Journal of Engineering and Technology
DergiPark AKADEMİK

Journal of Engineering and Technology

Twitter Users' Emotion, Emoticons and Scaling Metrics Based Categoric Interaction Analysis

Hafzullah İŞ [1*], Taner TUNCER [2]

[1] Department of Information Technology, Batman University, Batman, Turkey

[2] Department of Computer Engineering, Fırat University, Elazığ, Turkey

ARTICLE INFO

Article history:
Received 26 Nov 2018
Received in revised form
10 Dec 2018
Accepted 11 Dec 2018
Available online 28 Dec 2018

Key words:
Twitter
Quality Scaling Criteria
Semantic and Emoticons Analysis
Interaction Analysis

* Corresponding author.
E-mail address:
Hafzullah.is@batman.edu.tr

ABSTRACT

The popularity and use of social networks has also begun to increase in parallel with the worldwide increasing accessibility and means of access to the Internet. As one of the world's most popular social networks, Twitter is a platform where users are interacting through follow-up, sharing, messaging and appreciation tools, sharing their ideas and emotions in a variety of individual and corporate contexts. Therefore, Twitter is intense, dynamic and always an up-to-date data source. Identifying and correlating the physical and emotional interaction of users can be valuable in political, social, academic and commercial aspects. Users' physical networking with each other and emotional analysis can be done with many tools and applications. The character, tendency and impact analysis of the users can be used in the development of business intelligence applications and in the determination of social strategies.
In this study, a large Twitter user group is divided into four categories: political, Entertainment, Sports, Trade Marks. Then, the physical and emotional interaction of each category was revealed. The Physical interaction metrics determined as centrality, intensity, reciprocity and modularity while emotional interaction metrics were determined as resistance, passion, reach and emotionality. Positive, negative and neutral states of sharing were discussed in emotional measurement. Beside that, emoji-containing tweets have been transformed into texts and are especially included in emotion analysis. After all the metrics were calculated, physical and emotional interaction structures and overlap rates were revealed using "Interaction and Semantic Clustering Based Multinetwork Analysis" method.

1. Introduction

The point reached by today's information technologies, leads to profound changes in the functioning of social dynamics. Strategic Production, Intelligent Cities, Transportation, Industry 4.0, and the Internet of Objects create opportunities that facilitate human life in every aspect of life. the rapid development of Internet access facilities and tools plays an important role in bringing people's communication to digital platforms. Social networks that are rapidly developing and increasing in number of users day by day, provides all kinds of personal and corporate communication opportunities such as acquiring information, training, sharing content, expressing ideas, having fun, communicating, friendship and trading.

Social networks contain a great deal of information about members. By applying data mining techniques on scalable data in these networks, the patterns can be defined, information about the friendship relation, sharing, appreciation and interpretation of the users and characteristics, trends and interaction forms can be obtained. Increased opportunities and diversification of tools to access the Internet lead increase in the number of users of social networks and the frequency of usage. According to "We are Social Global Digital Report 2018" [1] The number of internet users in the world has been 4.021 million. More than half of the world's population is connected to the internete. According to the same report, the main reason for the reaching Internet such widely available rates is the smartphone users. In 2017, more than 200 million people became phone owners for the first time. From 7.6 billion people, 5,135 million of them, or 2/3, have become telephone owners. It turns out that more than half of these phones are smart phones. This is the biggest factor that facilitates access to the Internet.

According to the same report, 3.196 million people use online social networks at least once a month. 2.958 Million of these users, or almost 90% of them, have access to social networks from their mobile devices. According to Statista's January 2018 Social Networks Report [2], Twitter is among the most popular social networks and has over 330 million active users. These are equivalent to 30.6% of all social media users. According to the GlobalWebIndex data [3] users spend an average of 4.5 hours a day on the Internet, while a significant portion of this is devoted to social media use.

Twitter user acceleration is continuously rising. That is why it is very dense, dynamic and up-to-date source of data. The 140-character usage right provided to Twitter users enables them to share ideas, news and announcements, as well as share photos and videos. Twitter members can use the platform for fun, friendship, learning, education, marketing, and delivering ideas to masses. Government, corporations, businesses, organizers and individuals are making plans for policy development, product and marketing based on the trends and demands of Twitter users. Therefore, the identification of the emotional relationships as well as the physical connections of the users and their associated analysis can provide valuable information that can serve different purposes.

