DOI: 10.21917/ijsc.2021.0333

USER ACTIVITIES ANALYSIS IN LOCATION BASED SOCIAL NETWORK VIA ASSOCIATION RULES

Yahia Belayadi[1], Abdallah Khababa[2], Abdelouahab Attia[3], Zahid Akhtar[4]
[1,2]Department of Computer Science, Ferhat Abbas University, Algeria
[3]Department of Computer Science, Mohamed El-Bachir Ibrahimi University of Bordj Bou Arreridj, Algeria
[4]State University of New York Polytechnic Institute, United States of America

Abstract

In recent years, the field of the Internet of Things (IoT), including smart and wearable devices, has witnessed a tremendous advancement leading to the collection of a wide variety of information not only about users but also their activities via various systems such as social networks, apps and so on. Thus, the collection of this large amount of data allows social systems to reach a wide variety of targets and gives more visibility about users and their profiles. It can also help to improve the services and functionalities of the users. Besides, the analysis and prediction of user's activities in location-based social networks (LBSNs) have received much attention both from industries and research communities, especially in smart city developments, which give much importance to the automation of the LBSNs. In this paper, we present a new method based on association rules for user activity analysis in LBSNs. In particular, the Apriori algorithm has been applied to extract the consequential and advantageous rules to categorize users' profiles. Empirical evaluations on a publicly available large-scale real-world dataset, named Gowalla, demonstrate the effectiveness of the presented association rules-based system in analyzing users' activities via LBSNs.

Keywords:

Complex System, Social Networks, Association Rules, Apriori Algorithm, Gowalla Dataset

1. INTRODUCTION

Owing to technological developments and advances, mobile devices, particularly Smartphones and smart tablets, and wearable devices are being widely used in many daily applications. Moreover, the progress in Internet technology has made smart mobile devices quite a vital physical extension of human users. In fact they have become personal assistants and being widely employed to conduct various daily life activities as well as researching information such as travel plans, online shopping, online education, and job searches. The most current smart devices have several in-built sensors such as gyroscope, temperature sensor, accelerometer, camera, and GPS sensors, etc. The data obtained from these sensors can be used to determine user's trajectory patterns and for recognizing their moving behaviors.

Numerous location-based apps and social networks have emerged, which provide various services to users, e.g., Gowalla [1] that is also a topic of investigation in this paper. Location-Based Social Networks (LBSNs) typically incorporate social network facilities with mobile trajectory data [2]. LBSNs have been studied in the literature from different perspectives such as distributed representation, models of the neural network, prediction of social link, and recommendations. For instance, Gao et al. [4] have studied the content information on LBSNs. Authors in [4] investigated point of interest (POI) recommendations with different content information on LBSNs to determine user's sentiment indications and user interests. While on the contrary Qiao et al. [5] designed framework, named UP2VEC, which is based on a heterogeneous graph with a representation for learning, of users, and POI in LBSN. Whereas, Fan et al. [6] proposed a model for representation learning techniques for LBSNs called JRLM++, which modeled the chronological check-in sequence and social information without geographical information. Feng et al. [7] introduced new pair-wise metric embedding to design sequential POI transition. The PRME-G algorithm was developed to combine three factors, i.e., sequential transition, individual preference, and geographical influence. Chen et al. [8] considered the advantages of check-ins of related users for the recommendations in the LBSN. Huayu et al. [9] introduced a two-step structure for the POI recommendation system, which took into account the check-in information of three types of friends, i.e., social friends, location friends, and neighboring friends. Diem et al. [10] created a model to predict the user's next activity based on the past and current context of the user. This model contributed an outstanding tool for combining location prediction with transportation planning and operations processes. Likhyani et al. [11] investigated coarse-grained location types by the association between sparse location data and map information. While Shem-Tov et al. [12] designed a model to guide users in emergencies situations to a best nearest friend. He et al. [13] designed a deep neural network algorithm called NCF, which merges both generalized matrix factorization and multi-layer perceptron under one framework. The best- achieved performance of NCF was obtained by tuning hyper-parameters [14] and have been applied in different fields, including computer vision, semantic web [15], complex system and social networks analysis [3]. The data mining techniques are useful to extract information for data discovery as well as other predictors. For instance, building a high-quality mobile commercial application for recommendation systems with higher accuracy and eventually attracting customers to partnering businesses.

Thus, in this paper, we propose a novel method for user activity analysis based on the model of association rules [16] to jointly describe social network structure with the user's favorite places and behaviors. In the proposed approach, the comments by users and movements of users from place to place complement each other. Thus, it is a powerful way to describe heterogeneous data kinds in the LBSNs. In particular, our model considered two factors in the generation of the model of association rules, i.e. place visit by user and comment given by users. As a result, we presented the model as ($A \rightarrow C$ [support %, confidence %, and lift]), where A and C are the places visited by users and support, confidence, and lift are metrics to evaluate association rules.

The remainder of this paper is organized as follow: The proposed methodology is described in section 2 presents, section 3 presents experiment results with discussion. The conclusion is outlined in section 4.

2. PROPOSED METHODOLOGY

The Fig.1 depicts a block diagram of the proposed method of location-based association extraction model in a social network system. The process consists of three main stages:

Step 1: Preprocessing of a dataset, i.e., operations such as removal of null value, records with missing or wrong locations;

Step 2: Application of the Apriori algorithm [17], which comprises two steps

a) extracting overall frequents item sets,

b) extracting useful association rules;

Step 3: Visualization and interpretation of the rules.

