A MODEL TO MEASURE THE SPREAD POWER OF RUMORS

A PREPRINT

Zoleikha Jahanbakhsh-Nagadeh
Department of Computer Enginering
Naghadeh Branch, Islamic Azad University
Naghadeh, Iran
zoleikha.jahanbakhsh@srbiau.ac.ir

M.Reza Feizi-Derakhshi
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
mfeizi@tabrizu.ac.ir

Majid Ramezani
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
m_ramezani@tabrizu.ac.ir

Taymaz Rahkar-Farshi
Department of Software Engineering
AltÄśnbaÅ§ University
Istanbul, Turkey
taymaz.farshi@altinbas.edu.tr

Meysam Asgari-Chenaghlu
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
m.asgari@tabrizu.ac.ir

N. Nikzad-Khasmakhi
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
n.nikzad@tabrizu.ac.ir

Ali-Reza Feizi-Derakhshi
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
derakhshi96@ms.tabrizu.ac.ir

Mehrdad Ranjbar-Khadivi
Department of Computer Engineering
Shabestar Branch, Islamic Azad University
Shabestar, Iran
mehrdad.khadivi@iaushab.ac.ir

Elnaz Zafarani-Moattar
Department of Computer Engineering
Tabriz Branch, Islamic Azad University
Tabriz, Iran
e.zafarani@iaut.ac.ir

M. A. Balafar
Department of Computer Engineering
University of Tabriz
Tabriz, Iran
balafarila@tabrizu.ac.ir

February 28, 2020

arXiv:2002.07563v2 [cs.SI] 27 Feb 2020

Abstract

Nowadays, a significant portion of daily interacted posts in social media are infected by rumors. This study investigates the problem of rumor analysis in different areas from other researches. It tackles the unaddressed problem related to calculating the Spread Power of Rumor (SPR) for the first time and seeks to examine the spread power as the function of multi-contextual features. For this purpose, the theory of Allport and Postman will be adopted. In which it claims that there are two key factors determinant to the spread power of rumors, namely importance and ambiguity. The proposed Rumor Spread Power Measurement Model (RSPMM) computes SPR by utilizing a textual-based approach, which entails contextual features to compute the spread power of the rumors in two categories: False Rumor (FR) and True Rumor (TR). Totally 51 contextual features are introduced to measure SPR and their impact on classification are investigated, then 42 features in two categories "importance" (28 features) and "ambiguity" (14 features) are selected to compute SPR.

The proposed RSPMM is verified on two labelled datasets, which are collected from Twitter and Telegram. The results show that (i) the proposed new features are effective and efficient to discriminate between FRs and TRs. (ii) the proposed RSPMM approach focused only on contextual features while existing techniques are based on Structure and Content features, but RSPMM achieves considerably outstanding results (F-measure=83%). (iii) The result of T-Test shows that SPR criteria can significantly distinguish between FR and TR, besides it can be useful as a new method to verify the trueness of rumors.

Keywords Spread power of rumor · Ambiguity of rumor · Importance of rumor · Automatic rumor verification · Social networks

1 INTRODUCTION

With the advent of messenger and social networks, every user spreads posts without any editing, checking for factual accuracy, and clearly being accountable. These unverified posts are named rumor. Rumors are an inevitable social phenomenon and have had great importance throughout the history of human life because this social phenomenon transforms the political, social, and cultural life of humans. DiFonzo and Bordia [1] define rumor as âĂIJunverified and instrumentally relevant information statements in circulation that arise in contexts of ambiguity, danger or potential threat, and that functions to help people make sense and manage riskâĂİ. This unverified information may turn out to be true, or partly or entirely false; alternatively, it may also remain unresolved. Hence, throughout this study is adhered to DiFonzo and Bordia's [1] definition of rumors and classified rumors into two categories: True Rumor (TR) and False Rumor (FR). In this study, an FR is defined as "a piece of false and circulating information that based on special contextual features is designed for certain goals". TR is defined as " information that has been verified by trustworthy sources". Also, the person who creates and spreads rumors is named as a rumormonger.

The research described in this paper is considered the problem of rumor analysis in a different area from other researches and proposed a model to compute the spread power named it the Spread Power of Rumors (SPR). Despite notable works that explore information diffusion, rumor propagation in particular, on social media during crisis events, very few studies have looked at the speed of spreading rumors. To the best of the researchers' knowledge, there is no work in the area of SPRs, except Allport and Postman's[2] theory about the power of rumors.

The rumor power rule was provided by Allport and Postman [2] in 1947. They studied the psychology of rumors and introduced two factors of importance and ambiguity as effective factors in spreading rumors. This theory is based on two assumptions [2]: (i) people exert effort to find meaning in things and events; (ii) when people faced with ambiguity in an important matter, they try to find some meaning by retelling related rumors.

This means that the importance and ambiguity of rumors are vital variables that predict whether a rumor would be transmitted or not [2]. Allport and Postman defined the following formula to compute seriousness and propagation measure of rumor based on these two factors.

$$Power of Rumor \approx Importance \times Ambiguity \tag{1}$$

In formula 1, the relation between ambiguity and importance is not sum, but is multiplication; because if ambiguity or importance is zero then there will be no rumor. In other words, the power of a rumor is roughly equivalent to the multiplication of the importance and ambiguity surrounding the rumor. According to this theory, whenever the importance of a rumor is high, its influence rate goes up equally. Also, with increasing ambiguity in the case, the penetration rate of rumors rises. If one of these volatility or importance factors be zero, the influence rate of the rumor will be zero [2].

In other words, it is unlikely that an individual will attempt to spread a rumor that does not matter to him, although it is ambiguous. Also, the importance of the subject alone is not enough to spread the rumor, because the importance should be with the ambiguity that rumor reveals. For example, a rumor about choosing a presidential candidate after the announcement of the election results, though it is an important issue, due to the lack of ambiguity, it will not spread.

As another inference, a rumor about raising or lowering the percentage of banks' profits does not matter to anyone who does not have the money in the banks, and he will not pursue it. On the other hand, such rumors are less effective for bank employees and officials as they are aware of the exact news of the profit levels in the banks and there is no ambiguity for them, but it is different for ordinary people.

We used formula 1 as a rule to compute the power of rumors spread. This rule creates a motive for rumor propaganda, which means that rumor does not spread unless it is important for someone who listens to the rumor and spread it. Also, the rumor's ambiguity makes people interested in spreading it. Anxiety and worry are also a motivation for spreading stories containing fear and threat, which we hear in many cases. Therefore, the present research is focused on a set of features that determine the importance and ambiguity of the rumor and analyzed these features to calculate the SPR. Finally, the SPR criterion has been used as a characteristic in verifying rumors.

The rest of the paper is organized as follows: In Section 2, the problem definition and objectives are presented. Section 3 reviews a summary of related works. Section 4 describes the proposed Rumor Spread Power Measurement Model (RSPMM) and investigates a set of effective features to compute SPR. Section 5 describes the experiments and evaluations and in Section 6 conclusions of the paper is shown.