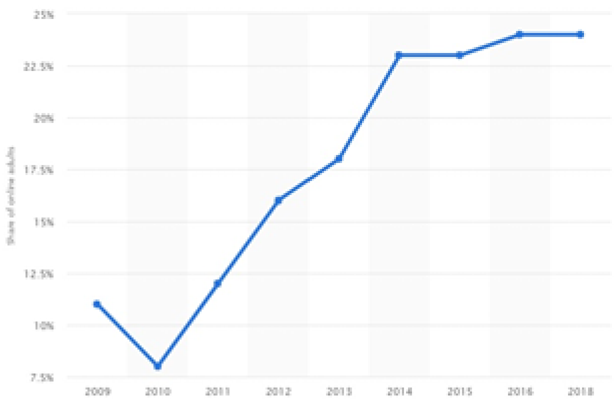

Figure 1. Twitter usage rates between 2009-2017

In this study, the information of Twitter users whose profile data are public, were taken and used by various means. All of the personal information of these people who were taken on Twitter was filtered and the data set was anonymized to conform to the ethical rules in order to protect the privacy of the users. Since the data retrieved with the Twitter APIs contains unnecessary information about a large number of parameters, the filters are applied to obtain faster and more efficient results, and the data set consisting of the average taken 12 columns is reduced to 4 columns.

200 Twitter user profile information and shares have been withdrawn. The last 1000 shares of these users have been examined and processed. As the latest shares of all users are drawn, the results obtained

on the basis of the algorithms and methods applied on the data set are of the most up-to-date. 200 user accounts; sports, politics, trademark and entertainment accounts. The average 50 popular accounts in each category are included in the study. From physical relationships of all accounts, profile physical quality scaling parameters such as density, reciprocity, centrality and modularity are calculated. The physical network interaction rates of all users in each Twitter member category have been subtracted from the quality scaling parameters. A total of 200,000 tweets from all accounts were analyzed in terms of the words they had in order to make emotional analysis base of three class; positive, negative and neutral.

All users in each category were scanned according to their emoji usage patterns and frequencies. And Classified according to have positive, negative, and neutral meaning. Then overlap rates were compared to word-based emotion analysis. The emphasis on expressing emotions in this way is measured in different angles. According to a report by Adweek magazine[18], 92% of internet users use emojis communication. Twitter has used 751 kinds of emoji more than 110 billion times between the years of 2014-16[19].

The emotions of users which have high usage rate are also studied in this paper and correlation of emotional state of the user and the physical structure of profile determined. Emojicons and word-based emotion analysis of all the profiles were made, and emotion interaction rates were revealed with various parameters. These parameters are strength, passion and reach. These parameters were compared with the physical interaction rates of the users and the overlap rates were determined by the "Interaction and Semantic Clustering based Multinetwork Analysis" method.

The rest of the paper is composed from the following parts: In the second part, literature research and academic studies are mentioned. In the third chapter, the acquisition, storage, processing and filtering of the data are discussed. In the fourth chapter, the categorical analysis of the data and the applied methodology are mentioned. Finally, in the conclusion section, the focal points of the article, the results, and the methods that can be used in the development are shared.

2. Related Works

As an effective social media tool for communication and interaction, Twitter has emerged as a pioneering platform for movements affecting world order such as Arab Spring. It plays an important role in representing individuals, corporations and states with their individual and institutional identities. The results of the analysis and the data obtained from this platform are seen to be very valuable in the construction of commercial enterprises and social strategies. In this respect, the usage habits on Twitter and the relationship of the profiles with each other continue to be subject to many academic studies. Studies mainly focus on the analysis of physical ties which including followers/following relationship, sharing, interpretation and emotions which based on vocabulary and emotions analysis. Profiles were examined physically in terms of reciprocity, modularity, centrality, diameter and density.

In social networks, classical community detection methods are usually based on graphical clustering algorithms that use the structural knowledge of group identity. Nodes are topologically divided into classes. In a study by Hsun-Hui Huang and their colleagues, entitled "Community Detection on Social Networks Based on the Clustering of Emotions" [4], community detection was done with a different approach than the analysis of the physical structure between classical users. In this study, physical interaction was only used topologically in clusters, but the sharing of users and the interactions created by these shares were scaled in the identification of the actual groups. Users who create high frequencies in a weighted interaction have come to the conclusion that they are a group.

