Int. J. Nonlinear Anal. Appl.
Volume 12, Special Issue, Winter and Spring 2021, 1143-1158
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2021.5594

Measuring and analyzing charisma on twitter using combination weighting and TOPSIS method

Reza Mohamaddoust[a], Javad Mohammadzadeh[a,*], Madjid Khalilian[a], Alireza Nikravanshalmani[a]

[a] *Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.*

(Communicated by Ehsan Kozegar)

Abstract

In this research, charisma has been measured at different levels of online social networks including charisma at the level of messages, individuals, and communities. First, the charisma-associated features have been extracted and then weighted by hybrid proposed methods. Eventually, measuring and ranking charisma has been investigated through technique for order preference by similarity to ideal situation (TOPSIS) as one of the leading multi-criteria decision-making methods. Through the proposed approach, the charisma of different messages, individuals, as well as implicit and explicit communities can be measured, ranked, and compared. In this research, eight datasets were collected from Twitter with different and diverse features. The results indicated that the charismatic messages and individuals of each dataset have been chosen properly and logically. Further, a method has been presented to measure the rate of charisma in every community which can be employed for comparing communities and predicting behavior in online communities.

Keywords: Charisma, Online social networks, Online community analysis, TOPSIS, Twitter

1. Introduction

A virtual community develops when like-minded users in social media join each other and begin to interact with each other. Communities are formed either explicitly or implicitly. In the explicit type of membership, the person is aware of their membership. For example, the membership occurs as subscribing to or following a user account. Nevertheless, if the individuals are not aware of the membership of each other, they form an implicit community. For example, when there is a specific

*Corresponding Author: Javad Mohammadzadeh

Email addresses: r.Mohamaddoust@pnu.ac.ir (Reza Mohamaddoust), j.mohammadzadeh@kiau.ac.ir (Javad Mohammadzadeh), Khalilian@kiau.ac.ir (Madjid Khalilian), Nikravan@kiau.ac.ir (Alireza Nikravanshalmani)

Received: June 2021 *Accepted:* September 2021

hashtag in individuals' accounts about a special topic for example in some posts, they constitute an implicit community, promoting a specific topic [49]. The identification and understanding of implicit communities have attracted a great deal of attention in recent years.

With development of social networks in recent decades and their extensive applications across different areas of human life, understanding the personal and collective behavior of individuals in these contexts is crucial. Accordingly, excessive research has dealt with investigating and understanding different aspects of online communities including the formation and detection of communities [17, 47], opinion leadership [38, 27], and influence [39]. Further, different concepts such as popularity, influence, activity, etc. have been defined in different uses and attempts have been made to quantify them. In this research, we have dealt with another very practical concept called charisma in the context of online social media (with a focus on Twitter as the social network of this research). According to the definition [39] has presented for measures and metrics, it is better to say that this research deals with "the charisma measure" in online social networks.

Metric is a simple mathematical statement which aids in providing essential information on the social network as a numerical value. Accordingly, it is possible to merge metrics to define (ranking) a measure. Indeed, a formula or an algorithm can be defined as offering a criterion for ranking every user in a network. Obviously, there are also more complex measures than mere combination of metrics [39].

In this research, charisma measure has been investigated at three levels: individuals, messages, and communities. In other words, this research tries to answer the following question: "how the charisma of messages, individuals, and communities in online social network can be measured and ranked?"

One of the effective methods for answering this question is use of multi-criteria decision-making (MCDM) methods. Various MADM approaches have been developed for ranking of alternatives including weighted sum model (WSM), analytic hierarchy process (AHP), simple additive weighting (SAW), multi-objective optimization on the basis of ratio analysis (MOORA), linear programming technique for multidimensional analysis of preference (LINMAP), Elimination Et Choix Traduisant la Realité (ELECTRE), and technique for order preference by similarity to ideal situation (TOPSIS) [26, 34, 51, 52].

TOPSIS is preferred to other approaches because of (i) its logical and programmable behavior; (ii) requirement of limited subjective inputs; (iii) its suitability for a large number of attributes and alternatives; and (iv) comparative consistency in the alternative ranking [28].

In this research, to measure the charisma, first charisma associated features are extracted based on Twitter communities. Then, the weight of each of these features is calculated, and the level of charisma of each alternative (message and individual) is determined and ranked using TOPSIS method. Via this method, the charismatic messages and charismatic individuals in the community can be easily identified. The issue of "charisma" is important in analyzing and predicting the behavior of users in social networks as well as the areas of marketing, sociology, and political sciences. Identifying charismatic individuals and charismatic messages can be very practical for marketing, generating effective content, and opinion leaders. Also, charismatic communities can be used for brand making as well as predicting and understanding the behavioral patterns of users.

Related work

The concept of charisma has long historical roots which dates back to the Greek word charis meaning charm, beauty, or allurement [33]. Nevertheless, its exact meaning has remained controversial ever since. Charisma is regarded as a personality trait associated with charm, magnetism, or likeability [7]. It has become a key part of the vocabulary through which we describe others; it is perhaps most often applied to the politicians, celebrities, and athletes who act as leaders in our modern society [20].

Charisma, is a concept that has led to wide and varied discussions among scientists. It was first used by Christian churches to qualify outstanding behaviors by people of religious character as a present from God [44, 12]. Later, it gained religious, social and political meanings and was associated with specific power, authority, and legitimacy types [4]. The concept is typically used with regards to the quality or qualities that make a person attractive, and people who are considered charismatic are thought to have superior characteristics [15].

There is still a lack of a clear definition for answering the question "what is charisma?" [20]. Meanwhile, the concept of charisma is used in different areas including management, political sciences, psychology, and other fields, expanding the aspects of this concept. Furthermore, to define the concept of charisma, some other concepts such as influence, inspiration, or environmental factors such as crisis or individual characteristics (followers) such as self-esteem have been propounded. Note that these concepts are intrinsically qualitative and not easily measurable. This research has tried to extract the influential variables affecting charisma and estimate their impact size based on the common points of charisma definitions.