2 Problem definition and objectives

Each social phenomenon requires a certain category of conditions for its emergence and sustainability. Rumor is also a social phenomenon that requires the realization of conditions for publication and acceptance. In 1974, Allport and Postman [2] provided a question: what are the conditions and factors needed to increase the intensity of rumor spread? They have outlined two fundamental conditions for spreading the rumor: first, the issue of rumor should be important to the audience. If the subject is interesting to the audience, rumors about that subject may be interesting to them, but this condition alone is not enough. The second condition for propagating the rumor is the existence of ambiguity in expressing the issue. Of course, rumors are more infectious; when little information is released through authoritative channels and uncertainties occur in society. Allport and Postman [2] defined the law of the power of the rumor based on the multiplication of importance and ambiguity. Allport and Postman's "basic law of rumor" is that the condition of the rumor's release depends on the existence of both factors of importance and ambiguity in the rumor. So far, this law has been presented as a theory and has not been practically studied on rumors.

Since the rumor text may contain effective text variables that increase the speed of its publication, so this article focuses on rumor content features to propose SPR as a new feature that may be effective in determining rumor validity. Thereby, in this study, the problem of computing SPRs as an unaddressed problem is formulated based on Allport and Postman's [2] theory about the power of rumors and the analysis of the textual content. The textual content has features that can be effective to compute two factors of importance and ambiguity. Accordingly, we seek to examine SPR as a function of several contextual features to investigate the role of words in computing the spread power of FRs. The main question of the research is:

Question: Is there a difference between the spread power of False Rumors and True Rumors?

To answer this question, it is necessary to calculate the propagation power for TRs and FRs to determine whether the criterion of power is a criterion in distinguishing TRs from FRs. According to Allport's law, the power of publication depends on two factors of importance and ambiguity, so it is necessary to calculate these two factors. In this paper, these two factors are calculated based on contextual features. Thereby, two sub-questions arise:

Sub-question A: What features show importance of a story to its reader? **Sub-question B:** What textual factors show ambiguity in the rumor?

The contextual characteristics of the FRs are analyzed and extracted to answer research questions. Thereby, a set of features is proposed to determine the importance and ambiguity score of the FRs. Our main objectives in this study are as follows:

- To provide a computational model for computing the SPR.
- To utilize SPR as a factor along with other extracted features to verify rumors.

So far, no research has been conducted on the problem of calculating SPR, and this research is the first work to solve this problem. Therefore, first, the content of 2431 FRs is deeply reviewed to identify and extract an extensive list of contextual features from FRs in the Persian language. Then, formulas are presented based on the introduced features to calculate each of the two "importance" and "ambiguity" factors. Moreover, weights are allocated to each of the features involved in the calculation of these two factors using a weighting technique. Finally, the spread power of the rumors is calculated by the multiplication of two factors of importance and ambiguity. In order to, Allport and Postman's theory is practically implemented to compute SPR.

3 Related works

Rumor phenomenon is studied by researchers in various research areas in the English language such as rumor detection, verification, and propagation modeling. However, in the Persian language is done very little work on analyzing the

Persian rumors. In the area of psychology, Allport and Postman [2] stated that the rumors are the product of two factors, "importance" and "ambiguity", which are two determining factors in the speed of rumor propagation. Harsin [3] presented the idea of the "Rumor Bomb", it means that a "Rumor Bomb" spreads the notion of the rumor into a political communication concept. In another research, Kumar and Geethakumari [4] explored the use of theories in cognitive psychology to estimate the spread of disinformation, misinformation, and propaganda across online social networks. They proposed an algorithm that would use social media as a filter to separate misinformation from accurate information.

In detection and verification areas, previous researchers analyzed various features at three levels: user information, content, and distribution network structure [5]. For example, Castillo et al. [6] and Kwon et al. [7] proposed a combination of linguistics and propagation features that can be used to approximate the credibility of information on Twitter. They focused on three aspects of features: temporal, structural and linguistic. Also, they identified critical differences in the spread of FRs and TRs. Castillo et al. [6] studied the propagation of rumors during real-world emergencies while Kwon et al. studied the propagation of urban legends (such as Bigfoot) on Twitter. Qazvinian et al. [8] addressed the problem of rumor detection in microblogs and explored the effectiveness of three categories of features: content-based, network-based features, and microblog-specific memes for identifying rumors. They also proposed four content-based features and presented the tweet with two different patterns: (1) Lexical patterns (all the words and segments in the tweet), (2) Part-of-speech patterns. Therefore, from each tweet, a set of features is extracted by corresponding to unigrams and bigrams of each representation.

Yang et al. [9] have done work similar to Castillo's work on Sina Weibo, a Chinese leading micro-blogging service provider that functions like a Facebook-Twitter hybrid. Yang et al. [9] grouped lists of features into five broad types (content-based, client-based, account-based, propagation-based, and location-based), and extracted the rumor detection features on Sina Weibo, for the binary classification purposes (rumor or not). Kwon et al. [7] and Yang et al. [9] show that the most significant features for rumor detection are emoticons, opinion words and sentiment scores (positive or negative). In 2015, Wu et al. [10] used all these effective features, plus two new semantic features: a topic model feature and a search engine feature. In 2015, Vosoughi [11] presented relatively new work to predict the veracity of rumors. He identifies salient features of rumors by examining three aspects of information spread: linguistic style used to express rumors, characteristics of people involved in propagating information, and network propagation dynamics. In 2017, Zamani et al. [12] addressed the problem of rumor detection on the Persian Twitter community for the first time. They explored and analyzed two categories of rumor features including structural and content-based features. They studied the effects of two sets of features to detect Persian rumors in two experiments. The first yields around 70% precision just based on structural features and the second more than 80% based on both categories of features. Then, they show how the features of users tending to produce and spread rumors, are effective in the rumor detection process.

In later years, researchers utilized semi-supervised and deep learning methods to solve the problem of rumor detection and verification. In 2019, Alkhodair et al. [13] proposed a semi-supervised learning solution for breaking news rumor detection that jointly learns word embeddings and trains a recurrent neural network. They also mitigated the cross-topic issue in breaking news rumor detection. Also, Liu et al. [14] proposed Long Short-Term Memory (LSTM) network based models for identifying rumors by capturing the dynamic differences between the spreaders and diffusion structures of the whole or only the beginning part (i.e., early rumor identification) of the spreading process. Xing et al. [15] proposed a novel attention learning framework via deep visual perception based recurrent neural network (ViP-RNN), considering both high-level feature interactions and contextual information. In particular, the proposed model is based on RNN for capturing the long-distance temporal dependencies of contextual information of relevant posts and composing low-level lexical features into high-level semantic interactions hierarchically by visual perception of convolutional neural network (CNN).

Zhao et al. [16] tackled early detection of rumors by determining clusters of potential rumors. For each cluster, a series of features describing it, are extracted. These features were two types of language patterns in rumors: the correction type and the enquiry type. The extracted features are fed to a J48 decision tree. The authors report an accuracy of 0.52 for their best run. In contrast, Zubiaga et al. [17] leveraged the context preceding a tweet with a sequential classifier. Their proposed method was based on the hypothesis that a tweet alone may not suffice to know if its underlying story is a rumor, due to the lack of context. They utilized a sequential classifier that learns the reporting dynamics during an event so that the classifier can determine whether the new tweet is a rumor or not.

Wang and Terano [18] proposed a study on the analysis of patterns of diffusion to detect rumors. These patterns combine structural and behavioral properties of rumor. They identified a series of short diffusion patterns, based on stance, that appear to be strongly related to rumors.

A recent study of rumors detection was presented by Kwon et al. [19] by examining rumor characteristics over different observation time windows. User, linguistic, network, and temporal features were adopted in this study in order to identify the significant differences between rumors and non-rumors for the first 3, 7, 14, 28 and 56 days from the initiation.