In the study by Yizhou Sun and his colleagues [5], the importance of focusing on heterogeneous information networks from a homogeneous structure has been emphasized. In addition to this, it has been stated that the unidirectional approach results in misleading results, where different types of objects need to be considered in terms of many parameters in the detection and evolutionary processes of clusters. In a study by Matthew J. Preisendorfer, entitled "Emoji Analysis, Relationships and Trust Modeling in Social Media"[6], was stated that Twitter's 140-character limit was often insufficient in expressing emotions. Emotions that can be used for this reason are very effective in defining the mood of the user. In addition that analyze of emoticons is more effective to make emotion analyse. In his research, he collects a large amount of Twitter data. He analyzed the emoticons to find correlations in social interactions. He examined updated events affect on social media interactions and behaviors. He stated that the usage and frequency of use of emoji can be followed like fingerprints of users. He also worked on emoticons analysis and trust models as user identification method. On the basis of the word, emotional analysis is often said to be inadequate. Instead, he pointed out that matching the emoticons in the user profile and emotions expressing emoticons leads to more accurate results. In paper called "Sentiment of Emojis" by Petra Kralj Novak and his friends stated that 10 billion emoticons were used in the last 2 years in Twitter, which is an effective expression way in transmission of emotion and thought [7]. 1.6 million tweets have scanned and the most used 751 emoticons found in 4% of these tweets. He stated that there is no difference in the positive, negative and neutral expressions of feelings in the 13 different language based scans.

There are a lot of valuable works on Twitter about analysis of emotions[8-12]. In a study by Jose Antonio Iglesias [13], one of the studies that forms the basis of physical linkage analysis of the methodology used in our study, the current social network analysis has shown that academic work focuses on users' interactions, information sharing, useful information about themselves, and associations. However, from the analytical point of view, it is stated that the profile information of the users, which they are not sufficient, contains very useful information and is worth to checking. In their articles, they have proposed a fuzzy system-based approach that automatically analyzes the Twitter users of certain communities according to their profile information and the classes they belong to. This approach can reveal common features by examining thousands of Twitter profiles at the same time. The basis of the approach is the large data analysis.

3. Data Collection and Observation

This section describes the process of collecting, storing, filtering and anonymizing data that collected from the Twitter social network. All the applications and tools which used to prepare the obtained data for analysis are discussed in detail.

a. Collection and Filtering

Depending on the nature and scope of the academic work the type, quality, size and source of the data to be withdrawn from social networks can vary. There are many web-based applications that can be used to extract data from Twitter. It is important that the collected data to be updated, comprehensive and not lost its integrity. Applying filters on a filtered data can negatively affect the result of the work because it breaks the quality of the data. There are too much applications such as Import.io, NodeXL and Netvizz which can also be used for Facebook, Instagram and RSS data capture. However, since it is very important that the method to be applied on the extracted data must collect in proper format, the application of Netlyticbased on Twitter Rest API V1.1 was used [14]. Netlytic is a data collection and visualization tool that allows you to conduct highly comprehensive analyzes on captured data. The use of Netlytic was deemed appropriate if the data coverage is too low. However, it has been tested that the use of NodeXL is correct when analyzing larger volume datasets. Since our data set is quite large in the first instance, the ELK stack model of Elastic Search has been used for storage and NoSQL has been used to process the data and facilitate large data processing. The data taken with Netlytic has been reduced to very low dimensions because it is filtered by deleting Links, Publication Date, Author, Title, and so on. If the data volume is larger, processing with Hadoopta MapReduce ensures more efficient results. However, as we have filtered and normalized the data we have drawn, it has been reduced to a lower size. There was no need to be hosted in such systems; that is why a high-performance PC with an I7 processor, 16 GB RAM, 1 TB HDD and 240 GB SSD disk capacity used for data store and analyze.

Filtered and cleaned data is stored in CSV format as clean text. All the algorithms required for network analysis can be applied on the data in this form. Data mining applications such as WEKA, SPSS, R and Gephi give successful results in analysis. From 200 Twitter users' profiles, periodically once a week for 3 months with Netlytic data was taken and saved to the system in csv format. The withdrawal of all shares of users will result in serious time, performance and storage needs, that is why the last 1000 tweets of users taken at each time. The average data taken for a user is 1 MB.

Table 1. Periodical pushed data from user accounts

Number of User	*Number of Tweet*	*Time Period*	*Data Size*
200	1000	12 Week, 3 Month	800 MB / Month

Amount of total crude data:200*4*3*1 MB=2400 MB
Amount of filtered data: 200*4*3*620 KB= 1450 MB
Parameters such as links, author, pubdates, which do not require the use of raw raw data and which cause unnecessary space and cause performance loss, have been cleared.

b. Processing and Analysis

Many metrics are used to determine the physical network connections of users whose data received from Twitter's profiles[15]. The analysis of the physical network connections of the data obtained with Netlytic was bilaterally approximated. These are who mention who and Who Reply who. There are many metrics used to scale the physical quality of the network such as diameter, density, reciprocity, centrality, modularity. These parameters calculated by using users' last 1000 tweets mention and replies statistics. In addition, the original names, posters with ties, ties, and follow-up ratios of the last shares were extracted and analyzed for use in determining the interaction rates of the ties. Each user's interactions with other users are examined by using Fruchterman Reingold that centralized personalities, DrL Layout that long distance connections are transferred in the background to highlight clusters and visualized with LGL layout.