Charisma refers to the ability of a leader to apply diffuse and intense influence over the values, beliefs, behavior, and performance of others through his own beliefs, behavior, and personal example [24]. A study [5] indicated that two Multifactor Leadership Questionnaire (MLQ) dimensions, idealized influence and inspirational motivation, can be merged to form a measure of charisma.

One of the characteristics emphasized in the definitions of charisma is the feature of influencing others. This influence can function as a driver for further activity of individuals. Regarding charismatic leadership, increasing the influence of the leader on others results in greater coordination and activity among the individuals for fulfilling the objectives. Extensive research has dealt with identifying influential nodes in complex networks or influential figures in social networks [39, 2, 25, 16, 50]. Nevertheless, there is no agreement on the exact meaning of an influential user [39]. Therefore, new influence measures are constantly emerging, each of which offers different measurement criteria. Although the concept of charisma is wider than the extent of influence or identification of influential individuals in the network, influence measurement methods in online social networks can inspire charisma measurement.

Among the research conducted regarding charisma in areas other than management and political sciences, one can mention [25] who investigated the features of charismatic trainers along with [36] who extracted the features of charismatic lectures based on Steve Jobs lectures. Generally, the factors that have been presented for measuring charisma in different research can be categorized into two groups:

1. Internal factors: They mean personal and behavioral factors of individuals. Examples include physical features (face, tone of voice) [20, 21], personality (self-confidence [24], self-esteem [30]), social (fame [20]), as well as behavioral skills (verbal and nonverbal skills, management skills such as articulating vision) [30]. The difference between the leader and followers as well as the ability and quality of leadership are associated with these features.

2. External factors: here, structural and environmental features are emphasized. Environmental conditions (e.g. critical conditions, war, peace, change, stability) [20, 30], as well as network features (e.g. network size, communication topology, virtual context or real world) are among the most important features of this group.

Since in this research TOPSIS method has been used for calculating and ranking charisma, we deal with the research in this area. Technique for order preference by similarity to ideal solution (TOPSIS) [26] is a ranking method and is one of the reliable approaches for analyzing multi-criteria decision-making methods [14]. It has been used to identify influential individuals in complex or social networks [25, 16, 50]. Using fuzzy TOPSIS method, [21] investigated charismatic leadership among the presidents of Turkey. Nevertheless, none of them has dealt with the issue of charisma in online social networks. As far as we know, this research has dealt with measuring different aspects of charisma in Twitter for the first time. Furthermore, the use of ranking of messages and individuals in Twitter communities through TOPSIS method along with presenting a practical weighting method is one of the innovations of this research.

In this research, we need datasets in which the number of followers as well as mentions of individuals and number of likes, number of retweets, and the number of replies of each Tweet are all known. One of the most well-known datasets extracted from Twitter context is the research by [32, 13, 46], though they did not cover any of the features required in this research (the number of likes, retweets, replies, mentions, and number of followers).

2. Dataset

The features of the datasets used in this research are summarized in Table 1. In the datasets called DS3 to DS8, the Tweets of six well-known politicians of Asian countries within the first six months of 2019 have been collected. Also, in DS1 and DS2 some tweets with special hashtags have also been collected (they constitute implicit communities). The official website of United Nations has introduced 21 March as the international day for the elimination of racial discrimination) and 25 June as Day of the Seafarer, with the hashtags of FightRacism and IAmOnBosrf. The hashtags in Twitter are used for classifying messages, developing ideas, and promoting special people or topics.

The datasets collected for this research include original tweets, i.e. the tweets whose writers are the user themselves, and do not include retweets or replies. Original tweets the user's creativity and beliefs more than any other of their actions (retweet, reply, etc.). For collection of tweets, REST API of Twitter was used.

3. Proposed approach

As mentioned earlier, in this research first the features associated with charisma were extracted after which these features were weighted. Eventually, using TOPSIS method, the charisma was measured and ranked at message and individual levels. Based on the ranking, the charismatic message, individual, and community (with the maximum charisma) can be identified.

4. Charisma features in Twitter

As mentioned in the section of related work, extensive research has been performed for extracting charisma features in different environments and different uses including [7, 20, 24, 36, 21, 30]. However, as far as we know, no research has dealt with extracting charisma features in social networks. The charisma should be measured based on the actions individuals can do in social networks such as Twitter (e.g. liking, retweeting messages, replies, mentions, follow-up relation, etc.). The important criteria agreed by most researchers for measuring charisma include great influence on beliefs, values, and behavior of others [24, 5, 18], never-ending dissemination of messages [19, 33], creating great

Table 1: The primary features of the datasets collected from Twitter

Name	Dataset	Number of messages	Number of users	Date
DS1	International Day for the Elimination of Racial Discrimination (#FightRacism)	2408	1695	From: 18^{th} March To: 27^{th} March
DS2	Day of the Seafarer (#IAmOnBoard)	1488	1014	From: 24^{th} June To: 4^{th} July
DS3	Tweets of @chedetofficial (Prime minister of Malaysia).	144	1	From: 1^{th} January To: 30^{th} June
DS4	Tweets of @jokowi (President of Indonesia)	463	1	From: 1^{th} January To: 30^{th} June
DS5	Tweets of @JZarif (Foreign minister of Islamic Republic of Iran)	115	1	From: 1^{th} January To: 30^{th} June
DS6	Tweets of @narendramodi (Prime minister of India)	1989	1	From: 1^{th} January To: 30^{th} June
DS7	Tweets of @RTErdogan (President of Turkey)	399	1	From:1^{th} January To: 30^{th} June
DS8	Tweets of @SMQureshiPTI (Foreign minister of Pakistan)	78	1	From:1^{th} January To: 30^{th} June

motivation for more effective and substantial activity [24, 5, 30], fame [20], as well as greater coordination and convergence in individual activities [20, 22]. Accordingly, some features should be chosen for finding charismatic messages, individuals, and communities which can cover these criteria more. Fig. 1 demonstrates these criteria for measuring charisma alongside their associated Twitter metrics. Indeed, to measure a high-level concept called charisma, we focus on more objective four qualitative criteria (influence, popularity, diffusion, and motivation), for whose measurement we work on their associated Twitter quantitative features (metrics). The criteria of charisma and their related Twitter metrics are demonstrated in Fig. 1.