Arabic rumors identification was presented by Floos [20], who focus on detecting rumors in Arabic tweets. The authors identified two types of tweets Rumors and News, then TF-IDF was computed for each term in the tweets. To classify the new tweet. The dot product was applied to the tweet vector and the news vector. The same steps are repeated for the rumors vector; the largest dot product result is used as a pointer to classify the new tweet. In another study, an information propagation model of a given microblog over a social network is proposed by Liu and Xu [21] based on that propagation patterns of rumors and credible messages are distinct from each other in social media. They analyzed and carried this model on Chinese social media site Sina Weibo.

Some studies used unsupervised approaches to detect rumors. One of these studies has been proposed and carried by Takahashi and Igata [22]. They investigated several characteristics of rumors including the retweet ratio, word distribution. The results of author's experiments concluded that retweet ratio tells no difference between non-rumors and rumors, but for a large sample size, it can get useful results. Also, the difference in word distribution characteristic appears to have an essential rule for rumors detection. Chang et al. [23] proposed a clustering-based approach for political rumors detection on Twitter. Their unsupervised classification method employs five structural and timeline features for the detection of extreme users. The extreme users are a particular kind of malicious user that tends to post false news on Twitter. These users have a large number of following, a high tweeting frequency, they show enthusiasm about the target topic and use extreme keywords in the description or tweets. The trustiness of news tweets is decided based on the number of extreme users. Chen et al. [24] addressed false rumor detection as an anomaly detection problem in Sina Weibo, not as a classification task, and viewed false rumors as anomalies. The authors proposed three categories of features including crowd wisdom, content features, and posting behavior. They first extracted the set of recent k Weibos of the user who propagated a post in addition to a set of needed features to these Weibos to detect a rumor; then performed a Factor analysis of mixed data (FAMD) on their proposed features to detect anomalies. Jain et al. [25] proposed an approach based on the premise that verified News Channel accounts on Twitter would provide more credible information than the unverified account of user (public at large). They defined four main steps for detecting rumors on Twitter. First, extracting tweets identifying topics based on clustering using hashtags and then collect tweets for each topic. Second, isolate the tweets for each topic in two categories of news and public tweets based on whether its tweeter is a verified news channel or a general user. Third, analyzing tweets of the same topic from verified Twitter accounts of News Channels and other unverified (general) users based on semantic and sentiment features to calculate mismatch ratio. And finally, if the mismatch ratio is greater than a threshold value, the topic is labeled as a rumor; otherwise, it is labeled as not rumor.

Some research has also focused on modeling the distribution of rumors in the social space [5]. For example, Zeng et al. [26] modeled the speed of information transmission to compare retransmission times across content and context features. They specifically contrasted rumor-affirming messages with rumor-correcting messages on Twitter during a notable hostage crisis to reveal differences in transmission speed. Doerr et al. [27] simulated a natural rumor spreading process on several classical network topologies. They also performed a mathematical analysis of this process in preferential attachment graphs and proved that the rumor spreading process disseminates a piece of news in sub-logarithmic time. That is, the spread speed of rumors is extremely fast on social networks. Based on these results, it can be claimed that the spread power of FRs is more than TRs.

The researches in these three areas try to judge the credibility of information based on a wide range of features related to the message, users, topics, and propagation pattern within Twitter, especially in the English language. In contrast, our research study rumors in a different area from others. We intend to formulate the theory of Allport and Postman [2] about the power of rumor. For this purpose, we will focus on analyzing contextual factors that increase the importance and ambiguity of the rumor, and we will compute SPR based on these two factors, also. The proposed features are drawn from extensive theories in social psychology and by analyzing the context of rumors in the Persian language.

4 Procedure for Computing the Spread Power of Rumors

Rumors about an issue are published when it is important in people's lives and the relevant information is incomplete or ambiguous. The existence of ambiguity may be due to the lack of transparency of the news or the receipt of contradictory types of news. According to this description, the chance of life and the prevalence of rumors increase based on two factors of "importance" and "ambiguity". Thereby, in the proposed model, a set of features is introduced to compute the score of these two factors and finally calculate the SPR score. The general structure of the proposed Rumor Spread Power Measurement Model (RSPMM) for computing the spread power of rumor is shown in Figure 1. As described in Figure 1, our model consists of the following steps:

1. Text pre-processing: Converting the text into a form that is predictable and analyzable for our task.

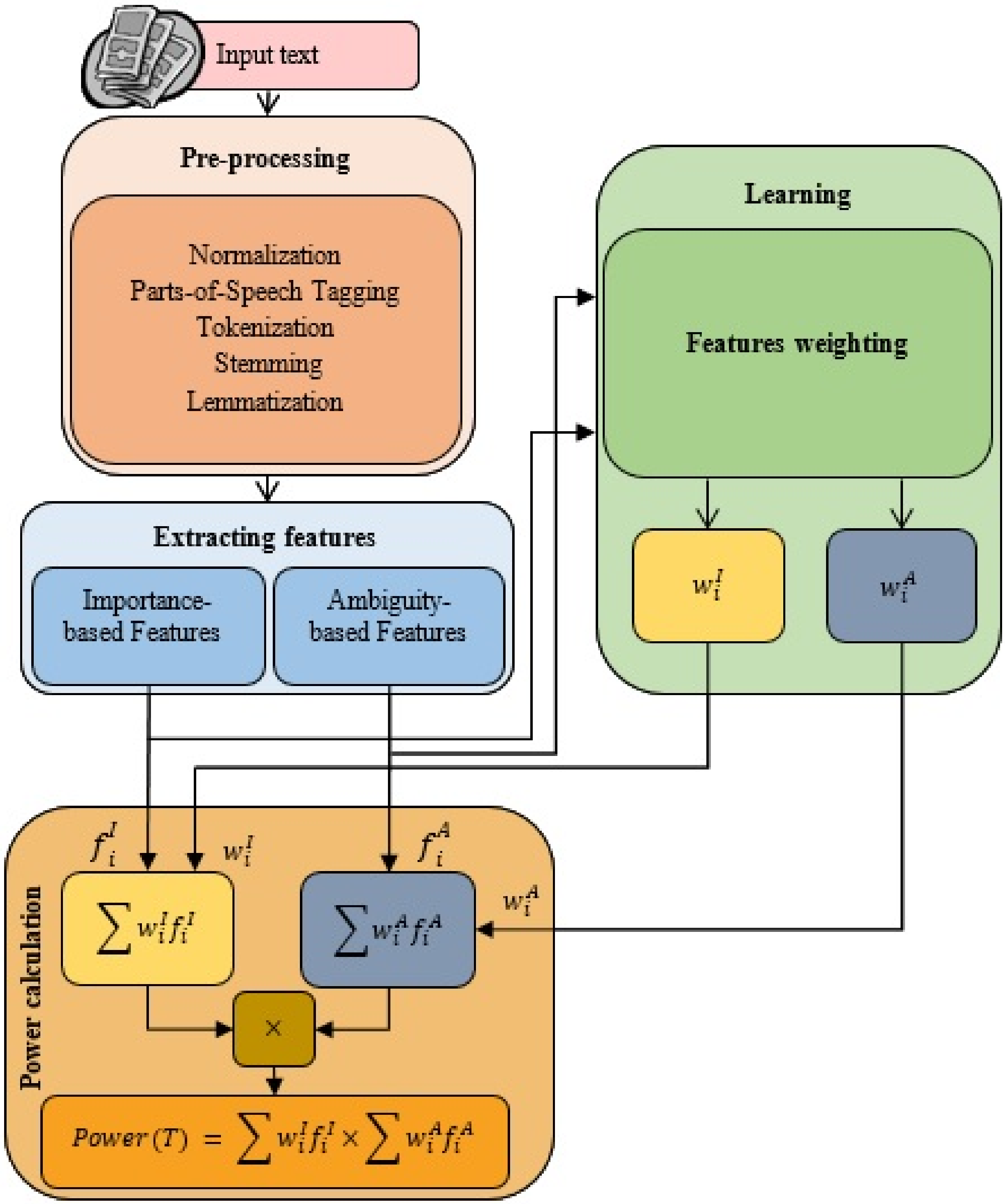

Figure 1: Proposed structure to compute spread power of rumor.