In figure 2, a Twitter user, who is one of 200 users, has determined the percentage of users who interact most according to network scaling metrics. In addition to the network physical quality scaling parameters, emotional analysis of the users was done on the basis of word and emoticons. For this, all the words and emoji used by the users are scanned and analyzed. In the analysis of the words, Support Vector Machines were used in the WEKA data mining platform. 30% of the data from all profiles were used as training data and 70% as test data.

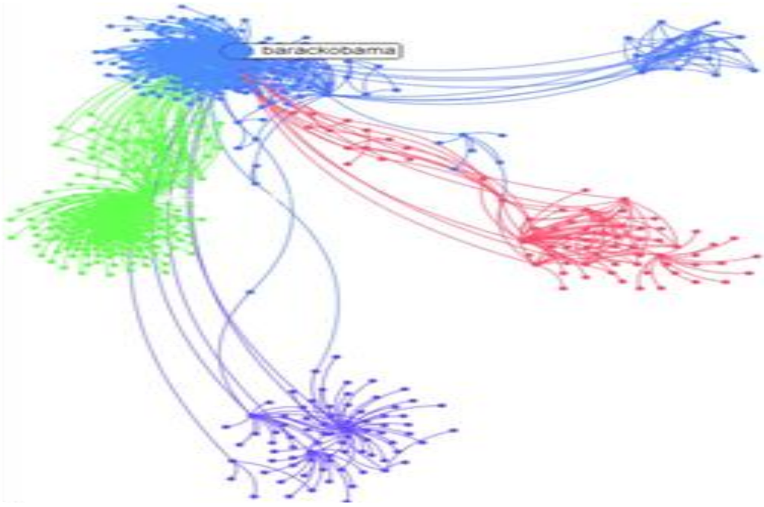

Figure 2. Weighted 4-Cluster DRL interaction network of popular user

All shares of 200 users have been evaluated. In these evaluations, the words shared in sharing were classified according to their being active and passive and expressing happiness and unhappiness.

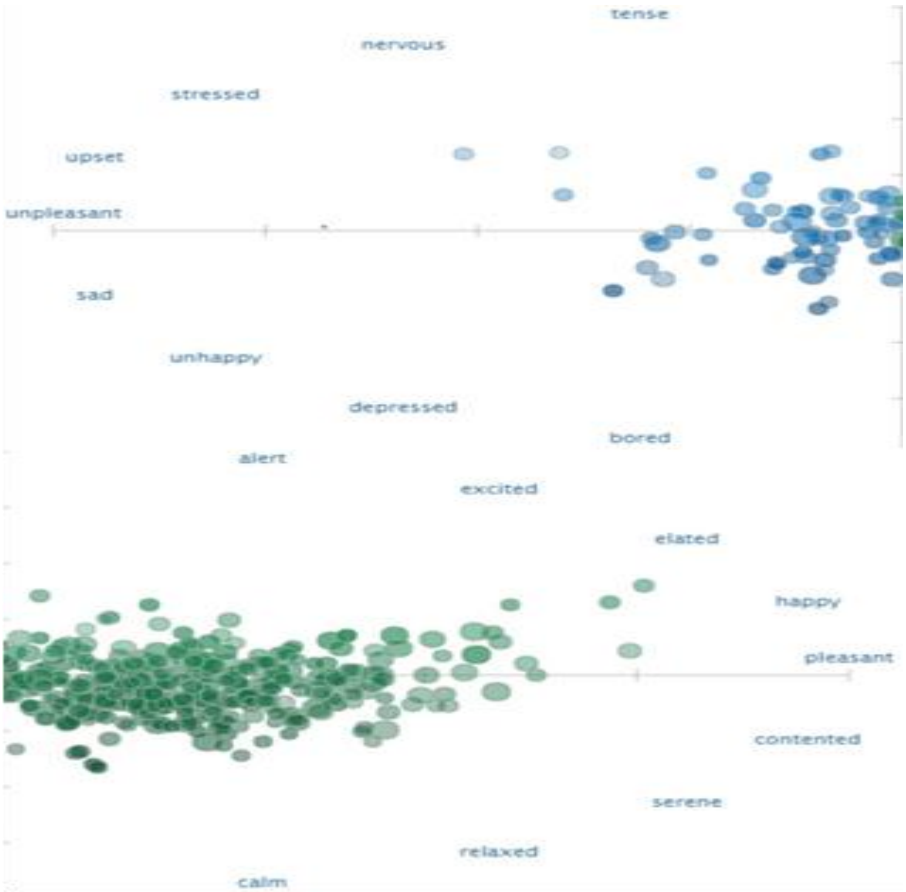

Figure 3. Words-based emotion analysis

Periodic emotional analysis of all users' periodic shots was performed on 4 different scales. These scales; Comfortable, Happy, Unhappy, Sad is in the form. As a result, these shares are combined into two classes, positive and negative. Figure 5 shows that the user has a positive emotional state of over 60%. Tweet Sentiment Visualization [16] was used to analyze and visualize this work.

