \in N(i,x)} s_{ix} \cdot (r_{xj} - b_{xj})}{\sum_{j \in N(i,x)} s_{ij}}$$

where $b_{xi} = \mu + b_x + b_i$, understanding that b_{xi}as the baseline estimate for rating r_{xy} being μ the overall mean event rating in the system, b_x is the rating deviation of user x which the average rating user x has rated μ and b_i is the rating deviation of event i. Adding them up gives the baseline deviation of user x and event i. The result of all this will then be combined by adding to the collaborative filtering piece as explained in previous section.

The Hybrid method will most definitely outperform the other two approaches because it combines the two methods presenting solutions to the problems the other two methods have. From this section we see how it is important to prepare the utility function in a manner that all angles are analyzed and the defaces of different utility functions can be solved and improve the effectiveness of the recommendation process.

The model drown bellow represents how our hybrid method approach will work. Given a user *(x)* that has his profile showing all his preferences, our hybrid method look to first optimize his event profile matching with his preferences and previous RSVP, similar to what is done in content based method where the TF-IDF was implemented to extract descriptions from social network. This social network features will both work with demographic data such as location and hometown and match them with comments or group discussions they have with other users. The event profile is then matched using the cosine similarity in a utility matrix. This utility matrix will not only have one sided ratings as seen in the collaborative filtering, but, it will combine both RSVP vectors, Event Ratings and the contextual information vector done by using TF-IDF.

This procedure is then transferred to the recommendation list, and the event with a highest matching data with the user's preference is then determined to be of highly recommendation.

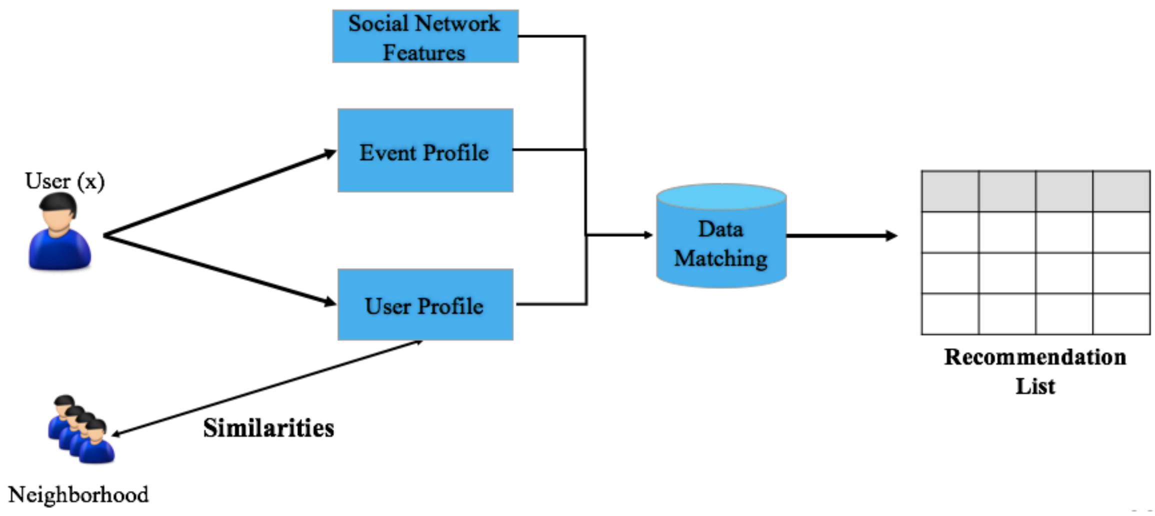


Figure 17: Hybrid method model with social network features

4.5. Data set limitations

During the data analysis process, I have encounter some difficulties dealing with the data set provided by the meetup API. Looking at the recommendation process one of the tasks is to know if the recommendation was successful. The success of the recommendation is measured by its usefulness to the user who was targeted. With such clarifications in the dataset I could have a clear indication if the user has attended a given event.