Every person in Twitter can perform some behavior, where the extent of influence of this behavior on others is regarded as the criterion of influence. One of the most important metrics for measuring influence on others in Twitter is the retweets of the messages of the person [39, 22, 41, 9, 42]. When a person retweets a tweet, this tweet is shown for all of the followers of that person. With retweet, more audience can see the tweet. Thus, one of the important metrics for measuring the criterion of influence is the number of times a tweet has been retweeted. The more one's tweets are retweeted, the more likely one is influential [49, 8]. This metric demonstrates a user's ability to generate content that is worth being passed along [9]. Another important metric related to the criterion of influence is the number of times a person is mentioned in the tweets [41, 42, 8, 23, 29]. The number of followers of an individual represents the size of the audience for that individual [9]. Further, [9] has shown that in Twitter those who have millions of followers are not necessarily more influential. Therefore, we did not consider the metric of number of followers for measuring influence.

1] defined information diffusion as the process through which a piece of information (e.g., a tweet) is spread and reaches users through interactions. When a person receives a message, they can decide whether to diffuse it or not (sharing it with friends). The more this interaction grows, the larger the body of information spread throughout the network will be. Retweet fulfills this function in the best way. Everyone who decides to diffuse a tweet retweet it to become available to all of their followers. Further, if any of the followers are also interested, they do the same. Therefore, the number of retweets of a tweet indicates the extent of its diffusion. A charismatic person should be able to encourage motivation and excitement in others through social networks. When the message written by a person is followed by more reactions of the audience (likes, retweets, replies), indeed that person has been able to induce more motivation in others. Therefore, more reactions to a tweet can suggest greater motivating power of that tweet. The number of replies is considered as a special metric for the activity measure [39, 37, 48]. Briefly, a person is said to be popular when they are recognized by many other users of the network [39]. Most researchers have considered the most important metric for measuring popularity as the number of followers [39, 35, 19, 1].

5. Weighting the features

In order to determine the weight of features or the effective metrics of charisma, the combined entropy-Shannon method [40], correlation coefficient methods, and the information obtained from the relationship between metrics and charisma have been used. All of the metrics and measures associated with charisma, similar to Fig. 1, have been tabulated in Table 2. As can be seen, the number of retweets is associated with three measures of "influence", "motivation", and "diffusion". Furthermore, the number of likes and number of followers are related to two measures, while the number of replies and the number of mentions each is associated with one measure. In other words, based on Table 1, the share of retweets, likes, number of followers, replies, and mentions can be obtained. Hence ultimately, for the metrics of retweets, likes, followers, replies, and mentions, we consider normalized weighted coefficients (the importance coefficient of metrics) as 3/9, 2/9, 2/9, 1/9, and 1/9.

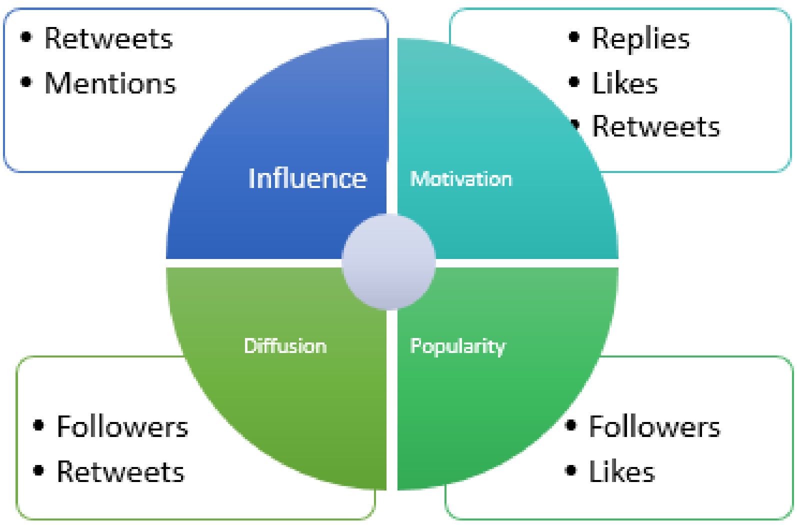

Figure 1: The measures associated with charisma and their pertinent metrics

The weight of metrics through the entropy-Shannon method has been shown in Table 3 for different datasets. Since DS3-DS8 are related to the tweets of one user account, thus the values of the number of followers and mentions of all tweets are the same. Therefore, the weight of these two metrics is considered as zero.

Table 2: The primary features of the datasets collected from Twitter

	Mentions	Replies	Followers	Likes	Retweets
Influence	✓				✓
Motivation		✓		✓	✓
Diffusion			✓		✓
Popularity			✓	✓	

The Spearman correlation coefficient formula for two variables such as x and y is as follows:

$$\rho = \frac{\sum_{i=1}^{n}(R(x_i) - \overline{R(x)}) \times (R(y_i) - \overline{R(y)}))}{\sqrt{\sum_{i=1}^{n}(R(x_i) - \overline{R(x)})^2 \times \sum_{i=1}^{n}(R(y_i) - \overline{R(y)})^2}} \tag{1}$$

$R(x_i)$ and $R(x_j)$ indicate the rank of the i_{th} variable of x and y. $\overline{R(x)}$ and $\overline{R(y)}$ represent the mean rank of x and y variables, while n indicates the total number of observations or samples.

Tau-Kendall b correlation coefficient is obtained through paired comparison of samples with each other. In this method, all paired variables are compared with each other, meaning that for n samples, $n(n-1)/2$ comparisons are required. If for two variables, when the first increases, then the second variable also grows, then we call them as concordant pairs. If with the increase in the first variable, its second variable diminishes, these two samples are called discordant pairs. Finally, if with increase or reduction of the first variable the second variable remains unchanged, they are called tied pairs. Tau-Kendall b formula is as follows:

$$Tau = \frac{N_s - N_d}{\sqrt{(N_s + N_d + T_y) \times (N_s + N_d + T_x)}} \tag{2}$$

Ns represents the number of concordant pairs, N_d indicates the number of discordant pairs, and T_x is the tied pairs based on variable x, while Ty denotes the tied pairs based on variable y.