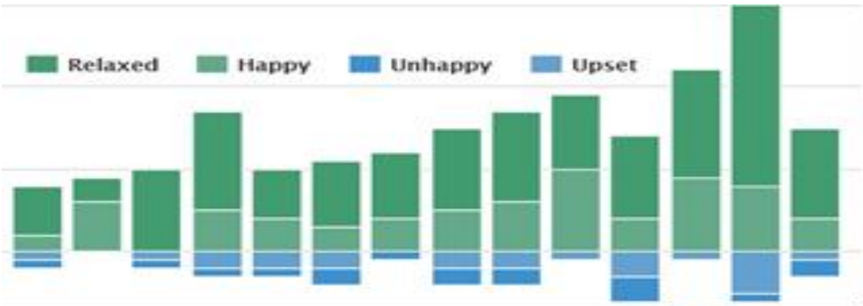

Figure 4. Emotional analysis of users in monthly time envelope

The results of the emotional analysis obtained were revealed on a category basis and categorical moods of the users were revealed. Physical metrics are compared and correlated. The meaning of physical and emotional interactions Metrics:
Density: The ratio of the number of up-to-date links in the network to the total number of possible connections.
Centralization: It determines the centroid nodes in the network.
Reciprocity: It is the ratio of all the correlations of the number of relativistic associations.
Modularity: The ratio of users in a network can interact with each other more than others and can create possible subsets.
Strength: In the social media, it refers to how much the subject matter account has been discussed. By dividing Existing mentions to all possible mentions can be handled.
Sentiment: The ratio of positive impressions to negative impressions.
Passion: The wager accounts for how often people talking about the account can speak.
Reach: It is the measure of the range of influence. It is the ratio of those who speak directly about an account. Scaling metrics formulas[15]:

$$Connects = \frac{N*(N-1)}{2} \quad (1)$$

$$Density = \frac{B}{\mathrm{P}} \quad (2)$$

$$Centrality = \sum_{s:=v:=t\in \mathrm{V}} \frac{\sigma\mathrm{st(v)}}{\sigma\mathrm{st}} \quad (3)$$

$$\text{Centrality Degree} = di * (N-1) \quad (4)$$

$$\text{Reciprocity} = L^{<->}/L \quad (5)$$

$$e_{ii} = |\{(u,v): u \in V_i, v \in V_i, (u,v) \in \mathrm{E}\}|/|\mathrm{E}| \quad (6)$$

$$a_i = |\{(u,v): u \in V_i, (u,v) \in \mathrm{E}\}|/|\mathrm{E}| \quad (7)$$

$$Modularity = \sum_{i=1}^{c} \left(e_{ii} - a_i^2\right) \quad (8)$$

Generally emoticons do not pass texts when emotional analysis of the shares is done. As a result they can not be semantically analyzed. Emoticons can be overlooked in situations where even a word is very important to express that a share is positive, negative or neutral.

However, studies have shown that emoticons that shared in tweets make people feel emotions as possible as in real life. Emoticons have begun to be used with a growing popularity in the social media in recent years in expressing emotions. In particular, Twitter's inadequate number of characters in a conversation is one of the most encouraging elements of emoticons use. Especially since the number of characters used in Twitter is insufficient, the use of emoticons is important for expressing emotions. Studies [17-19] have stated that emoticons are used for communication in a very dense manner and provide a very clear sense of completeness when emotions are expressed. The most influential word of Oxford's choice in 2015, laughing emotions, has been a sign of their popularity [20]. Emoticons should not be ignored in order to make realistic analyzes. Shares containing emoticons have been identified with Netlytic. The emoticons in these shares were converted to text using emoticons translate [21]. Later, Sentigem's emotional analysis program[22] used to re-processed related emoticons expressed tweets as positive, negative and neutral situations. Re-evaluations of emoticons containing emotions into emotional analysis play a crucial role in achieving realistic outcomes.