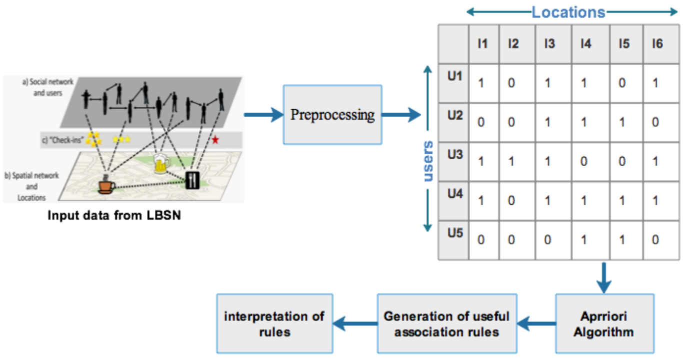


Fig.1. Block diagram of user activity-based association rules.

2.1 LOCATION SOCIAL NETWORK

The concept of location social network [18] [19] corresponds to a 3-tuple $G=(V,E,C)$, where V is the set of nodes (i.e., users), E is a set of edges that designates the social connections between users, and C is the users' check-in dataset. All these parts of data are usually collected from a user checking in at a definite time and place.

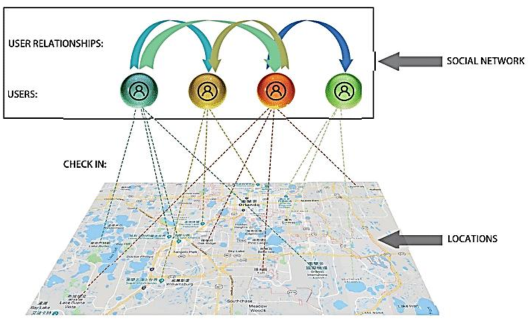


Fig.2. Location-based social network LBSNs architecture [20]

The Fig.2 illustrates an LBSN architecture [20] that uses the GPS features to locate users and inform them and transmit their location and location-tagged media content such as texts, photos, and video by smartphones. This architecture encloses two layers: (i) Online Social Networks (OSNs) Layer: in this layer, users form their social networks apps make interactions in LBSNs; (ii) Physical Location Layer: mobile users who hold their mobile devices can broadcast their location by checking-in and checking-out with location-based services apps.

2.2 ASSOCIATION RULES

The search for association rules between Boolean attributes in large databases has been initiated by Agrawal et al. [16]. Apriori [17] is the algorithm that generates all association rules from a transaction table. This algorithm is divided into two phases. The first is to find all the frequent itemsets (set of attributes), i.e., the itemsets whose support is greater than or equal to a minimum threshold (Minsup) given by users. The second one is to identify the association rules from the frequent itemsets generated in phase one. The association of rules with confidence is greater than or equal to the minimum threshold (Minconf) given by users too. Finally, we state that some rules are accepted, while others are rejected by the Apriori algorithm.

Suppose that we have association rules defined as ($A \rightarrow C$ [support%, Confidence %, and Lift]) with $A \cap C = \emptyset$. (i.e. if A then C). A and C are frequent itemset with a measure of rule, including support %, Confidence %, and Lift indicating the reliability, precision, and validity, respectively [21]. Also, A is called the left-hand-side (LHS) and C is called the right-hand-side (RHS). Support measure the reliability of an association rule. This measure computes the frequency of an item or itemset appearing in the database and calculated as:

Support(AC)=Number of A and C/Total number of transaction (1)

Confidence indicator of precision of an association rule is given as:

Confidence(AC)=support(AC)/support(A) (2)

Lift is an indicator of validity of an association rule, which also can be read as a ratio of likelihood, i.e., under H_0: A and C is independent, which is given by:

lift(AC)=support(AC)/(support(A)×support(C)) (3)

Apriori Algorithm

Input: D: A set of transactions, *Minsup*: Threshold of min support

Output: L: Frequent Itemsets

Step 1: L_1 = {1-frequents itemsets)

Step 2: k=2;

Step 3: While L_{k-1} non vide do

- **a.** C_k=Apriori-Gen (L_k)
- **b.** For each t of D do
  - **i.** C_t=Subset (C_k,t); {the itemset candidate contenus in C_k}
  - **ii.** For each c of C_t Do
    - **1.** *c.count*++ ;
  - **iii.** End For
- **c.** End For
- **d.** L_k={c *de* C_t/*c.count*≥*minsup*};
- **e.** k++;
- **f.** End while

Step 4: Return UL_k;

Generation rules algorithm - Algorithm-gen-rules;

Step 1: For each frequent itemset f_i do

  a. Generate all sub-itemsets I_j of f_i

Step 2: End For

Step 3: For each sub-itemset I_j of f_i Do

  a. generaterule($I_j \rightarrow (f_i - I_j)$)

  b. if *Minconf* ≥ *threshold min cof*

Step 4: End For

2.3 DATA PREPARATION

Before applying the Apriori algorithm, both matrices user-location and user-comment (presented in Fig.3.) are created.

2.3.1 User-Location Data:

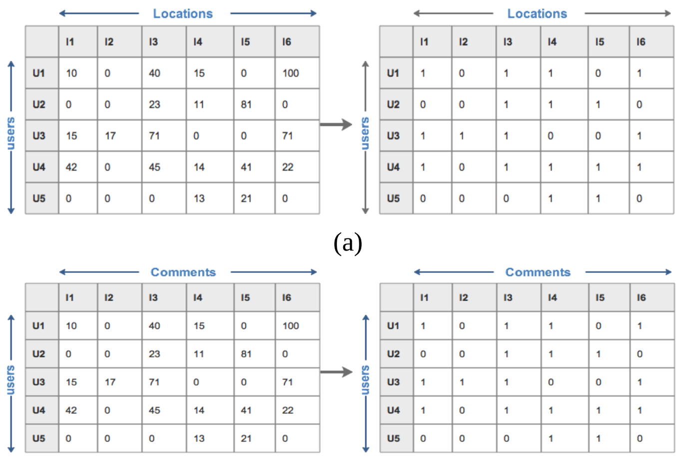
Let U and L denote the number of users and locations, respectively. First, the frequency $F_{u,l}$ of each user that is located in location l is computed. The representation of $F_{u,l}$ is:

$$F_{u,l} = \begin{cases} f_{u,l} & \textit{Number of user visited location} \\ l & \textit{Otherwise} \end{cases} \quad (4)$$