Since the methods used for weight calculations (entropy-Shannon, Spearman as well as Tau-Kendall correlation methods) have no preference over each other, simple averaging can be used for the combination of the weight of metrics in these two methods. These weights are considered as a W_{1*n} matrix, with n representing the number of metrics. The weight extracted based on charisma (Table 2) is considered as a V_{1*n} matric, where n denotes the number of metrics. Then, through the following formula, the combined weight, i.e. the final weight of the metrics for each dataset is obtained:

$$W_j = \frac{V_j W_j}{\sum_{j=1}^{n} V_j W_j} \tag{3}$$

Indeed, the entropy method performs the weighting based on the validity of data themselves, while correlation coefficient methods those so based on the relationship between the variables. After averaging the weights in these methods, considering the significance factor of each metric regarding charisma, the final weight is calculated. The combined (final) weights for the metrics of number of retweets, likes, followers, replies, and mentions for different datasets are shown in Table.

6. Measuring and ranking charisma

Once the weight of each feature was determined, the charisma can then be measured through TOPSIS, and the obtained value can also be ranked further. TOPSIS involves five steps as follows.

Table 3: The results of weighting the metrics of different datasets based on the entropy, Spearman, Tau-Kendall, and hybrid method (final); L: number of likes, RT: number of retweets, RP: number of replies, F: number of followers, and M: number of mentions

Dataset	Method	L	RT	RP	F	M
DS1	entropy-Shannon	0.169	0.167	0.228	0.230	0.206
	Spearman	0.249	0.245	0.185	0.196	0.126
	Tau Kendall b	0.248	0.249	0.192	0.181	0.130
	Hybrid	0.238	0.354	0.108	0.217	0.083
DS2	entropy-Shannon	0.093	0.107	0.186	0.284	0.329
	Spearman	0.266	0.255	0.157	0.18	0.142
	Tau Kendall b	0.263	0.259	0.165	0.166	0.148
	Hybrid	0.226	0.339	0.092	0.230	0.113
DS3	entropy-Shannon	0.261	0.412	0.327	0	0
	Spearman	0.345	0.340	0.315	0	0
	Tau Kendall b	0.353	0.345	0.302	0	0
	Hybrid	0.312	0.535	0.153	0	0
DS4	entropy-Shannon	0.190	0.370	0.440	0	0
	Spearman	0.340	0.343	0.318	0	0
	Tau Kendall b	0.343	0.348	0.309	0	0
	Hybrid	0.291	0.531	0.178	0	0
DS5	entropy-Shannon	0.234	0.325	0.441	0	0
	Spearman	0.340	0.338	0.322	0	0
	Tau Kendall b	0.346	0.341	0.314	0	0
	Hybrid	0.310	0.508	0.182	0	0
DS6	entropy-Shannon	0.235	0.209	0.556	0	0
	Spearman	0.327	0.344	0.329	0	0
	Tau Kendall b	0.326	0.348	0.327	0	0
	Hybrid	0.312	0.475	0.213	0	0
DS7	entropy-Shannon	0.130	0.131	0.739	0	0
	Spearman	0.327	0.344	0.329	0	0
	Tau Kendall b	0.326	0.348	0.327	0	0
	Hybrid	0.288	0.455	0.257	0	0
DS8	entropy-Shannon	0.221	0.358	0.421	0	0
	Spearman	0.342	0.331	0.327	0	0
	Tau Kendall b	0.347	0.333	0.320	0	0
	Hybrid	0.305	0.515	0.179	0	0

Step one: the decision matrix X_{mn} is created with m observations or alternatives (it can be for example the messages or individuals of a community in Twitter) and n criteria (the features or metrics chosen for charisma) from the datasets of interest.

$$X_{ij} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix} \tag{4}$$

Step two: normalization of the decision matrix with the following formula:

$$r_{ij} = \frac{X_{ij}}{\sqrt{\sum_{j=1}^{m} X_{ij}^2}} \tag{5}$$

Step three: creating the weighted decision matrix through multiplying the values of the normal matrix by the weights of criteria:

$$\nu_{ij} = W_j \times r_{ij} \tag{6}$$

Step four: first, the maximum value of each criterion (column) from the weighted normal matrix is considered as the positive ideal (v_j^+), while the minimum value of each column is regarded as the negative ideal (v_j^+). Next, the sum of the Euclidean distance of all of the elements of each column until the positive ideal or negative ideal of its corresponding column is calculated by the following formula.

$$s_i^+ = \sqrt{\sum_{j=1}^{n} (v_j^+ - v_{ij})^2} \tag{7}$$

$$s_i^- = \sqrt{\sum_{j=1}^{n} (v_j^- - v_{ij})^2} \tag{8}$$

Step five: the relative closeness to the ideal solution for each alternative (column) is calculated by the following formula:

$$C_i = \frac{s_i^-}{s_i^- + s_i^+} \tag{9}$$

Now, based on the value of C_i, the alternatives can be ordered in order to achieve the ranking of interest. Indeed, Ci value functions as the magnitude of the charisma of each alternative. In this way, the charisma is measured and is represented as a numerical value. The alternative with the maximum C_i value can be chosen as the charismatic alternative. Further, for different uses, a threshold limit can be defined for determining the charismatic alternatives. Also, the alternatives can be used based on C_i values for different uses through classifier or clustering methods.

In order to measure the charisma of a community, first we should define a concept called ideal community. The charisma of a community is maximum when all messages are as charismatic as the top message of the community. This means that the extent of charisma of an ideal community is equal to multiplication of the number of alternatives (messages or individuals) by the magnitude of charisma of the ideal alternative (charismatic message or the charismatic individual) in the community. Once the charisma value of the ideal community is known, the rate of charisma of a community can be calculated as follows:

$$C_{community} = \frac{\sum_{i=1}^{m} C_i}{C_{community}^+} = \frac{\sum_{i=1}^{m} C_i}{m \times C^+} \tag{10}$$

Where C_i represents the charisma of the i_{th} alternative in the dataset, $C_{community}+$ shows the charisma of the ideal community, m is the number of community alternatives, and C^+ denotes the charisma magnitude of the ideal alternative. Nevertheless, the value of C^+ can be used as the maximum value of charisma from the average charisma of for example the top 10 alternatives. When $C_{community}$ is known, the charisma rate of different communities can be compared with each other. Larger values indicate that there is less discrepancy between the charisma of a community and its ideal state. This value for any community is between zero and one.