2. Feature extraction: Analyzing and extracting the characteristics of a fast-spreading rumor to compute the score of these two factors.
3. Learning: Weighting content features and determining the degree of importance of each feature in predicting classes.
4. Power calculation: Computing the power of rumor based on two score importance and ambiguity.

The steps of the proposed model are described in more detail below.

4.1 Text Pre-processing

Online texts such as posts propagated on social networks and messenger usually contain lots of noise and uninformative parts (such as symbols, special characters). Therefore, five steps of pre-processing include tokenization, normalization, part of speech tagging, stemming, and lemmatization are performed on the rumor texts.

4.2 Feature extraction

As mentioned in the previous section, two influential factors in the propagation of rumors are the importance and ambiguity. Therefore, the context of rumors is analyzed using text mining and NLP techniques to extract useful contextual features into two categories of importance and ambiguity that can distinguish FRs from TRs.

The content of FRs is more captivating and attractive than TRs; therefore FRs are quickly accepted and propagated by audiences without any review of its accuracy. Also, the plot of an FR is not out of two modes; either it is expressed based on imagination, lies, and slander, or it is published an event that its acceptance depends on the state of the audience's public opinions and its publication time. Accordingly, rumormongers use the power of words to express FRs to captivate the audience and gain their trust. Thereby, FRs make a sense similar to the truth for the audience so that the audience accepts it and propagates it, even if its validity is doubtful.

The main purpose of this research is to implement the theory of Allport and Postman on rumors. Therefore, it is necessary to first study the contextual features of a lot of Persian FRs and identify the textual features that increase the importance and ambiguity of the rumor. Then, SPR is computed based on these features. First, 42 features have been introduced to compute SPR; 28 features to determine the importance of the text and 14 features to compute the ambiguity of the text. Table 1 shows the features introduced to calculate SPR. In Table 2, a set of features is illustrated that these features are first used in RSPMM to calculate SPR. However, after evaluating these features on the dataset, it was found that these features have little ability to distinguish FR and TR, so these features are removed from the list of features.

Table 1: Summaries of contextual features and their component dependent variables and measures to compute two factors of importance and ambiguity. The proposed new features are marked with the "*" mark.

#	Abbreviation	Feature	Description
		Emotional	
1	E_Tag	Emotiveness [28]	The ratio of adjectives plus adverbs to nouns plus verbs.
2	F	Fear*	The ratio of the number of sentences containing fear-based words to the total number of sentences in the text.
3	SU	Surprise*	The ratio of the number of sentences containing surprise-based words to the total number of sentences in the text.
4	D	Disgust*	The ratio of the number of sentences containing disgust-based words to the total number of sentences in the text.
5	Sad	Sadness*	The ratio of the number of sentences containing sadness-based words to the total number of sentences in the text.

6	AN	Anger*	The ratio of the number of sentences containing anger-based words to the total number of sentences in the text.
7	MV	Motion Verbs*	Verbs of motion in two categories: (1)transitional (i.e., top to down, down to top, left to right, right to left, and multi-directional), (2)Self-contained motion (such as, oscillation, dilatory, rotation, wiggle, wander, rest).
8	CW	Consecutive Words*	consecutive repeated words, for example, «توجه توجه» / "tevajjoh tavajjoh"/"Attention Attention" and so on.
9	CC	Consecutive Chars*	the existence of consecutive repeated characters in a word. For example, آسلاااااامممم / "salÃćÃćÃćÃćÃćÃćÃćm"/"hellllllloooo".
10	PS_Sent	Positive Sentiment [6]	The ratio the number of positive words in the text to the sum of positive and negative words.
11	NG_Sent	Negative Sentiment [6]	The ratio of the number of negative words in the text to the sum of positive and negative words.
12	SA_Thrt	Speech Act_Threat*	With these speech acts we can promise for hurting somebody or doing something if hearer does not do what we want.
13	SA_Req	Speech Ac_Request*	Politely asks from somebody to do or stop doing something.
14	Adj_SUP	Superlative Adjective*	Simple Adjective + Suffixes «ترین» /tarin/ Number + («اُی» + «مین» /[o]min/
15	Adj_CMPR	Comparative Adjective*	Simple Adjective + Suffixes + «تر» پسوند /tar/
16	AFF	Affective*	Words that cause emotion or feeling, such as, «اخطار» "/"ekhtar"/" Warning", «دلخراش» delkharash"/"irritant".
17	Start	Start phrase*	The beginning sentence of the text.
18	End	End phrase*	The ending sentence of the text.
		Newsworthy	
19	RT	Relative Time*	Words like «امشب» / "emÅąab" / "tonight", «امروز» / "emroz" / "today", «اخیراً» / "akhiran" / "recently" and so on.
20	ADJ_QTY	Quantity Adjective [28]	Words such as «برخی» / "barkhi" / "some, «همه» / "hame" / "all" and so on.
21	ND	Numerical Digits [6]	The number text characters used to show numerals. For example, the numeral "56" has two digits: 5 and 6.
22	NE	Named Entity [29]	In three classes, the person's name, the organization, the location.
23	LD	Lexical Diversity [28]	Which is the percentage of unique words or terms in all words or terms [28].
24	CER	Certainity [28]	The ratio of certainty-based words to the sum of certainty and uncertainty-based words.