Then, we define the user-location based on binary data form the interaction matrix $ML \in R^{U \times L}$ of users and the visited places as:

$$ML_{u,l} = \begin{cases} 1 & \text{if } f_{u,l} \geq 1 (\textit{i.e. user visited location}) \\ l & \textit{Otherwise} \end{cases} \quad (5)$$

At this point, a value of 1 in $M_{u,l}$ indicates that the user's u visited the location l. However, it does not mean u likes it. Likewise, a value of 0 does not necessarily mean u do not like it. It may be because the user is not aware of the location l.



(a)

(b)

Fig.3(a). Matrix of users who visited location, (b) Matrix of user's comments

2.3.2 User-Comment Data:

Let U and C denote the number of users and comments, respectively. First, the frequency $C_{u,c}$ of each user that has made comment about check-in spot is computed as:

$$C_{u,c} = \begin{cases} c_{u,c} & \textit{Number of comment} \\ l & \textit{Otherwise} \end{cases} \quad (6)$$

Then, the user-comment based on binary data form interaction matrix $MC \in R^{U \times C}$ from users and visited places are defined as:

$$MC_{u,c} = \begin{cases} 1 & \textit{if } c_{u,c} \geq 1 \\ l & \textit{Otherwise} \end{cases} \quad (7)$$

At this point, a value of 1 in $M_{u,c}$ indicates that the user's u put comment c.

2.4 DATASET GOWALLA

We evaluated our proposed methodology on the publicly available datasets named Gowalla. Gowalla is an LBSN, which was launched at the beginning of 2009 permitting users to share their locations or geo-tagged information such as comments and photos with friends through check-ins. From the Gowalla service, the database Gowalla was created by collecting public data. It has 30,367 geo-referenced spots in New York City with information concerning 357,753 visits for 19,183 users of Gowalla. The Gowalla dataset is composed of files such as:

1. spots.txt that contains information about spots in New York, including ids, names, and geospatial coordinates,
2. users.txt file with Gowalla users, containing ids, names, hometowns, and geospatial coordinates,
3. highlights.txt that is a tab-separated file with spots marked as highlights by Gowalla users, containing spot and user ids together with a textual description and type of spot,
4. users-spots.txt that is a tab-separated file with spots visited by Gowalla users, containing spot and user ids, together with a Boolean value indicating if the user frequently visits the spot or not.

The properties of Gowalla data are summarized in Table.1, while categories are outlined in Table.2.

Table.1. Properties of Gowalla dataset

Designation	Number
Users	19183
Activities	8334
Spot id	30367
Connected hometown	11129
Frequency cheek-in	357753
Categories of spot or highlights.txt	9

Table.2. Categories of Gowalla dataset

Number	Category
1	Architecture and building
2	Art and Culture
3	College and education
4	Entertainment
5	Food
6	Nightlife
7	Park, Nature and recreation
8	Shopping and services
9	Travel and Lodging

In this section, we describe the empirical evaluation of the proposed methodology using Gowalla data. For implementation, we used both R language and Tanagra tools, which permit applying the Apriori algorithm. The data preprocessing step included removing records with missing or wrong locations. The data preprocessing step included removing records with missing or wrong locations. Then, the matrix data ML (users-locations), which indicates the locations visited by users using the process (section) using Eq.(4) and Eq.(5), was built.

Also, the same process for MC (users-comments) by using Eq.(6) and Eq.(7) was performed. The objective is whether to keep this user visit or choose a location from the possible connections to a place that contains this type. Next, the Apriori algorithm was applied. We performed two kinds of experiments. The First one is to investigate the locations frequently visited based on users-spots.txt. While, the second one is to study the behavior of users based on data given in highlights.txt and offered by Gowalla services. Thus, we have generated both data matrices (user-locations and user-comments).

2.5 EXPERIMENT 1: ASSOCIATIONS RULES WITH USERS-LOCATIONS

Here, we describe some rules extracted by Tanagra software and R. Besides, we present rules analysis by graphs including scatter plot for rules, two-key plot, graph of rules and parallel coordinate plots. This type of visualization has been generated by the ArulesViz package [22] and vis network package [23]. We selected rules based on each category (see Table.2).

R provides a package named “arules”, which permits applying several algorithms for extraction of association rules. Besides, it offers some functions to plot graph and statistics about numbers of associations rules depending on parameters such as maxlen that indicates length of rules (i.e., number of items), supp: support of rules and Confis: the confidence of rules. Then, we import the file .CSV by instruction: file<-read.csv(“filename.csv”), where filename.csv contains data matrix user-locations. In the next step, association rules were extracted by:

rules=apriori(file, parameter=list(supp=α, conf=β, minlen=γ));

where α and β range [0,1], and γ should be more than 1.

Moreover, Tanagra software permits to apply apriori algorithm by fixing parameters min support and min confidence. The Table.3 presents an example of extracted rules by Tanagra software. The parameters are 13%, 75% for support and confidence, respectively, with max length of items 3.

In Table.4, we put some rules with description. In addition, in Fig.4, the path way of these rules in map using Google maps have been presented.