Table 2. User's category based physical metric measurements

Metric	*Politics*	*Entertainment*	*Sports*	*Trade Mark*
Density	0.0013439	0.00167189	0.002521990	0.00100971
Reciprocity	0.0287868	0.00715563	0.00481463	0.011881188
Centralization	0.15738962	0.25149489	0.42233328	0.415662287
Modularity	0.75121542	0.57872088	0.395469055	0.347034119

Table 3. User's category based emotional metric measurements

Metric	*Politics*	*Entertainment*	*Sports*	*Trade Mark*
Strength	%90	%52	%66	%16
Reach	%11	%50	%55	%12
Passion	%68	%22	%21	%23
Sentiment (P|N)	12|1	11|1	2|1	4|1
Interaction Value	~%75	~%75	~%75	~%25

As shown in Table 3, Politics, Entertainment, Sports and Trade Mark based four categories contructed and in each category 50 person Twitter accound analyzed. For each category of users 4 physical scaling metrics calculated and their avarage results found. These metrics are: Density, reciprocity, centralization and modularity. In calculation all metrics Netlytic was used.

According to Table 4, 200 user classified in 4 categories and for each category the emotional scaling metrics calculated. These metrics are strength, reach, passion and sentiment. SocialMention was used to do these calculations[23].

c. Proposed Approach and Methodology

Analyzes made are handled as a whole in this part of the article. The methodology of the methods applied in the analysis is emphasized. There are many ways to scale the effectiveness of social media accounts. However, it is important that it is a fast and performance method. In the study of social network accounts, the theory of homophily what mean people of similar character attract each other is the cause of basic interaction. "Normalized Absolute Information" can be used for the value measurement. The investigation of the relationship network of a Twitter user is a topic that can be dealt with in many ways. These; such as the number of followers, likes, shares, the structure of the bond with the people they interact with, and the type of sharing. All of these qualities are very important parameters. In this article many of the users are excluded because the interaction links of the users are analyzed. The focus is on reciprocity, density, centrality and modularity, which form the basis of interaction. These metrics, which can be applied in every situation according to the static and dynamic state of the network, are seen as valuable parameters.

In addition to the network of physical interaction, the network of emotional interaction has also been examined in consideration of many parameters. These parameters are; strength, reach, passion and sentiment. In the detection of the sentiment, the cases of positive, negative and neutral sharing of the user are discussed. As it is known, sometimes even a single word is very important in correct classification. In general, emoticons are ignored when classification is made. This means one-sided and incomplete analysis. However, in this study all tweets with emoticons were scanned and sorted. Emoticons have been replaced by textures that have their own meaning. Once this is done, all tweets have been reviewed again for semantic analysis. If there are tweets with 5% emoticons and they are excluded from the assessment, the result will be an average deviation of 5%. All user accounts and shares that are taken from physical and emotional evaluation and categorically analyzed are analyzed by "Interaction and Semantic Clustering based Multinetwork Analyze" method. In this method, all shares are categorized. The physical metric measures and the emotional metric measures of each category are compared and the overlapping and overlapping situations of the interactions are tried to be clarified. The positive and negative acceptance threshold values for each metric are different. If the threshold value is appropriate, the profile will have a metric of 1, meaning that the interaction rate is high. If the threshold value is not appropriate, the metric will be 0. Physical metrics constructed from 4 different parameters.

Table 4. Category based metrics thresholds

Metric	*threshold*	*Politics*	*Entertainment*	*Sports*	*Trade*
Density	>0.002	0	0	1	0
Reciprocity	>0.01	1	0	0	1
Centralizaton	<0.4	1	1	0	0
Modularity	>0.5	1	1	0	0
Interaction	-	%75	%50	%25	%25

The contribution rate of each of these is 25% to the physical interaction value of the network. The percentage of the total of four different metric threshold values represents the "Interaction Value". In emotional interaction, each metric value contribution is considered as 25% to the interaction value if it is value above 50% threshold. Reciprocity is the proportion of bonds that show two-way communication based on the total number of existing relationships. Measured by determining the number of total reciprocals in the network.

Taking a higher value means there is strong two-way interaction. A low reciprocity value indicates that many enterprises are unilateral. Interaction is high if knowledge transaction is based two-way flow. Political and commercial accounts were predominantly reciprocal, while entertainment accounts (magazine programs, talkshow ..) were low. In political accounts with high reciprocity rates, clinging to ideas and reaching over 90% of the interaction on ideas has been one of the criteria for parallel physical and emotional interaction. In political accounts with a high level of physical interaction, emotional interaction is also quite high. The semantic analysis is positive at a rate of 12/1. This proves that there is a lot of interest in positive sharing in political calculations.