Imp	25	SA_Dec	SA_Declerative*	Transfer information to hearer, this type of SAs commits the speaker to something being the case.
	26	SA_Quot	SA_Quotations*	These are speech acts that another person said or wrote before.
	27	ADJ_ORD	Ordinal Adjective*	Number + («ام», «م») /om/
	28	SM	Spelling Mistake [28]	The ratio of misspelled words based on typographical errors to total number of words.
Ambiguity	29	UCER	Uncertainity [28]	Word that indicates lack of sureness about someone or something, where the ratio of uncertainty-based words to the sum of certainty and uncertainty-based words is calculated.
	30	SV	Sensory Verb	A verb that describes one of the five senses: sight, hearing, smell, touch, and taste. For example, «شنیدن» /"Åąenidan"/hear, «حس کردن» /"hes kardan"/feeling, «دیدن» /"didan"/see, and so on.
	31	QW	Question Word [6]	A function word used to ask a question, such as what, when, where, who, âĂę
	32	QM	Question Mark [6]	The ratio of the number of sentences containing the question mark '?' or multiple question marks "?????" to the total number of sentences of the text.
	33	EM	Exclamation Mark [6]	The ratio of the number of sentences containing the exclamation mark to the total number of sentences of the text.
	34	SA_Ques	Speech Act_Question*	These are usual questions for information or confirmation.
	35	PRO	Pronoun [6]	A personal pronoun in 1st, 2nd, or 3rd person.
	36	Tntv	Tentative	The adjective tentative describes something that is uncertain and unsure.
	37	NEG	Negation [19]	Units of language, e.g., words (e.g., not, no, never, incredible), affixes (e.g., -nâĂŹt, un-, any-).
	38	Antcpnt	Anticipation [19]	Coming or acting in advance: clouds anticipant of a storm. 2. Expectant; anticipating: a team anticipant of victory.
	39	ADV_EXM	Example Words*	such as words are all form such as or like /hamÄ■on / «همچون», /hamanand/ «همانند», (i.e. for example) /masalan/ «مثلاً», âĂę
	40	IF	Conditional words*	The conditional conjunctions (such as, if) that are frequently found in conditional sentences.
	41	GT	GeneralTerms [6]	Refers to a person (or object) as a class of persons or objects that includes the person (or object).
	42	UNT	Un_Trust*	Words like lack of trust, distrust, suspicion, mistrust, doubt, disbelief, dubiety, wariness and so on.

4.2.1 Characteristics of an important rumor

Two categories of textual features, newsworthy and emotional, are introduced to compute the importance score of the FR, that highlight the newsworthy and the emotional level of the text.

- ***Computing the emotional score of a rumor***
  Based on our result of student's t-test on the several thousand rumors in Persian language, it became clear that rumormongers increase the emotional aspect of the text by utilizing textual features such as adjectives, adverbs, motion verbs, emotional words and so on. These features are described in the following:

  1. **Emotiveness:** Adjectives(Adj), Adverbs(Adv) describe things and modify other words so change our understanding of things.
  $$E\_Tag(T) = \frac{|Adj(T)| + |Adv(T)|}{|Noun(T)| + |Verb(T)|} \tag{2}$$
  Where, $E_Tag(T)$ is the ratio of adjectives ($|Adj(T)|$) plus adverbs ($|Adv(T)|$) to nouns ($|Noun(T)|$) plus verbs ($|Verb(T)|$) in the text *T*, which is selected as an indication of expressivity of language [28].
  2. **Emotional words (6 features):** Rumormongers use the power of emotional words to cause fear, concern, and hatred in the audience. In this study, the National Research Council Canada (NRC) [30] emotion lexicon is utilized to obtain the emotional score of words in the context of Persian rumors in eight basic emotions which are anger, fear, anticipation, trust, surprise, sadness, joy, and disgust. Saif et al. [30] in November 2017 provide versions of the lexicon in over one hundred languages such as the Persian language. The English terms in this lexicon have been translated using Google Translate. Researchers of the current paper manually reviewed and corrected each of the machine translation words with the help of two linguists. The experimental results (Table 4) show that rumormongers use five categories of Emotional words including, Fear (F), Surprise (SU), Disgust (D), Sadness (Sad), Anger (AN) in order to increase the emotional score, so in the current paper, these categories are evaluated on the Persian rumors. Additionally, Motion Verbs (MV) is often used in rumors to enhance that score. For this purpose, the set of MVs in the Persian language is used that is collected by Golfam et al. [31]. They extracted and analyzed 126 MVs from Dadegan site[1] and Persian Language Database[2] and other written sources. The list of MVs is narrowed down by selecting MVs which often appear in FRs.
  $$EM\_Words(T) = \frac{\sum_{i=1}^{|S(T)|} |EM\_Wrd(S_i(T))|}{|S(T)|} \tag{3}$$
  Score of each of emotional features $F_{EM\_Words} = \{F, SU, D, Sad, AN, MV\}$ is calculated by formula 3, the fraction of sentences containing emotional feature $|EM\_Wrd(S_i(T))|$ to all sentences $|S(T)|$ of the text *T* is calculated.
  3. **Consecutive Words or Characters (2 features):** Consecutive words or characters within a sentence in every language is syntactically incorrect. However, rumormongers try to emphasize the main subject of rumor by repeating consecutive words in the FR. For example, «توجه توجه» /Attention Attention, is a Consecutive Word (CW) and words like «سلاااااااام» /"salÃćÃćÃćÃćÃćÃćÃćm"/"Hellllllllloooo" and «هشداااااار» /"HoÅądÃćÃćÃćÃćÃćÃćr"/"Alaaaaarm" are words with Consecutive Chars (CC).
  $$CC(T) = \frac{\sum_{i=1}^{|S(T)|} |CC(S_i(T))|}{|S(T)|} \tag{4}$$
  $$CW(T) = \frac{\sum_{i=1}^{|S(T)|} |CW(S_i(T))|}{|S(T)|} \tag{5}$$
  In formulas 4 and 5 are respectively computed, the fraction of sentences containing CC and CW to all sentences $|S(T)|$ of the text *T* is calculated. Where $|CC(S_i(T))|$ and $|CW(S_i(T))|$ indicates whether sentence S_i of the text *T* contains CC and CW or not respectively.
  4. **Sentiment (2 features):** News sentences usually do not convey any sentiment. For example, in the sentence *"The Coaches of the Persepolis and Oil teams of Tehran, Branko and Ali Daei planted a tree seedling in the league organization on the occasion of the arbor day."*, there is no excitement.

[1] www.dadegan.ir
[2] www.pldb.ihcs.ac.ir

Nevertheless, rumors contain several characteristic sentiments (e.g., anger) compared to other types of information [32]. On the other hand, there is a general lay belief that FRs are dominated by negative sentiment and polarity [33]. Despite the fact that rumors contain negative polarity, they almost are expressed in positive polarity. The concepts of sentiment polarity of Zhang and Skiena [34] is utilized and calculated the sentiment score of the rumor text using a lexicon-based method (Formulas 3 and 4).

$$PS\_Sent(T) = \frac{|PSntm(T)|}{|PSntm(T)| + |NSntm(T)|} \tag{6}$$

$$NG\_Sent(T) = \frac{|NSntm(T)|}{|PSntm(T)| + |NSntm(T)|} \tag{7}$$

Where $|PSntm(T)|$ is the number of positive terms and $|NSntm(T)|$ is the number of negative terms in the text *T*. The NRC Emotion Lexicon [30] is used to obtain the sentiment score of Persian words in one of the positive (1), negative (-1), or neutral (0) polarities.

5. **Speech Act (SA) (2 features):** The importance of a message from the person's point of view increases when the message transmits critical news and motivates fear in the audience. Hence, individuals spread rumors when they feel anxiety or threat. Also, to increase the speed of the release of a rumor, the rumormonger asks audiences to notify the message as soon as possible to his or her relatives. Rumormongers warns the audience that if they do not inform others, a bad event is may happen. Therefore, identifying speech acts of rumors can help to distinguish FRs from TRs. Based on the obtained results in [35], FRs are often expressed in four SA classes, including threat (SA_Thrt), declaration (SA_Dec), question (SA_Ques), request (SA_Req). In this study, the two types of SA "SA_Thrt" and "SA_Req" are considered as two effective factors in increasing the emotional score of a text.
6. **Start and End of the text (2 features):** The beginning and ending sentences of FRs can have a significant effect on attracting the attention of the audience because the author expresses the main purpose in these sentences. Therefore, the first sentence based on emotional words, the final sentence based on both emotional words and Words associated with the request are analyzed. Both of these features have a binary value, which indicates whether the text has start and end sentences with the specified attributes.