Table.3. Association rules by Tanagra (Number of rules: 9)

Number	Antecedent	Consequent	Lift	Support (%)	Confidence (%)
1	“ Ed Sullivan Theatre =true”	“World Trade Center=true”	153.1	0.131	100
2	“Ellis Island=true”	“Statue of Liberty=true”	40.82667	0.261	80
3	“Grand Central Terminal=true” –”Radio City Music Hall=true”	“Rockefeller Center=true”	34.79545	0.131	100
4	“Castle Clinton National Monument=true”	“The Museum of Modern Art (MoMA) = true”	24.43085	0.196	75
5	“Rockefeller Center=true” –”The New York Public Library=true”	“Grand Central Terminal=true”	19.62821	0.131	100
6	“JFK John F. Kennedy International=true” - “Madison Square Garden=true”	“Grand Central Terminal=true”	19.62821	0.131	100
7	“Grand Central Terminal=true” – “Hard Rock Cafe=true”	“Times Square= true”	6.80444	0.131	100
8	“The Museum of Modern Art (MoMA)=true” –”Hard Rock Cafe=true”	“Times Square= true”	6.80444	0.131	100
9	“Madison Square Garden=true” - “Bryant Park=true”	“Times Square= true”	6.80444	0.131	100

Table.4. Some rules with descriptions

Rules	Sup. %	conf.%	lift	description
{Radio .City. Music. Hall, The. New .York. Public. Library} => {Times. Square}	1.3	66.67	4.53	This rule indicates that 1.30% of users of Gowalla are interested in places Radio. City. Music. Hall and The New. York. Public. Library favorite the Times. Square places with 66.67% precision
{Grand. Central. Terminal Madison. Square. Garden} => {JFK. John. F. Kennedy. International}	1.3	50	3.71	These rules show that 1.30% users located in Grand. Central. Terminal station and the touristic place Madison. Square. Garden goes to JFK. John. F. Kennedy. International airport with 50% precision
{Ellis. Island} => {Statue. Of .Liberty}	2.61	80	4.08	This rule presents that 2.61% of the visitors of Ellis Island prefer to profit the occasion to visit Statue. of. Liberty with 80% precision

{Ed Sullivan Theatre} => {1 World Trade Center}	1.31	100	1.53	This rule shows that the 1.31% of users of Gowalla who are located in spot “Ed Sullivan Theatre” has been located in World Trade Center of 100% precision
{Bowling. Green} => {Katz .s .Delicatessen}	0.65	50	6.96	This rule indicates that 6.5% of users who are located in Bowling. Green garden is located in Katz’s. Delicatessen restaurant with 50% precision
{Federal. Hall. National .Memorial} => {Bryant. Park}	0.65	100	4.03	This rule informs us that 0.65 % of Gowalla users who visited the Federal. Hall. National. Memorial favorite the Bryant. Park with 100% precision
{Temple. Bar} => {The. Museum of. Modern. Art. MoMA.}	0.65	50	1.62	The rule shows that 0.65% of users who went to Temple. Bar like to visit the. Museum. of. Modern. Art. MoMA with 50% precision
{Lucky. Strike} => {Frying. Pan}	0.65	50	9.56	The rule indicates that 0.65% of users located in Lucky. Strike restaurant are located in Frying. Pan restaurant with 50% precision
National.9.11. Memorial. Museum} => {Tekserve}	0.65	50	9.56	This rule signifies that 0.65% of users located in 9/11Memorial.Museum restaurant are located in Tekserve with50% precision
{Roosevelt. Island. Station} => {City. Field}	0.65	100	3.47	This rule show that 0.65% of users Gowalla located in Roosevelt. Island station are going to the city. Field stadium with 100% precision

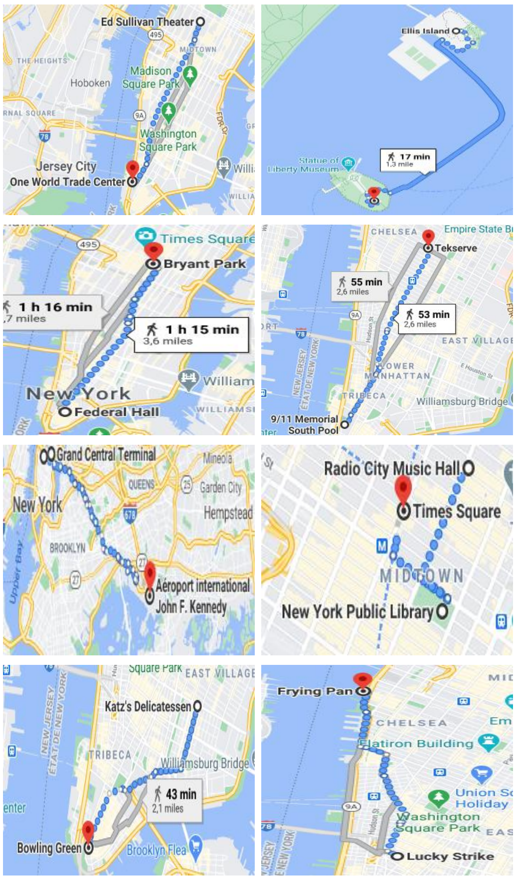


Fig.4. An illustration of some paths between antecedent and consequent of rules visited by Gowalla users in NYC