In popular accounts, the reply is too low, mention is high. The most mentioning or responding users in inferiority centrality come to the forefront. Out-of-center centrality refers to high active twitter users who interact with their environment in the network. While the centrality of sports and trademark accounts is quite high, the centralization of political and entertainment accounts is very low. The high center of gravity means that the network is authoritarian, connected to a certain family, and has a high level of auto control. If centrality is low, information on the network means little control of free circulation.

Density is the ratio of available connections to the total number of possible connections in a network. In other words, it is calculated by dividing the number of existing links to the number of possible links. If the density value is 1, it can be said that all users are connected to each other. This measure is complementary to the diameter because it also expresses the speed of information flow.

Centralization refers to the presence of people who act as bridges in communication on a network. The withdrawal of these contacts causes the lack of communication on the network to weaken. If the centrality is high, the value is 1; if the centrality is low, the value is close to 0. A low decentralization means less commitment. If centralism is high, there is institutionalism, control, centralism. If it is low; comfort, freedom, flexibility and diversity.

Modularity means that the interaction is concentrated among the groups, does not reflect the network in general, and is high in the structure of the networks that the cluster stands for. The high level of modularity means that they are a tight base that is fond of each other, values each other and communicates well. However, apart from the groups, it can be said that all of you are low.

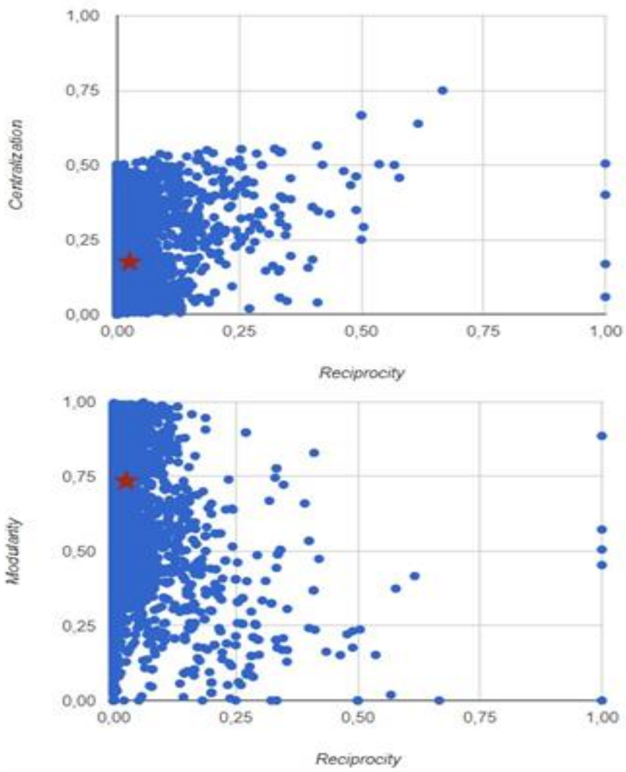

Figure 5. Twitter Test User Reciprocity Based Modularity and Centralization Correlated Interaction Fusion

When the overlap rates of physical and emotional interaction are examined;

- In politic accounts, physical interaction ratio is 75%, paralelly emotional interaction ratio is also 75%i
- In entertainment accounts, physical interaction ratio is 50%, bu emotional interaction ratio is 75%.
- In sports accounts, physical interaction ratio is 25% but in contrast emotional interaction ratio is 75%.
- In trade mark accounts, physical interaction ratio is 25% and parallelly emotional interaction ratio is 25%.

However, the physical interaction in sportive calculations has been the result of no effect on emotional interaction. Ultimately, the cumulative effect of physical and emotional interaction was achieved at a rate of 67.5%.

4. Discussion

More effective methods should be developed for the evaluation of metrics that used to measure the emotions of users. The method used here is based on the study of experimental achievements. The given metrics result calculated according to time series simple moved average method. But, to get the more effective result, weighted moving average time series must be applied to dataset.

5. Conclusion

The Internet continues to change the forms of communication and interaction of mankind quite quickly and profoundly. The tremendous increase in accessibility and means of access to the Internet has increased the number of users astronomically. As a natural consequence of this, social media users are increasing in number and usage rates. Twitter is one of the most active social media platforms in the world. As a result of examining the relationships of users in this platform, physical interaction can be revealed. These inferences are possibly can be the base of many valuable works. However, these implications need to be transformed into a simple commercial business intelligence, or more basis for a comprehensive corporate and social strategic policy. It is also essential to identify users' trends, characters and interaction networks. For this, it is important to conduct physical and semantic analysis of profiles and shares. Performing emotional analysis of user tweets and associating them with physical network interactions will help identify more valuable patterns.