7. Experimental results and discussions

In the real world, it has frequently been observed that a person is chosen as the leader and millions of people followed him or her (e.g. Gandhi, Nelson Mandela, Imam Khomeini in Iranian revolution). Generally, those who have considerable charisma are very few, while those with little charisma are abundant. This is exactly like the concept lying in the power-law distribution for the distribution of degrees of vertices in many real-world networks [10]. The distribution of the degrees of the number of followers in social networks follows power-law distribution. Any network with power-law distribution is called scale-free [49, 6]. The distribution of degrees of features (metrics) in online social media many times follows this distribution [49]. This feature is also observed for the value of charisma. The charisma at the three levels of messages, individuals, and community is inspected further.

8. Charisma of messages

Fig. **??** indicates the values of charisma of messages in DS1 and DS2 alongside their ranks. The trend of both diagrams is the same and is similar to the power-law function ($y = ax^b$). In the top ranks of DS1, the values of charisma are significantly different compared to the respective ranks of DS2. However, from rank 67 onwards, DS2 has larger charisma values at similar ranks, though this difference is very trivial. Indeed, from the ranks around 60 onwards, these two diagrams are close to each other and this trend continues for the rest of the ranks. Indeed, in less than 1% of the top data, different values of charisma are observed. However, in the remaining 99%, the values are similar and close to each other in both diagrams. Nevertheless, the tweets collected for DS1 and DS2 belong to different time periods and based on two different hashtags (two different communities).

Fig. 3 demonstrates charisma values of the top 150 messages for DS3-DS8 datasets (which are related to the tweets of the user account of six politicians (note that if only they had 150 tweets within the six-month period of interest). As can be seen in Fig. 4, the percentage of messages with zero number of likes, retweets, replies, or mentions in both datasets is over 20%, 40%, 80%, and85%, respectively. Also, these two datasets have 23% and 20% cases zero values for all of these four metrics, accounting for a significant portion of all messages of the two communities.

The trend of diagrams indicates that all of them follow the power-law distribution or Pareto distribution [3] in some way. The diagram of these distributions has a long tail, while the early part (head part) is short.

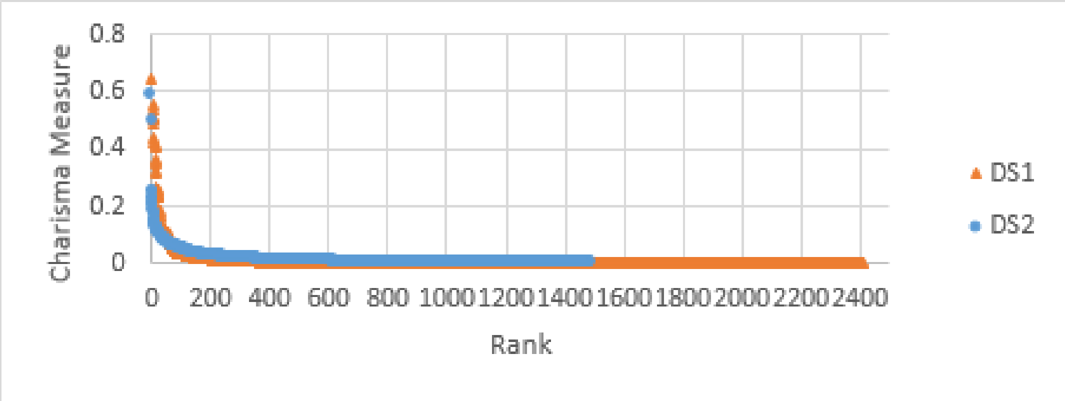

Figure 2: The ranking and values of charisma of messages in DS1 and DS2 datasets

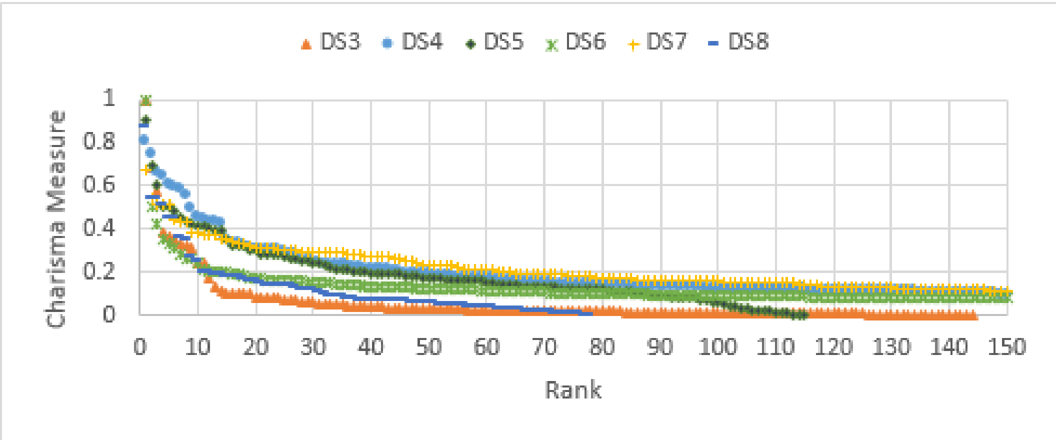

Figure 3: The ranking and values of charisma of messages in DS3-DS8 datasets

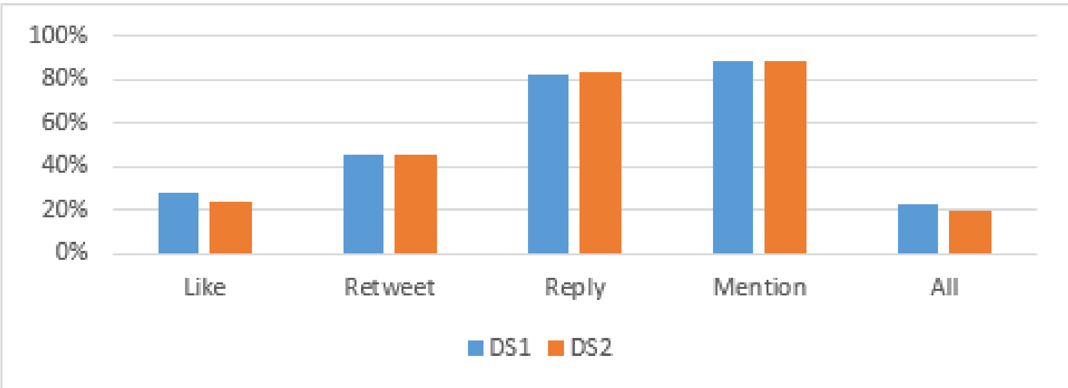

Figure 4: The percentage of messages with zero values in the metrics: number of likes, retweets, mentions, and replies both individually and entirely