These emotional-based features have a different impact on measuring the emotional score of the text. Therefore, it is necessary to consider the influence coefficient of each feature in distinguishing the emotional measure of FR texts from TR. Therefore, the feature weighting techniques are used to obtain these coefficients, so that for each of these parameters, the best coefficients are obtained, that these coefficients indicate the effect of that feature in distinguishing FRs and TRs.

$$Emotional(T) = \sum_{j=1}^{|F|} \frac{\sum_{i=1}^{|S(T)|} W\_F_j * |F_j(S_i(T))|}{|S(T)|} \tag{8}$$

Where, $|S(T)|$ is the total number of sentences in the text *T*. W_F_j is weight of feature F_j. The weight of feature calculates in section 5.3 by the PSO algorithm. $F_j(S_i(T))$ compute the fraction of sentences of the text *T* containing features
$F_{Emotional} = E\_Tag, F, SU, D, Sad, AN, MV, PS\_Sent, NG\_Sent, CW, CC, SA\_Thrt, SA\_Req, Start, End.$

- ***Computing the newsworthy of a rumor***

The main reason for accepting an FR as valid news by a person is that it is newsworthy. News can be defined as "newsworthy information about recent events or happenings, especially as reported by news media.". In this study, the set of the textual variables, including timing, statistical information, human interest, lexical diversity, certainty, and spelling mistakes are introduced to evaluate the newsworthy of text. These factors detailed below.

7. ***Relative Time (RT):*** The timing or novelty of news is of particular importance. Making news of the day with past events with a new angle or view is part of the daily work of journalists. If a story happens today, it is news. However, if the same thing happened last week, it is no longer interesting. Also, rumormongers use RTs in FRs to apparently display new news, or try to pretend that an important event will happen soon. For example, for over three years, the rumor "Recently, Google has put an Internet voting to change the name of the Persian Gulf" be released on social networks. In order to, we extracted words with the ADV_TIME tag as relative times.

$$RT(T) = \frac{\sum_{i=1}^{|S(T)|} |RT(S_i(T))|}{|S(T)|} \tag{9}$$

Where, $|RT(S_i(T))|$ indicates whether sentence S_i of the text *T* contains RT or not.

8. ***Statistical Information (SI) (2 features):*** Zhang et al.[36] found that rumors that are short and contain numbers are more likely to be true than those that are long and do not contain any quantitative details. According to studies conducted on Persian FRs, it can be seen that rumormongers try to attract the attention of the audience with fake statistics.
9. ***Named Entity (NE):*** Most people follow the topics that are discussed about celebrities. Celebrities as NEs are high-interest items for individuals. They publish related news to famous people based on popularity or disgust. Therefore, rumormongers utilize the names of famous people, scientists, philosophers, or organizations or institutions to increase the newsworthy of the rumor. In this study, NEs are extracted using a Hidden Markov Model (HMM)-based model [37] in three classes, the person's name, organization, and location.
$$NE(T) = \frac{\sum_{i=1}^{|S(T)|} |NE(S_i(T))|}{|S(T)|} \tag{10}$$
Where $|NE(S_i(T))|$ indicates whether sentence S_i of the text *T* contain NE terms or not.
10. ***Lexical Diversity (LD):*** Rumormongers try to attract the attention of audiences to the issue of rumor. Therefore, they repeat the important and emotional words on the subject of rumor. Therefore, the use of repetitive words in the text reduces the LD of the text. Thereby, FRs have a few LD due to the high repetition of tokens.
$$LD(T) = \frac{|V_{uniq}(T))|}{|V(T)|} \tag{11}$$
Therefore, the ratio of the number of unique vocabulary ($|V_{uniq}(T)|$) to the total number of vocabularies $|V(T)|$ in the text *T* is calculated as the LD score of the text *T* [28].
11. ***Certainty (CE):*** Rumormongers use certainty-related words to hide their lies about FR story. The existence of the certainty-related words in FRs increases the audience's trust to the subject. So the audience mistakenly believes the rumor and sends it to others. The criterion of certainty as a characteristic to increase the newsworthy of rumor for the audience is calculated by formula 12.
$$CER(T) = \frac{|V_{Cer}(T))|}{|V_{Uncer}(T)| + |V_{Cer}(T)|} \tag{12}$$
Where, $|V_{Uncer}(T)|$ and $|V_{Cer}(T)|$ are respectively the number of uncertainty-related vocabularies and certainty-related vocabularies in text *T*.
12. ***SA_Dec and SA_Quot (2 features):*** These two types of SA give a formal concept to text. Therefore, rumormongers utilize this type of SAs to express their stories formally.
13. ***Superiority Features(SF) (2 features):*** Rumormongers attempt to upgrade (one thing or person is superior to another) or diminish the main element of rumor using the superiority attributes (Adj_SUP) and ordinal adjectives (Adj_ORD). Both of these features have a Boolean value for each sentence, meaning it is checked whether the sentence contains these attributes or not.
Finally, the Newsworthy score of text is computed by formula 13.
$$Newsworthy(T) = \sum_{j=1}^{|F|} \frac{\sum_{i=1}^{|S(T)|} W\_F_j * |F_j(S_i(T))|}{|S(T)|} \tag{13}$$
Where,
$F_{Newsworthy} = TI, Adj\_QUA, NumDigit, NE, LD, CER, SA\_dec, SA\_uot, Adj\_SUP, Adj\_ORD.$
Also, the weights W_F_j is weight of feature F_j that indicates its impact. Then, the importance score of the text *T* is calculated by formula 14.
$$Importance(T) = \sum_{j=1}^{|F|} \frac{\sum_{i=1}^{|S(T)|} W\_F_j * |F_j(S_i(T))|}{|S(T)|} \tag{14}$$
Where, $F = F_{Emotional} \cup F_{Importance}$. It should be noted that the sum of Emotional and Newsworthy weights is equal to one.

4.2.2 Characteristics of an ambiguous rumor

The second component of rumors management is ambiguity. The essence of rumors is in their ambiguity, where the ambiguity of evidence makes the process of rumors spreading more widely [39]. The ambiguity arises when, either the news is received in distorted form, or the person received contradictory news, or one cannot understand such news. The ambiguous expression of news challenges the audience. In the following, a set of features is introduced that the presence of those features in the text causes ambiguity in subject.

Table 2: A summary of the content features that are introduced and evaluated, but are not considered in RSPMM for calculating the spread power of rumor.