Without such confirmation, the data set can only present results that I could only come with some assumption but not a concrete and firm idea of the user's presence in the event. Meetup API provides in their JSON data set methods to see the RSVPs of a given event and of a user that a given event was recommended to. But that however is clearly of the user's presence because many people might give a positive RSVP but in the end they don't attend the event. This then makes it hard to answer if RSVPs can serve as a useful metric for user's profile when building a content based recommendation system.

When retrieving information from the API, meetup limits the results to only 200 tuples per query run. This makes it difficult to track the if there are any sets repeating. With such difficulty there was a need to limit at times the extraction of the dataset which took some time to organize it.

The data set also does not provide the user's feedback. I find it particularly important for such systems using social network based recommendations having a feed from the activities. Some systems require users to provide a rating but I suppose that a questionnaire to respond to quality of the recommendation would help develop the hybrid system. This was another difficult I found, because with this dataset I can only conclude with assumptions and probabilities and not a fair concrete answer based on the quantity of statements given by users after recommendations.

This in particular is a point that I would recommend future researches to observe when dealing with behavioral research such as this. The ideal way would be to simultaneously test all this assumption being given by the current data set and run questionnaire to a seatrain amount of users that have attended the events and match their appreciations in relation to their ratings they had provided to the system.

5. Conclusion

In this study, I have conducted an online data sample experiment to understand the impact recommendations have on the user's decisions. The study integrated some theories of social influence to decisions and recommendation systems looking at the same time from a practical point and a theoretical point of view. The results presented help understand some key factors that determines ways that individuals are influenced by a recommendation thus to say that it has an impact when a user is making his decisions. To achieve all subsequence judgements to my possible hypothesis I conducted several tests on the online data extracted using the meetup.com API.

5.1. Research findings

Based on the research objectives carried by this research I was able to come with following findings:

1. How do new users to recommendation system respond to recommendations?
   - ❖ From the analysis carried based on the utility function new users in a content base recommendation and collaborative filtering methods will not respond well to recommendations. Therefore, there is a need of an influential factor for him to have a positive response. This factors are mainly based on the latent information the user provides to the recommendation system such as who are his SN friends and what type of activities they are engaged in. this type of data will allow the system to build a user profile with accuracy and his neighborhood. Improving the recommendation quality and have the impact desired to the new user which is a positive response.
2. How do users respond to events near them?
   - ❖ From the results it was clear that location is a key factor in recommending events near the users. Most of the users tend to respond to events near them. From the data set is also clear that there are many users who are influenced by events in big cities near them. This came to a little conflict of the distance factor, but that incorporates the social activity factor where user tend to have what's best for them so some are insignificant if the event is far from them.
3. How does contextual information play in the recommendation influence?
   - ❖ The intrusion rate found on this research clarifies the importance of contextual information. I have noticed that there was a big number of users from other cities or groups engaged in activities from other groups or cities. This is mainly because of the social network interaction that the system provides when recommending events to the

users. Information such as demographic data and feeds play their importance in such recommendation as users tend to have the sense of belonging in their surroundings.

I have conducted also experiments on the data to be able to retrieve possible infiltrations of members belonging to a different group and having different preferences and judged the possible reasons for such infiltration. From this stand point I have examined the factor of some members having registrations in other cities would accept recommendations to attend events in other cities having different preferences as well.

Based from the data collect it was possible to demonstrate the high structured content recommendations have higher impact in relation to basic structured recommendations. Groups that would have more interactions among users, it was possible to see that such groups could impact users from other groups, cities and ethnicity.

The impact of predicted recommendations for future events are to be seen with higher influence when there is greater contextual information such as ratings, previous experiences, group leaders RSVPs. Groups that tend to have better contextual information their events where highly numbered and at times being represented with user limitation to attend or even closed events. With this results we can see that social networks serve a high factor in this impact to users decisions because many of its users tend to describe past events experiences not only with a simple ranking but with contextual information such as comments and photos.

In contrast to more traditional systems, where user would simply interact with recommendation with a simple rating, event recommender systems provide clear data due to the fact that it uses social networking style to its recommendation. So members taking part in this events see the recommendations has a positive answer to their doubts or needs to socialize.