9. Charisma of individuals

As mentioned earlier, to measure the charisma of everyone, the metrics related to their messages should be combined with each other. Each of the datasets DS3-DS8 is related to the tweets of a politician. Thus, the charisma which politician is calculated based on the sum of all their tweets.

However, for DS1 and DS2, a person may have several messages, where the charisma of each individual is calculated based on the charisma of their messages as well as the number of followers and mentions as stated earlier. The number of unique users of these two datasets is shown in Table 1.

Fig. 5 displays the rank and values of charisma of individuals in DS1 and DS2. The distribution of the charisma values of individuals, as with the charisma values of messages, has very few individuals with large charisma values, with the majority of the individuals having low charisma. The diagrams of both communities are similar to the charisma of their messages, and the same results can be repeated also here. In some research, an individual (individuals) with a large charisma value is known as a community leader or coordinator. As can be seen for DS1, four individuals (user account) out of 1695 individuals have obtained the maximum level of charisma with a huge difference compared to others. The first and second ranks whose charisma values are greater than 0.6 are related to the user account of @UNHumanRights and @UN. The first is the official account of the United Nations human rights office) with around 2.5 million followers, and the second is the official account of the United Nations with more than 11 million followers. The third and fourth ranks are related to the official Twitter account of secretary-general of the UN, Mr Antonio Guterres with around 600,000 followers and the account of Mr Justin Trudeau, the 23rd Prime Minister of Canada with around 4.5 million followers. As can be observed, these Twitter accounts belong to them as individuals and completely related to the hashtag #FightRacism. For DS2, which is also related to tweets with hashtag #IAmOnBoards, four user account have gained the maximum charisma made a great difference with others. The first rank which is considered the charismatic user account of this community @IMOHQ publishes the latest news of the international Marine organization (IMO as a specialized agency of the United Nations). The second rank is related to the user account of United Nations @UN, and the third rank was obtained by @MarineInsight, the user account of a famous website, which presents maritime news, articles, career guidance, and discussions. It was followed by the user account owned by @AminaJMohammed, the Deputy Secretary-General of the United Nations. As was observed, in this dataset, the charismatic individuals are all well-known figures and related to the hashtag of interest.

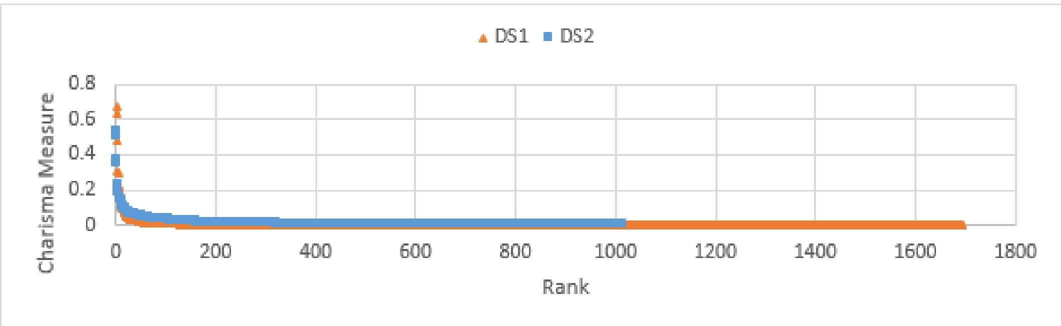

Figure 5: The ranking and values of charisma of individuals in DS1 and DS2

10. Charisma of community

As was also mentioned in calculating the community charisma, the total charisma of the ideal alternative (C^+) can be used instead of the maximum value of the charisma of alternatives from the average charisma of n top alternatives (n can be any number between 1 and the total number of alternatives (m)). Fig. 6 demonstrates the rate of charisma of different datasets based on top 1, 10, and 50 alternatives of each dataset. For C^+ value, as we involve more ranks for the averaging, its value decreases. Therefore, the value of charisma of the ideal community also diminishes, suggesting that the rate of charisma of the community increases. According to the formula $C_{community} = \frac{\sum_{i=1}^{m} C_i}{C^{+}_{community}} = \frac{\sum_{i=1}^{m} C_i}{m \times C^{+}}$, the rate of the committee charisma has an inverse relationship with the number of alternatives (m) and the charisma value of the ideal alternative. One of the most important reasons why DS8, DS5, and DS3 in the average of the top 50 alternatives have the maximum value is that these three datasets have only 78, 115, and 144 messages (alternatives) respectively (compared to other datasets, they have a low m), chosen based on the power-law diagram when the top 50 individuals are chosen. Indeed, a major part of the values with low charisma (the tail of the diagram in Fig. 3) has been used in calculating the ideal alternative, causing reduction of C^{+} value. Furthermore, based on the mentioned formula, one can attribute the low value of charisma in DS1 and DS2 datasets to the large m and having a large number of alternatives with very low charisma, causing reduction of the total charisma of the community. To better compare the rate of charisma of two different communities, it is better to consider the value of m equal in both communities in some way (so that the effect of m compared to the charisma of communities would be neutralized). Then, the effect of other variables including the charisma values of the ideal alternative and the total charisma of community alternatives in the charisma rate of the two communities can be compared in different states.