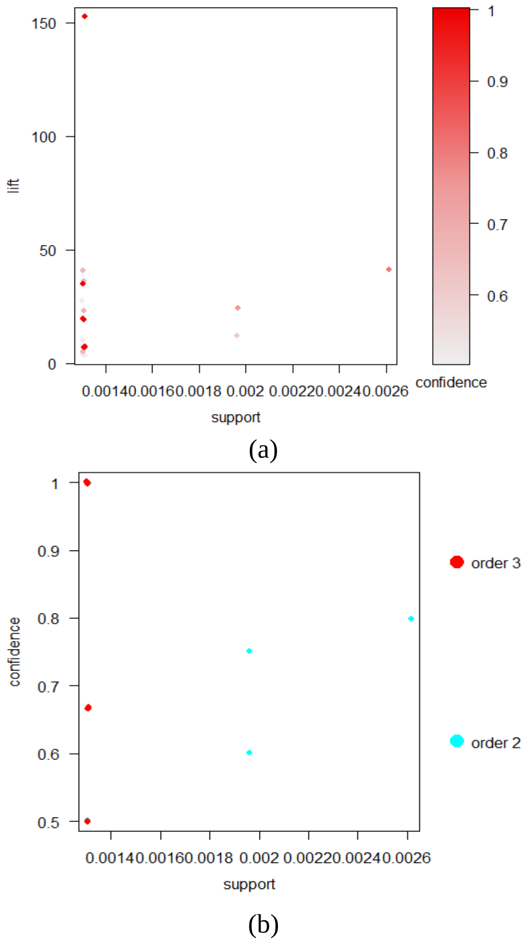


Fig.5. Visualization of: (a) scatter plot for 19 rules; (b) two-key plot

We present some results by graphs generated using R software for 19 rules in Fig.5, where the scatter plot for 19 rules with the two-key plot is shown. The Fig.5(a) presents the diversity of rules according to support, confidence, and lift measures. The Fig.5(b) shows the order of rules according to the support and confidence measures.

The Fig.6 illustrates the parallel coordinates plot for 19 rules of order 2 and 3, respectively. From this figure, it is clear to read rules (i.e., {Ellis. Island}→{Statue of Liberty} or {Ed. Sullivan. Theater}→{X1. World. trade. Center} and so on

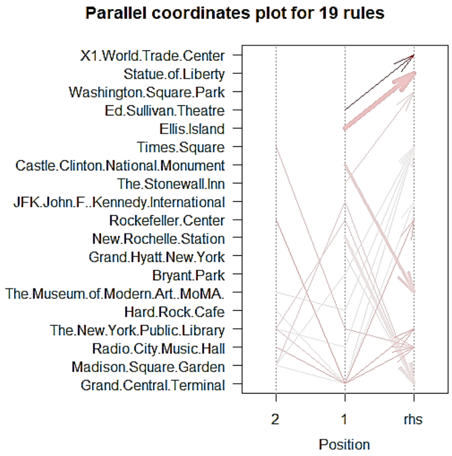


Fig.6. Parallel coordinates plot for 19 rules

2.6 EXPERIMENT 1: ASSOCIATIONS RULES WITH USERS-COMMENTS

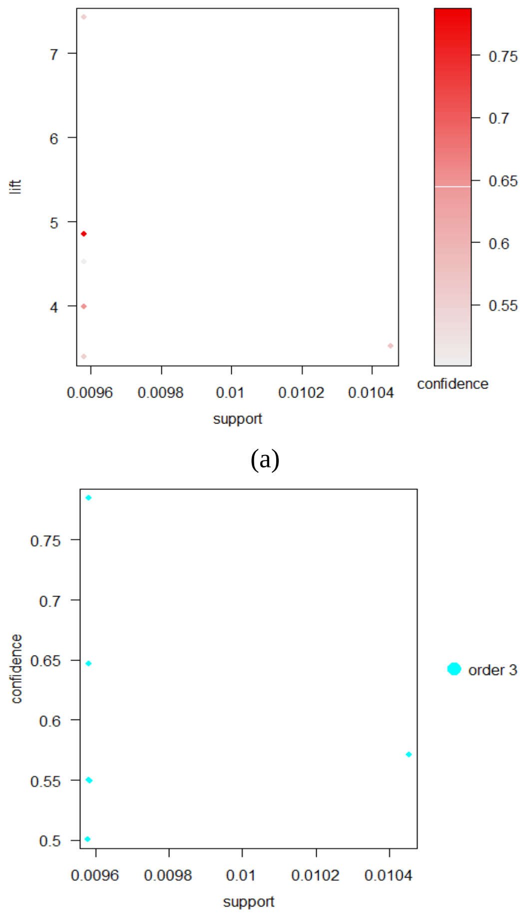
The limitation of the ubiquitous "check-in" based systems can be overcome by leveraging the location with user comments. Gowalla added a new service called "Highlights," which aims to define places with user experiences rather than check-ins. It works as follows the Gowalla user tags real-world places, e.g., name of favorite "Date Night" spot, best "Tacos Place", favorite "Watering Hole" among other things. Every answer can reveal a bit of the user's identity, and where they spend their time. Hence, with the services named "experience-data" linked with certain places, users can learn from the interesting experiences of their friends when they check-in at those places. For instance, if you check-in at a certain coffee shop, you might be notified that a friend of yours had an experience not too long ago at that very same location. Likewise, if you are looking for the best coffee shop in a city your visiting, you can visit the highlights page to see what users deem the best coffee in town. When coupled with a user's social graph, things get interesting and provide a unique additional experience using Gowalla.