With this I can say that the findings of this study have significant important implications for future studies. On a more practical way my research was conducted on already existing dataset. Improvements could be done if similar study is taken on real-time online data and utilize the finding to develop a recommender system based on social network contextual data incorporations. From my current findings it's possible to picture that such research would present even hidden key factor that could serve as a higher impact to user's decisions.

From my studies it clearly and safe to be said that more researches need to be done. From the work that I have conducted its clear that further researches are needed to understand the impact of recommendation systems on user's decisions combining this with their preferences. Issues such as trust, decision bias, and preference realization are not clearly analyzed on this research. It is important to analyze such facts because they are directed linked to the context of recommendations in any kind of applications. Occasionally such issue could be introduced into the recommendation models and have a higher precision and timing to satisfy and serve as an even higher factor to user's decision making. With that been said I can state that this researches are still young and there is still much that needs to be investigated. Therefore, the combinations of both behavioral studies and information systems are important. Because is important to understand from time to time how people behave and come up with better concepts to be introduced to systems.

In this manner of thinking and analyzing the results I got, I can complete my research by answering my research questions:

1. RSVPs are important in to a given extend. In many systems they would serve as good variant since it provides at least an intension of the user to attend the event. But even thou he does not attend, the simple fact of the person giving his RSVP represents that he showed some interest in the recommendation. Smarter systems could then develop and link the RSVPs with other variables to understand the user's preference and build a personalize user profile or the user's neighborhood.
2. Depending from user's to user's there are different factors that would indicate such key factors. This was analyzed in my research I could see that his social network interaction is a big factor, because it provides a lot information that cannot be simply analyzed and generated by ratings, but also other methods or algorithms.

So thou recommendation systems and applications still need improvements to be able to personalize and meet many needs searched by users, it is clear that such system serves a as a great impact in the user's decision making process. It benefits the user with issues hidden information, actions or attitude reflection from his previous experiences and many other key factors. So I conclude by saying that recommendation systems are capable to impact user's in their decisions even thou there are many improvements needed.

6. References

1. Goldsmith, E. Resource management for individuals and families (5th ed.). Upper Saddle River, NJ: Pearson, 2013.
2. W. Hill, L. Stead, M. Rosenstein, and G. Furnas, "Recommending and Evaluating Choices in a Virtual Community of Use," Proc. Conf. Human Factors in Computing Systems, 1995.
3. G. Adomavicius, A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions" IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 17, NO. 6, JUNE 2005.
4. J.S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Artificial Intelligence, July 1998.
5. M. Balabanovic and Y. Shoham, "Fab: Content-Based, Collabora- tive Recommendation," Comm. ACM, vol. 40, no. 3, pp. 66-72, 1997.
6. G. Shani, David Hekerman, and Ronen I. Brafman. "An mdp-based recommender system". Journal of Machine Learning research, 6: 1265-1295, 2005.
7. P. Lazarsfeld, Katz E. "Personal influence: The part played by people in the flow of mass communications", Glencoe, IL: Free Press New York, P.F.1955.
8. Kozinets, R. V., de Valck, K., Wojinicki, A. C., & Wilner, S. J. S.). Networked narratives: Understanding word-of-mouth marketing in online communities. *Journal of Marketing, 74*, 71–89. 2010
9. Tomoharu Iwata, Kazumi Saito, Takeshi Yamada, "Modeling User Behavior in Recommender Systems based on Maximum Entropy", WWW '07 Proceedings of the 16th international conf. on World Wide Web, Pages 1281-1282, May 2007.
10. G. Adomavicius, Jesse Bockstedt, Shawn Curley, Jingjing Zhang, "Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects" Information system research, volume 24, issue 4, September 5, 2013
11. Augusto Q. de Macedo, Leandro B. Marinho, "Context-Aware Event Recommendation in Event-based Social Networks", RecSys '15 proceedings of the 9th ACM conference on Recommender Systems, Pages 123-130, september 2015.
12. www.meetup.com
13. Jacowitz, K.E., and Kahneman, D. "Measures of anchoring in estimation tasks," Personality and Social Psychology Bulletin, (21) pp. 1161–1166, 1995.
14. http://meetupapi.readthedocs.io/en/latest/meetup_api.html#meetup.api.meetup.api.Client.GetGroups.