It is recommended that for choosing the number of alternatives used in C^{+}, in the header and tail part of the diagram, the values of charisma be considered in descending order. A good logic for consideration of the charisma of the ideal community is choosing more alternatives from the header part of these diagrams, so that the ideal state of the community would not have very large values. Generally, selection of the number of alternatives involved in C^{+} can be done based on the use and area of interest, and even it can be combined and used with clustering methods in machine learning (which will be mentioned as the future works of this research). Furthermore, by measuring the level of charisma of an individual, a brand, a news agency, a sports team, or committee over time, one can observe the trend of its charisma over time and compare accordingly.

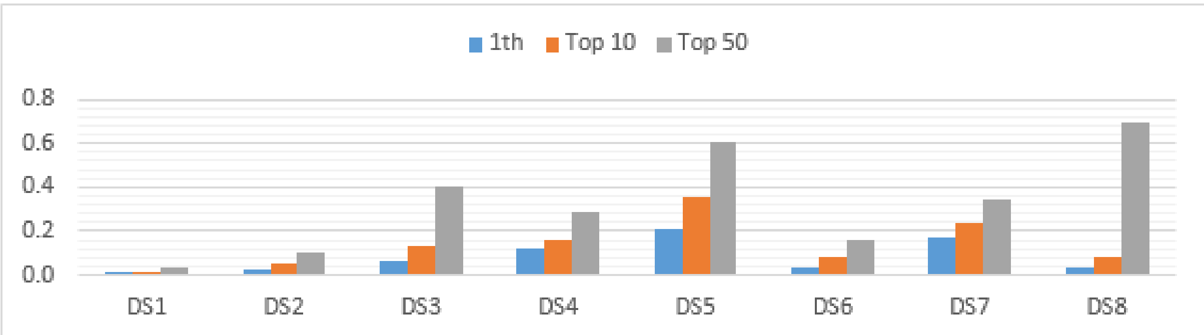

Figure 6: The rate of charisma of communities based on the average value of different charismas for the ideal alternative

11. Conclusion and future works

In this research, first the concept of charisma was examined in different references. Then, the criteria and metrics affecting charisma in Twitter social network were extracted. In order to give the weights to the metrics or their features associated with charisma, entropy-Shannon method along with Spearman correlation coefficient and Tau Kendall b methods were used. Based on the importance of each metric in the charisma (significance factor), the combined or final weight of the features was extracted. In order to measure charisma at the level of messages and individuals and for ranking them in the community, the well-known TOPSIS method was used. Next, we presented a method for measuring the rate of charisma of communities, indicating the discrepancy between every community and the ideal state of the charisma of that community. We measured and ranked charisma at the levels of messages, individuals, and community for all of the eight datasets presented in this research. Ultimately, we thoroughly investigated the behavior of charisma values in these datasets. Although in the real world collected datasets have significant differences with each other, substantial behavioral similarities existed in the distribution of charisma values for messages and individuals as well as some metrics. Further, the way charisma of communities could be compared properly and the main reasons for changes in the charisma values across different communities were also discussed. According to the findings, the results presented for ranking the charisma of user accounts and the charisma of messages in datasets are logical and defensible.

The methods presented in this research can also be generalized to other social networks such as Facebook, Instagram, and others with some minor modification. Since this research is one of the first in inspection of charisma across online social networks, future research can work on examining charisma in other practical areas such as increasing the charisma of brands in online communities, extracting the features of leaders or coordinators of online communities, extracting the features of charismatic messages through text mining methods, exploring the images of multimedia messages, and generally increasing the number of features affecting charisma for predicting the value of charisma across different levels.

References

[1] A. Aleahmad, et al., O*LFinder: Finding opinion leaders in online social networks*, Journal of Information Science, 42(5) (2016) 659-674.

[2] Z.Z. Alp, and Ş.G. Öğüdücü, *Identifying topical influencers on twitter based on user behavior and network topology*, Knowledge-Based Systems, 141 (2018) 211-221.

[3] B.C. Arnold, *Pareto and generalized Pareto distributions, in Modeling income distributions and Lorenz curves*, (2008) 119-145.

[4] Ş. Aslan, *The relation of charismatic leadership and organizational citizenship behavior: The role of'tenure'and'salary'variables*, Journal of Human Sciences, 6(1) (2009) 256-275.

[5] B.J. Avolio, B.M. Bass, and D.I. Jung, R*e-examining the components of transformational and transactional leadership using the Multifactor Leadership*, Journal of occupational and organizational psychology, 72(4) (1999) 441-462.

[6] A.L. Barabási, and E. Bonabeau, *Scale-free networks*, Scientific american, 288(5) (2003) 60-69.

[7] J.M. Beyer, *Taming and promoting charisma to change organizations*, The Leadership Quarterly, 10(2) (1999) 307-330.

[8] M. Bouguessa, and L.B. Romdhane, *Identifying authorities in online communities*, ACM Transactions on Intelligent Systems and Technology (TIST), 6(3) (2015) 30.

[9] M. Cha,, et al. *Measuring user influence in twitter: The million follower fallacy. in fourth international AAAI conference on weblogs and social media*. 2010.

[10] A. Clauset, C.R. Shalizi, and M.E. Newman, P*ower-law distributions in empirical data*, SIAM review, 51(4) (2009) 661-703.

[11] T. Cleff, *Exploratory data analysis in business and economics*, 2014.

[12] J.A. Conger, and R.N. Kanungo, *Charismatic leadership in organizations: Perceived behavioral attributes and their measurement*, Journal of organizational behavior, 15(5) (1994) 439-452.

[13] M. De Domenico, et al., *The anatomy of a scientific rumor*, Scientific reports, 3 (2013) 2980.

[14] N. Delgarm, B. Sajadi, and S. Delgarm, *Multi-objective optimization of building energy performance and indoor thermal comfort: A new method using artificial bee colony (ABC)*, Energy and Buildings, 131 (2016) 42-53.

[15] E.C. Demircioğlu, *Evaluation of Charismatic Leadership from Management Perspective*, Uluslararası Akademik Yönetim Bilimleri Dergisi, 1(1) (2015) 52-69.

[16] Y. Du, , et al., *A new method of identifying influential nodes in complex networks based on TOPSIS*, Physica A: Statistical Mechanics and its Applications, 399 (2014) 57-69.