Number of rules : 1992

N°	Antecedent	Consequent	Lift	Support (%)	Confidence (%)
1	"Grassroots Tavern=true" - "WCOU Radio (Tile Bar)=true"	"Boka=true" - "Swift Hibernian Lounge=true"	2341,00000	0,043	100,000
2	"Swift Hibernian Lounge=true"	"Boka=true" - "Grassroots Tavern=true" - "WCOU Radio (Tile Bar)=true"	2341,00000	0,043	100,000
3	"Grassroots Tavern=true" - "Swift Hibernian Lounge=true"	"Boka=true" - "WCOU Radio (Tile Bar)=true"	2341,00000	0,043	100,000
4	"Grassroots Tavern=true"	"Boka=true" - "Swift Hibernian Lounge=true" - "WCOU Radio (Tile Bar)=true"	2341,00000	0,043	100,000
5	"Boka=true" - "Grassroots Tavern=true" - "Swift Hibernian Lounge=true"	"Sophie's=true"	2341,00000	0,043	100,000
6	"Boka=true" - "Grassroots Tavern=true"	"Swift Hibernian Lounge=true" - "Sophie's=true"	2341,00000	0,043	100,000
7	"WCOU Radio (Tile Bar)=true"	"Boka=true" - "Grassroots Tavern=true" - "Swift Hibernian Lounge=true"	2341,00000	0,043	100,000
8	"Swift Hibernian Lounge=true" - "WCOU Radio (Tile Bar)=true"	"Boka=true" - "Grassroots Tavern=true"	2341,00000	0,043	100,000
9	"Boka=true" - "Grassroots Tavern=true" - "Swift Hibernian Lounge=true"	"WCOU Radio (Tile Bar)=true"	2341,00000	0,043	100,000
10	"Boka=true" - "Grassroots Tavern=true"	"Swift Hibernian Lounge=true" - "WCOU Radio (Tile Bar)=true"	2341,00000	0,043	100,000
11	"WCOU Radio (Tile Bar)=true"	"Boka=true" - "Grassroots Tavern=true" - "Sophie's=true"	2341,00000	0,043	100,000
12	"Sophie's=true" - "WCOU Radio (Tile Bar)=true"	"Boka=true" - "Grassroots Tavern=true"	2341,00000	0,043	100,000
13	"Boka=true" - "WCOU Radio (Tile Bar)=true"	"Grassroots Tavern=true" - "Swift Hibernian Lounge=true"	2341,00000	0,043	100,000

Fig.7. Generated rules by Tanagra tools. 1992 rules were extracted, we illustrate the 13 first rules.

The Fig.7 shows some rules generated by Tanagra software. Once the rules are created, we can then review and filter them this can be done in several ways using both graphs and by simply inspecting the rules. For example, if we fixed α=0.009, β=0.5 and γ=2 then rules=apriori(file, parameter=list(supp=0.009, conf=0.5, minlen=2))



(a)

(b)

Fig.8. Visualization of: (a) scatter plot for 6 rules; (b) two-key plot

Table.5. Rules generated by parameters supp=0.09, conf=0.5, minlen=2

Rules	Support	Confidence	Lift
[1] {Live. Music, People. Watching => {Pizza}	0.095819	0.785714	4.849462
[2] {People. Watching, Pizza => {Live. Music}	0.095819	0.55	7.428235
[3] {Scenic. at .Night, Tourist. Trap => {Pizza}	0.095819	0.55	3.394624
[4] {Pizza, Scenic at. Night} => {Tourist. Trap}	0.095819	0.5	4.519685
[5] {Live. Music, Tourist. Trap} => {Pizza}	0.095819	0.647059	3.993675
[6] {Guilty. Pleasure, Tourist. Trap} => {Pizza}	0.10453	0.571429	3.526882

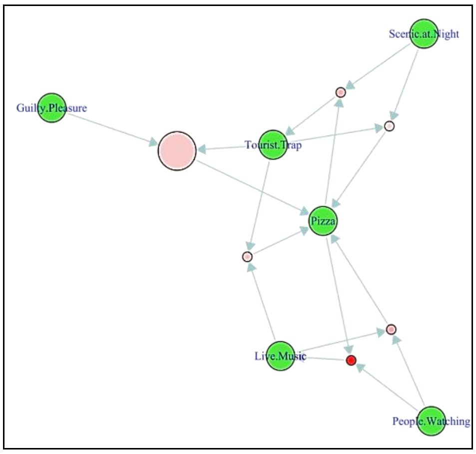
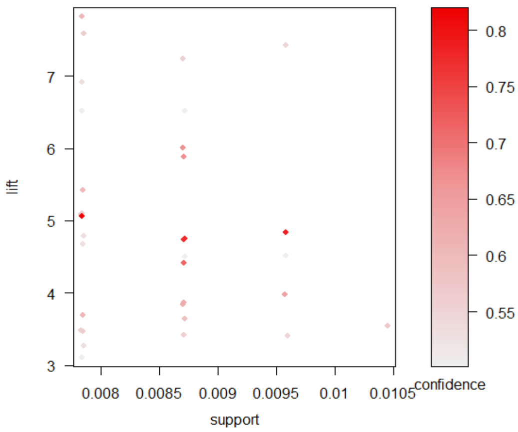
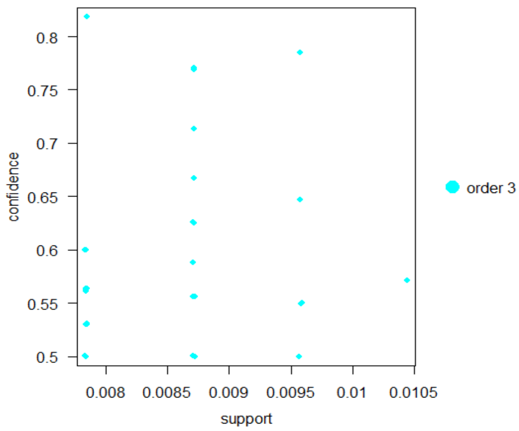


Fig.9. Illustration of a graph with 6 rules

The Fig.8 presents graph of 6 rules without layouts restrictions. In this case, the Vis AR tools [24] was not able to illustrate all rules. Also, in the above graph, the size of the nodes is based on the highest Support and the color signifies the highest Lift. The incoming lines correspond to the LHS and the outgoing lines characterize the RHS.