In the study; the physical network interaction structure of an intensive Twitter user data set has been determined using different metrics. Before semantic analysis, emoticons in the users' tweets were converted into text. Thus, semantic analysis of all tweets is done as a whole. Afterwards, the emotional interaction scaling metrics of all tweets were calculated. Finally, categorical basis, physical interaction values of users and emotional interaction values are compared.
As a result, According to analysis, found that with 67.5% ratio the results of network and semantic analysis support each other. And it is determined that users networks and their moods are influenced by each other. Diversification of the dataset used in the study, more effective and valuable results can be obtained by increasing the physical and emotional metrics.

Acknowledgement

Data in this paper was previously published as a conference paper in "International Engineering and Technology Symposium (IEST'18), May 2018".

6. References

[1] S. Kemp, "We are Social Global Digital Report 2018" (Online). Available: https://wearesocial.com/uk/blog/2018/01/global-digital-report-2018 last retrieved on March 30, 2018.

[2] Statista, "Most Famous Social Networks 2018" (Online). Available: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ last retrieved on March 30, 2018

[3] Global Web Index, “Global Social Media Research Summary 2017”(Online) Available: http://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/ ,last retrieved on March 30, 2018.

[4] H. H. Huang, H. C. Yang, “Semantic Clustering-Based Community Detection in an Evolving Social Network” in 6. International Conference on Genetic and Evolutionary Computing, IEEE, 2012.

[5] Y. Sun, J. Tang, J. Han, M.Gupta, and B. Zhao. “Community Evolution Detection in Dynamic Heterogeneous Information Networks” in Proc. KDD MLG, 2010.

[6] M. J. Preisendorfer, “Social Media Emoji Analysis, Correlations and Trust Modeling”, MSc. Thesis, SUNY Polytechnic Institute, 2018.

[7] P. K. Novak, J. Smailovic, B. Sluban, I. Mozetic, “Sentiment of Emojis” arXiv:1509.07761v2 [cs.CL], 2015.

[8] L. Zhao, C. Zeng, “Using Neutral Networks to Predict Emoji Usage from Twitter Data”, Stanford, 2017.

[9] R. Socher, A. Perelygin, J. Y. Wu, J. Chuang..,”Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank” Stanford, 2015.

[10] G. Guibon, M. Ochs, P. Bellot, “From Emojis to Sentiment Analysis”, WACAI, France, 2016.

[11] A. Bartolome, M. Bock, R. K. Vinayagam, R. Krishnamurthy, “Sentiment and Topic Analysis”, CS 6604 Digital Libraries, F. Report, 2017.

[12] A. Yeşilyurt, Ş.E. Şeker, “Metin Madenciliği Yöntemleri ve Twitter Duygu Analizi”, YBS Sözlük, Cilt 4, Sayı 2, 2017.

[13] J. A. Iglesias, A. G. Cuerva, A. Ledezma, A.Sanchis, “Social Network Analysis: Evolving Twitter Mining” in International Conference on Systems, Man, and Cybernetics, SMC, Budapeşt, Hungary, 2016.

[14] Netlytic, (Online), https://netlytic.org/home/?page_id=10834, last retrieved on March 30, 2018.

[15] Wikipedia, (Online), https://wikipedia.org/wiki/Metrics/, last retrieved on March 30, 2018.

[16] Tweet Sentiment Vizualization (Online), https://www.csc2.ncsu.edu/faculty/healey/tweet_viz/tweet_app/ last retrieved on March 30, 2018.

[17] B. Hutchins, “The Emoji Infographic: Stats to Back Up Your Obsession”, https://www.meltwater.com/blog/the-emoji-infographic-stats-to-back-up-your-obsession/ last retrieved on: March, 2018

[18] B. Shaul, ”Report: 92% of Online Consumers Use Emoji (Infographic)“, http://www.adweek.com/digital/report-92-of-online-consumers-use-emoji-infographic/, last retrieved on 2018

[19] N. Shah, “Introducing emoji targeting”, https://blog.twitter.com/marketing/en_us/a/2016/introducing-emoji-targeting.html, last retrieved on 2018.

[20] Oxford Dictionary (Online), “Oxford Dictionaries Word of the Year 2015”, http://blog.oxforddictionaries.com/2015/11/word-of-the-year-2015-emoji/, last retrieved on 2018.

[21] Emoji Translator (Online), http://superemojitranslator.com /emoji-translate, Last retrieved on March, 2018.

[22] Sentigem (Online), http://sentigem.com/#!, Last retrieved on March 2018.

[23] SocialMention (Online), http://www.socialmention.com, Last retrieved on March, 2018.