[17] H. Fani, et al., *User community detection via embedding of social network structure and temporal content*, Information Processing & Management, (2019) p. 102056.

[18] J.B. Fuller, et al., A quantitative review of research on charismatic leadership, Psychological reports, 78(1) (1996) 271-287.

[19] D. Gayo-Avello, *Nepotistic relationships in twitter and their impact on rank prestige algorithms*, Information Processing & Management, 49(6) (2013) 1250-1280.

[20] A. Grabo, B.R. Spisak, and M. van Vugt, *Charisma as signal: An evolutionary perspective on charismatic leadership*, The Leadership Quarterly, 28(4) (2017) 473-485.

[21] D. Gul, and A.S. Uludag, D*etermination of the Most Charismatic Leader Using Analytic Hierarchy Process and Fuzzy TOPSIS: An Application in Turkey*, International Business Research, 9(7) (2016) 80-97.

[22] J. Henrich, M. Chudek, and R. Boyd, *The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders*, Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1683) (2015) 20150013.

[23] J. Herzig, Y. Mass, and H. Roitman. *An author-reader influence model for detecting topic-based influencers in social media*, in Proceedings of the 25th ACM conference on Hypertext and social media. 2014.

[24] R.J. House, W.D. Spangler, and J. Woycke, *Personality and charisma in the US presidency: A psychological theory of leadership effectiveness*, in Academy of Management Proceedings, Academy of Management Briarcliff Manor, NY 10510, 1990.

[25] J. Hu, , et al., *A modified weighted TOPSIS to identify influential nodes in complex networks*, Physica A: Statistical Mechanics and its Applications, 444 (2016) 73-85.

[26] C.L. Hwang, and K. Yoon, *Methods for multiple attribute decision making, in Multiple attribute decision making*, (1981) 58-191.

[27] L. Jain,and R. Katarya, *Discover opinion leader in online social network using firefly algorithm*, Expert Systems with Applications, 122 (2019) 1-15.

[28] P.P. Kalbar, S. Karmakar, and S.R. Asolekar, *Selection of an appropriate wastewater treatment technology: A scenario-based multiple-attribute decision-making approach*, Journal of environmental management, 113 (2012) 158-169.

[29] G. Katsimpras, D. Vogiatzis, and G. Paliouras, D*etermining influential users with supervised random walks*, in Proceedings of the 24th International Conference on World Wide Web. 2015.

[30] K. Kim, F. Dansereau, and I.S. Kim, *Extending the concept of charismatic leadership: An illustration using Bass's (1990) categories, in Transformational and Charismatic Leadership: The Road Ahead 10th Anniversary Edition*, Emerald Group Publishing Limited, (2013) 165-194.

[31] W.R. Knight, *A computer method for calculating Kendall's tau with ungrouped data*, Journal of the American Statistical Association, 61(314) (1966) 436-439.

[32] J. Leskovec, and J.J. Mcauley, *Learning to discover social circles in ego networks*, in Advances in neural information processing systems. 2012.

[33] B. MacLachlan, *The age of grace: Charis in early Greek poetry*, 2014.

[34] A. Mardani, et al., *Multiple criteria decision-making techniques and their applications–a review of the literature from 2000 to 2014*, Economic Research-Ekonomska Istraživanja, 28(1) (2015) 516-571.

[35] R. Nagmoti, A. Teredesai, and M. De Cock, *Ranking approaches for microblog search*, in Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01. 2010.

[36] O. Niebuhr, J. Voße, and A. Brem, *What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice*, Computers in Human Behavior, 64 (2016) 366-382.

[37] A. Pal, and S. Counts, *Identifying topical authorities in microblogs*, in Proceedings of the fourth ACM international conference on Web search and data mining. 2011.

[38] C.S. Park, and B.K. Kaye, T*he tweet goes on: Interconnection of Twitter opinion leadership, network size, and civic engagement*, Computers in Human Behavior, 69 (2017) 174-180.

[39] F. Riquelme, and P. González-Cantergiani, *Measuring user influence on Twitter: A survey*, Information Processing & Management, 52(5)(2016) 949-975.

[40] C.E. Shannon, *A mathematical theory of communication*, Bell system technical journal, 27(3) (1948) 379-423.

[41] A. Silva,, et al. *ProfileRank: finding relevant content and influential users based on information diffusion*, in Proceedings of the 7th Workshop on Social Network Mining and Analysis. 2013.

[42] D. Simmie, M.G. Vigliotti, and C. Hankin, *Ranking twitter influence by combining network centrality and influence*

observables in an evolutionary model, Journal of Complex Networks, 2(4) (2014) 495-517.
[43] A. Towler, , et al., *How charismatic trainers inspire others to learn through positive affectivity*, Computers in Human Behavior, 32 (2014) 221-228.
[44] M. Weber, *The theory of economic and social organization*, Trans. AM Henderson and Talcott Parsons. New York: Oxford University Press, 1947.
[45] A.D. Well, and J.L. Myers, *Research design & statistical analysis*. 2003.
[46] J. Yang, and J. Leskovec, *Patterns of temporal variation in online media*, in Proceedings of the fourth ACM international conference on Web search and data mining. 2011.
[47] X. You, Y. Ma, and Z. Liu, *A three-stage algorithm on community detection in social networks*, Knowledge-Based Systems, 2019.
[48] J. Yuan, et al., *Topology-based algorithm for users' influence on specific topics in micro-blog*, JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE, 10(8) (2013) 2247-2259.
[49] R. Zafarani, M.A. Abbasi, and H. Liu, *Social media mining: an introduction*, 2014.
[50] A. Zareie, A. Sheikhahmadi, and K. Khamforoosh, *Influence maximization in social networks based on TOPSIS*, Expert Systems with Applications, 108 (2018) 96-107.
[51] E.K. Zavadskas, et al., *Hybrid multiple criteria decision-making methods: A review of applications for sustainability issues*, Economic research-Ekonomska istraživanja, 29(1) (2016) 857-887.
[52] E.K. Zavadskas, Z. Turskis, and S. Kildienė, S*tate of art surveys of overviews on MCDM/MADM methods*, Technological and economic development of economy, 20(1) (2014) 165-179.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.