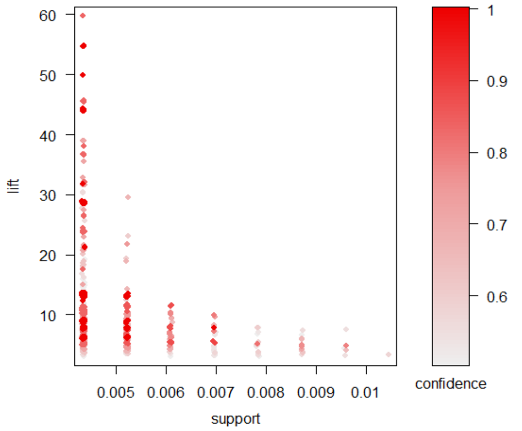
(a)



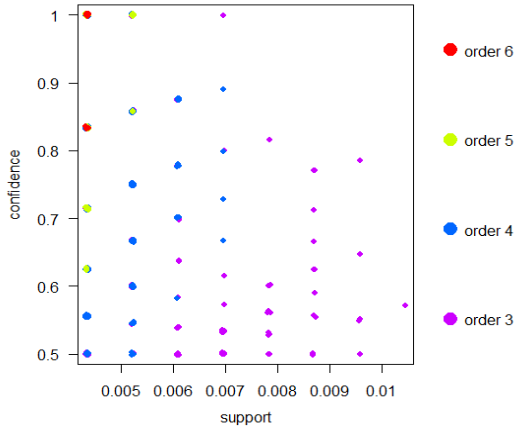
(b)

Fig.10.Visualization of: (a) scatter plot for 33 rules; (b) two-key plot

By command rules=apriori (file, parameter=list (supp=0.007, conf=0.5, minlen=2)), we generated 33 rule(s) shown in Fig.9. While, by instruction “apriori (file, parameter=list (supp=0.004, conf=0.5, minlen=2))”, we generated 767 rules illustrated in Fig.10 in the form of scatter plots and two-key plot, where the number of itmset are 3, 4, 5 and 6 that are presented by “order”. The Fig.12 shows the grouped matrix of 763 rules. This representation explains better relation between rules and item sets.



(a)



(b)

Fig.11. Visualization of: (a) scatter plot for 767 rules; (b) two-key plot

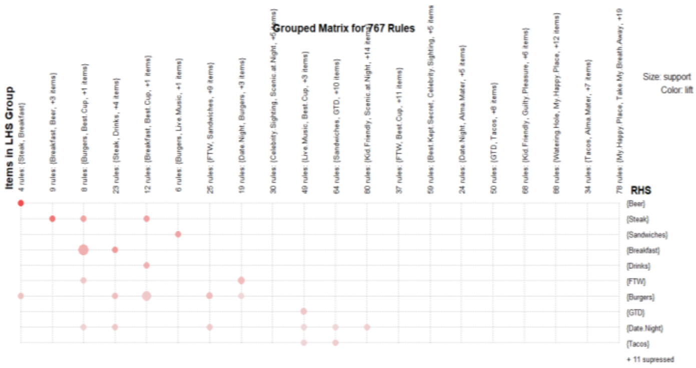


Fig.12. Grouped matrix for 767 rules (LHS→RHS)

The display of association rules using a graph structure permits to improve the visualization of association rules by customizing the available layout. This can ameliorate the illustration and augment the interoperability while visualizing huge sets of rules.

Table.6. Some rules with descriptions

Rules	Sup.%	Conf.%	Lift	Description
{Guilty .Pleasure, Celebrity Sighting} => {Tourist. Trap} => {Pizza}	0.0871	0.6667	6.02	This rule informs us that the tourist visitors' comment on Guilty Pleasure and like Celebrity Sighting is about 8.71 % with 66.67% precision.
{Best .Kept. Secret, Pizza} => {Best. Cup}	0.0871	66.667	5.88	This rule indicates that people who comment best kept secret and favorites as meal pizza like Best Cup with 8.7% of reliability and 66.67% precision.
{Pizza, Scenic. at. Night} => {Tourist. Trap}	0.0958	0.5	4.52	This rule demonstrates that 9,58% of persons who likes to watch Scenic. at. Night favorites to take pizza are tourists with 50.00% precision
{Live. Music, People. Watching} => {Pizza}	0.0785	0.958	4.85	This rule shows that people who favorite place to hear live music. And Highlights generally ask for eating Pizza with 7.875% of reliability and 95.8% precision
{Take My Breath Away} => {Best Kept Secret}	0.114	1	2.44	This rule shows that more than 11% of the users of the Gowalla site who like cinema and look for it, and comment on, also prefer to frequent the green spaces with 100% precision
{Scenic at Night} => {My Happy Place}	0.091	1	3.14	This rule shows that about 9% of the users of the site like places that guarantee beauty and charming landscapes such as they look to New York at night with its magnificent lights then they comment "My happy place". This rule appears with precision 100%.
{Celebrity Sighting, WTF => {Take My Breath Away}	0.091	1	4	This rule indicates that 9.10% of persons who like Celebrity Sighting and stay in the middle of open space comments "take my breath away" with 100% precision
{Alma Mater, Pizza} => {Burgers}	0.114	1	2.933	This rule indicates that people who visitors of Alma Mater that favorite Pizza as meal also takes Burgers with11.4% of reliability and 100% precision.

In addition, the graph representation is appropriate for visualizing rules that contain a lot of consequents and antecedents. We have discovered over 14,906 rules, some of them are interesting rules while several rules might be useless. Thus, selecting the threshold for support and confidence is an important step to get the stronger rules. Some of these rules are presented and discussed in Table.6.