#	Abbreviation	Feature	Description
1	Wrd	Word	Number of words in the rumor.
2	Vrb	Verb	Number of verbs in the rumor.
3	S	Sentence	Number of sentences in the rumor.
4	WL	Word Length	Average length of words in the rumor (Number of character /Wrd)
5	SL	Sentence Length	Average length of sentences in the rumor (Wrd/S)
6	RD	Readability	Using a machine learning model for Persian text readability assessment [38]
7	DDT	Depth of Dependency Tree	The grammatical complexity of the text (depth of its dependency parse tree)
8	Cls	Clause	Average of pause in sentences

14. ***Uncertainty (UNCer):*** Rumormongers try to highlight the ambiguity of the subject by using suspicious and surprising words, so that challenge the mind of the audience about what is behind the rumor's rumor. Therefore, a collection of uncertainty-based words in the Persian language is extracted to measure the uncertainty score in rumor text.
$$UNCer(T) = \frac{|V_{Uncer}(T))|}{|V_{Uncer}(T)| + |V_{Cer}(T)|} \tag{15}$$

15. ***Sensory Verbs (SV):*** When a rumormonger creates a rumor, there is a clear sign in the sentence that indicates that he has personally seen or heard what he speaks about it, or it is the result of his reasoning and speculation. These signs are SVs that create evidentiality in rumors. Evidentiality is a grammatical category that its role is to show the source of information. Of course, these verbs appear in cases where the rumormongers want to increase the rumor's credibility, so they use these verbs as a means to emphasize the rumor.
$$SV(T) = \frac{\sum_{i=1}^{|S(T)|} |SV(S_i(T))|}{|S(T)|} \tag{16}$$
The phrase $|SV(S_i(T))|$ has a Boolean value that indicates whether the sentence S_i of the text *T* has a sensory verb or not.

16. ***Question Tokens (2 feature):*** The question sentences cause ambiguity and disturb the reader's mind. Therefore, Question Words (Q_W), Question Mark (Q_M), and SA_ques are considered as factors that create ambiguity in the content of the rumor and calculated by formulas 17 and 18.
$$Q\_W(T) = \frac{\sum_{i=1}^{|S(T)|} |QW(S_i(T))|}{|S(T)|} \tag{17}$$
$$Q\_M(T) = \frac{\sum_{i=1}^{|S(T)|} |QM(S_i(T))|}{|S(T)|} \tag{18}$$
Where $|QW(S\_i(T))|$ and $|QM(S\_i(T))|$ separately means that the sentence S_i of text *T* has at least one question word or question mark or not.

17. ***Exclamation Mark (EM):*** A punctuation mark is usually used after an interjection or exclamation to indicate strong feelings or high volume (shouting) or to show emphasis. Exclamation mark used for any other purpose, as to draw attention to an obvious mistake, in road warning signs, (in chess commentaries) beside the notation of a move considered a good one, (in mathematics) as a symbol of the factorial function, or (in logic) occurring with an existential quantifier.

18. ***Pronouns (Pro):*** Rumormongers in their rumors use the fewer self-reference (first-person singular pronoun) and more group-reference (first-person plural pronoun) and other references (third-person pronouns) to create non-immediacy and uncertainty.
$$PRO(T) = \frac{\sum_{i=1}^{|S(T)|} |Pronoun(S_i(T))|}{|S(T)|} \tag{19}$$

Where, $|Pronoun(S_i(T))|$ is the number of sentences containing pronoun (i.e., third-person and first-person plural pronouns).

19. ***Tentative (Tentv):*** The adjective tentative is used to describe what is unclear. Therefore, rumormongers utilized these types of words to create a sense of hesitation in the audience and engage minds.

$$Tentv(T) = \frac{\sum_{i=1}^{|S(T)|} |Tentative(S_i(T))|}{|S(T)|} \tag{20}$$

Thus, a fraction of the sentences S_i of the text *T* containing the tentative-based words ($|Tentv(S_i(T))|$) is calculated by formula 20.

20. ***Negation (Neg):*** In rumors, the use of negative words refers to two purposes: (1) creating negative emotions, (2) an unusual expression of the news event. In Persian language seven negative prefixes are used to build words with negative or contrastive meaning. These prefixes are: (1)'bi-'/im-(e.g, impolite), (2)un-, in- (e.g., injustice), no-),(3)'zed-'/unti-(e.g., unti-security), (4)'gheir-'/un-(e.g., Unnecessary), (5)'ne-'/'na-'/not(e.g., nemidanam/I do not see, nasalem/unhealthy), (6) hich-/no- (e.g., nobody), (7)la-/without.

$$Neg(T) = \frac{\sum_{i=1}^{|S(T)|} |Negation(S_i(T))|}{|S(T)|} \tag{21}$$

A fraction of the sentences containing the negative prefixes are calculated by formula 21.

21. ***Anticipant (Ancpnt):*** Many people are interested in predicting many events, so they try to anticipate the most likely problems, but it is impossible to be prepared for each eventuality. The rumormongers also intend to create fear and turmoil in the community by anticipating unpopular events that have not yet happened.

$$Ancpnt(T) = \frac{\sum_{i=1}^{|S(T)|} |Anticipant(S_i(T))|}{|S(T)|} \tag{22}$$

Where, $|Anticipant(S_i(T))|$ is a binary value that indicate whether the sentences S_i of the text *T* contains the anticipate-related words or not.

22. ***UNTrust (UNT):*** In expressing news about famous people or important factors of society, the existence of words containing untrust causes doubts about that subject in the mind of the audience.

$$UNT(T) = \frac{\sum_{i=1}^{|S(T)|} |UNTrust(S_i(T))|}{|S(T)|} \tag{23}$$

Where, $|UNT(T))|$ is a fraction of the sentences S_i of the text *T* containing the untrust-based words are calculated. Finaly, the ambiguity score of a text is calculated by formula 24.

$$Ambiguity(T) = \sum_{j=1}^{|F|} \frac{\sum_{i=1}^{|S(T)|} W\_F_j * |F_j(S_i(T))|}{|S(T)|} \tag{24}$$

Where, $|F|$ is $F_{Ambiguity} = \{UNCer, SV, QW, QM, SA\_ques, EM, PRO, Tntv, Neg, Antcpnt, UNT\}$ and W_F_j is weight of features F_j.

4.3 Learning

In this step, Particle Swarm Optimization (PSO) [40] as a weighting algorithm is used to find optimal weights for each feature that these weights are used in the process of calculating SPR. The purpose of feature weighting is to determine the degree of importance of each feature in predicting two classes FR and TR. Because, different features can have different levels of importance for class prediction in classification problems, so the weight value of each feature will also be effective in calculating the spread power of rumor. To gain this degree of importance, feature weighting techniques are used attempting to influence the distance function by giving different weights to different features. The output of weighting algorithms is the weighted value of each feature of the dataset. Therefore, high-weight features will be more effective in the classification results. The process of feature weighting as follows:

1. Feeding the learning component by the extracted features of the previous step into two categories of importance and ambiguity and employing PSO to obtain optimal feature weights and kernel parameters.
2. Utilizing the cross-validation method to separate dataset into training and testing set.

Table 3: Distribution of manually annotated rumors used for training and evaluating the rumor verification system.

Dataset	FRs	TRs	Description
Twitter [12]	783	783	Crawling Twitter rumors from two Iranian websites, Gomaneh.com and Shayeaat.ir which publish Persian rumors and annotating by Zamani et al..
Telegram [41]	882	882	Crawling Telegram rumors by using provided API by Computerized Intelligent Systems (ComInSys)[3] Lab of the University of Tabriz, between May 1, 2017 and November 30, 2018 and annotating in two class including TR and FR.

3. Setting up parameters of PSO for each training set, generating randomly all particlesâĂŹ positions and velocity, setting up the learning parameters, the inertia weight, and the maximum number of iterations,
4. Updating the velocities of all particles at each iteration.
5. Training SVM classifier according to particles values.
6. Calculating the corresponding fitness function for each particle.
7. Obtaining the best gene weights and best kernel parameters values.
8. Training SVM classifier with obtained parameters.
9. Updating the inertia weight and return to step 4.