3. CONCLUSION

This paper investigated user's activities in a complex social network platform Gowalla. A new framework-based association rule has been designed to investigate users' activities. Our system consists of four main steps, i.e., preprocessing data, data-matrix preparation of both users-location and user-comments, applying the Apriori algorithm to extract associations rules, and finally interpret rules. The experimental results on a publicly available Gowalla database show that the presented scheme is a very powerful and efficient tool in analyzing LBSNs. Also, the proposed model with association rules, can be considered as a useful tool that can be added to existing state-of-the-art techniques. As a future work, we will consider the application of vertical and horizontal clustering on data matrix (user-location/user-comment) before applying association rules.

REFERENCES

[1] M. Ye, P. Yin, W.C. Lee and D.L. Lee, "Exploiting Geographical influence for Collaborative Point-of-Interest Recommendation", *Proceedings of International ACM SIGIR Conference on Research and development in Information Retrieval*, pp. 325-334, 2011.

[2] C. Yang, M. Sun, W.X. Zhao, Z. Liu and E.Y. Chang, "A Neural Network Approach to Jointly Modeling Social Networks and Mobile Trajectories", *ACM Transactions on Information Systems*, Vol. 35, No. 4, pp. 1-28, 2017.

[3] J. Bao, Y. Zheng, D. Wilkie and M. Mokbel, "Recommendations in Location-Based Social Networks: A Survey", *Geoinformatica*, Vol. 19, No. 3, pp.525-565, 2015.

[4] H. Gao, J. Tang, X. Hu and H. Liu, "Content-Aware Point of Interest Recommendation on Location-Based Social Networks.," *Aaai*, Vol. 15, pp. 1721-1727, 2015.

[5] Y. Qiao, X. Luo, C. Li, H. Tian and J. Ma, "Heterogeneous Graph-Based Joint Representation Learning for Users and POIs in Location-Based Social Network", *Information Processing and Management*, Vol. 57, No. 2, p. 102151-102163, 2020.

[6] W.X. Zhao, F. Fan, J.R. Wen and E.Y. Chang, "Joint Representation Learning for Location-Based Social Networks with Multi-Grained Sequential Contexts", *ACM Transactions on Knowledge Discovery from Data*, Vol. 12, No. 2, pp. 1-21, 2018.

[7] S. Feng, X. Li, Y. Zeng, G. Cong and Y.M. Chee, "Personalized Ranking Metric Embedding for Next New Poi Recommendation", *Proceedings of International Conference on Artificial Intelligence*, pp. 2069-2075, 2015.

[8] X. Chen, Y. Zeng, G. Cong, S. Qin, Y. Xiang and Y. Dai, "On Information Coverage for Location Category based Point-of-Interest Recommendation", *Proceedings of AAAI Conference on Artificial Intelligence*, pp. 1-8, 2015.

[9] H. Li, Y. Ge, R. Hong and H. Zhu, "Point-of-Interest Recommendations: Learning Potential Check-Ins from

Friends", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pp. 975-984, 2016.

[10] D. To, D. Si and Y. Chen, "Traveler's Next Activity Predication with Location-Based Social Network Data", *Proceedings of International Workshop on Prediction of Human Mobility*, pp. 15-23, 2019.

[11] A. Likhyani, D. Padmanabhan, S. Bedathur and S. Mehta, "Inferring and Exploiting Categories for Next Location Prediction", *Proceedings of International Conference on World Wide Web*, pp. 65-66, 2015.

[12] I. B. Shem Tov and S. Bekhor, "Extracting Travel Demand for Emergency Situations using Location-Based Social Network Data", *Transportation Research Procedia*, Vol. 45, pp. 111-118, 2020.

[13] X. He, L. Liao, H. Zhang, L. Nie, X. Hu and T.S. Chua, "Neural Collaborative Filtering", *Proceedings of International Conference on World Wide Web*, pp. 173-182, 2017.

[14] C.C. Aggarwal, "*Data Mining: The Textbook*", Springer, 2015.

[15] P. Ristoski and H. Paulheim, "Semantic Web in Data Mining and Knowledge Discovery: A Comprehensive Survey", *Journal of Web Semantics*, Vol. 36, pp. 1-22, 2016.

[16] R. Agrawal, T. Imielinski and A. Swami, "Mining Association Rules Between Sets of Items in Large Databases", *Proceedings of International Conference on Management of Data*, pp. 207-216, 1993.

[17] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules", *Proceedings of International Conference on Very Large Data Bases*, pp. 487-499, 1994.

[18] Z. Zou, X. Xie and C. Sha, "Mining User Behavior and Similarity in Location-based Social Networks", *Proceedings of International Conference on Parallel Architectures, Algorithms and Programming*, pp. 167-171, 2015.

[19] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and S. Y. Philip, "A Comprehensive Survey on Graph Neural Networks", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 32, No. 1, pp. 4-24, 2020.

[20] F. Hao, J. Zhang, Z. Duan, L. Zhao, L. Guo and D.S. Park, "Urban Area Function Zoning Based on User Relationships in Location-Based Social Networks," *IEEE Access*, Vol. 8, pp. 23487-23495, 2020.

[21] F. Berzal, I. Blanco, D. Sanchez and M.A. Vila, "A New Framework to Assess Association Rules", *Proceedings of International Symposium on Intelligent Data Analysis*, pp. 95-104, 2001.

[22] M. Hahsler, "A Rules Viz: Interactive Visualization of Association Rules with R", *The R Journal*, Vol. 9, No. 2, pp. 1-13, 2017.

[23] B.V Almende and B. Thieurmel, "Vis Network: Network Visualization using 'Vis. Js' Library", Available at https://cran.r-project.org/web/packages/visNetwork/index.html, Accessed at 2017.

[24] L. Yang, "Visualizing Frequent Item Sets, Association Rules, and Sequential Patterns in Parallel Coordinates", *Proceedings of International Conference on Computational Science and Its Applications*, pp. 21-30, 2003.