4.4 Power calculation

After, features extracted from the step of feature extraction and weights extracted by PSO for each feature. First, two score of importance and ambiguity is calculated based on features and their weights. Finally, the power calculates by the multiplication of these two factors.

5 EXPERIMENTS AND RESULTS

This section evaluates the SPR criterion as a new proposed criterion that can be used to validate rumors. For this purpose, three experiments are performed on data sets of Twitter and Telegram to evaluate SPR and answer the research questions: (1) investigating the significant difference of each of the selected features between two groups of FRs and TRs. (2) measuring SPR based on proposed new features. (3) application of power criterion as an effective feature in validating rumors. Experimental results illustrate the efficacy and efficiency of the proposed new features in rumors verification. In the experiments presented subsequently, 10-fold cross-validation is used. Also, to evaluate the performance of RSPMM using Random Forest (RF) classifier, calculate Precision, Recall, F-measure based on two classes of FR and TR. In this section, the experimental details described.

5.1 Data Collection and Dataset of Rumors

This study evaluates SPR on Persian rumors from two different sources including Twitter and Telegram. The detailes of these two datasets in table 3 is described.

5.1.1 Twitter

Twitter is a micro-blogging social network service where users can publish and exchange short messages of up to 280 characters long; these messages called tweets. Accessibility, speed, and ease-of-use have made Twitter a valuable social medium for a variety of purposes that use of it is exponentially growing. It also is a useful source for gathering data and serves as one of the foremost media for research in NLP.

5.1.2 Telegram

Telegram is an instant messaging service. Due to the popularity of the Telegram in Iran and the dissemination of messages through it, SPR published on the Telegram is tested and evaluated. Therefore, a few thousand Persian posts from Telegram channels in Iran are crawled and extracted by using provided API by ComInSys lab of the University of Tabriz. Due to the lack of available publicly available dataset of labeled rumors for Telegram, the researchers collected a set of Persian posts from Telegram in various topics. Then, several channels of Telegram (i.e., Fars News Agency, Iranian Students' News Agency (ISNA), Tasnim News Agency, Tabnak, Nasim News Agency (NNA), Mehr News Agency (MNA), Islamic Republic News Agency (IRNA)) and three websites (i.e., Gomaneh. com, Wikihoax.org, Shayeaat.ir) are selected as trustworthy sources and used for verifying of rumors. Three undergraduate annotators manually annotate collected data as TR and FR by crawling mentioned trustworthy sources. Consequently, a dataset of 882 Persian FRs and equal numbers of TR from Telegram posts added to the dataset by crawling several Telegram news channels. FRs are tagged with '0', and TRs with '1'.

5.2 Data Analysis

In order to better understand the impact of various features on calculating SPR, a set of statistical analyses such as the T-test is performed on each feature. Also, the distribution of features in both TR and FR categories is represented by boxplots. In this study, 51 features are proposed to use in the process of calculating SPR. Based on the experimental results, it was found that some of these features could not distinguish FR from TR, or did not have a good impact on the calculation of SPR.

5.2.1 Student,s t-test

Since our samples are independent, an independent samples T-test is run on the proposed features. An Independent Samples t-test compares the means and P-value of each feature for two groups FR and TR. StudentâĂŹs paired t-test is performed on the 56 features that were originally proposed to calculate the spread power of rumor. NULL hypothesis is rejected if $P < 0.05$. In this study, the null hypothesis is defined as follows:

Null hypothesis: The spread power of FRs is equal to the TRs.

Therefore, StudentâĂŹs t-test is carried out and the result of P-value for each of the features listed in Table 4, with the hypothesis that each feature appears with a different frequency in FRs and TRs and can discriminate between them. The P-value results ($<= 0.05$) demonstrate that most of these features reveal statistically significant differences between FR and TR texts.

According to the results of the t-test on the 51 features listed in Table 4 and the classification of FRs and TRs based on these features, 42 features (Table 1) out of 51 features are selected as effective and informative features in distinguishing FRs from TRs and are useful features in calculating SPR. Features that have been removed from the list of selected features including word count, sentence count, the mean length of sentences (LS), average word length (LW), the narrative speech act (SA-Narrtv), readability (RD), Clause (Cls) and the depth of dependency tree (DDT) (Table 2).

Table 4: The average frequencies in 43 proposed features in TR and FR, AND t-test student results (those values that are greater than 0.05 are italicized).

	Word	**Sentence**	**LS**	**LW**	**RD**	**Clause**	**DDT**	**Verb**
Average frequency in FR	93.39	4.64	23.48	1.85	0.543	0.093	0.192	0.057
Average frequency in TR	53.37	2.05	29.11	1.97	0.593	0.094	0.227	0.030
P-value	*1.77E-11*	*9.59E-21*	*3.04E-05*	*3.87E-16*	*0.017*	0.426	*2.29E-09*	*1.81E-13*
	PS_Sent	**NG_Sent**	**IF**	**NER**	**Antcpnt**	**E_Tag**	**UNT**	**Adj_ORD**
Average frequency in FR	0.392	0.301	0.054	0.441	0.557	0.227	0.909	0.058
Average frequency in TR	0.455	0.234	0.033	0.140	0.435	0.224	0.935	0.032
P-value	*0.002*	*0.001*	*0.010*	*7.83E-32*	*5.1E-10*	0.330	*9.19E-16*	*0.005*
	NER	**Emotional**	**Importance**	**Ambiguity**	**Newsworthy**	**Power**	**AFF**	**Adj_EXM**
Average frequency in FR	0.779	0.248	5.490	2.803	0.235	0.056	2.597	0.355
Average frequency in TR	0.509	0.237	4.272	2.658	0.299	0.0487	2.015	0.444
P-value	*1.54E-36*	0.121	*1.82E-27*	*0.003*	*1.71E-28*	*0.002*	*2.4E-12*	*2.32E-30*
	Adj_Qua	**MV**	**SM**	**CC**	**CW**	**LD**	**PRO**	**RT**
Average frequency in FR	0.192	0.486	0.052	0.494	0.025	0.761	0.175	0.380
Average frequency in TR	0.227	0.609	0.040	0.008	0.007	0.813	0.106	0.310
P-value	*2.76E-09*	*3.67E-08*	*0.002*	*2.23E-08*	*0.001*	*6.28E-12*	*6.83E-06*	*0.001*
	Adj_CMPR	**Unec**	**Ce**	**QW**	**QM**	**EM**	**Adj_SUP**	**Neg**
Average frequency in FR	0.076	0.606	0.150	0.101	0.593	0.090	0.056	0.081
Average frequency in TR	0.032	0.611	0.088	0.043	0.543	0.034	0.038	0.70
P-value	*3.18E-05*	0.345	*5.45E-12*	*2.15E-07*	*0.017*	*1.12E-30*	*0.039*	0.182
	Start	**END**	**AN**	**Sad**	**D**	**F**	**SU**	**Tntv**
Average frequency in FR	0.502	0.441	0.476	0.543	0.284	0.610	0.350	0.209
Average frequency in TR	0.200	0.140	0.324	0.411	0.250	0.433	0.230	0.154
P-value	*3.45E-29*	*7.83E-32*	*4.58E-11*	*1.57E-08*	0.057	*3.50E-14*	*3.14E-08*	*0.001*
	SA_Ques	**SA_Req**	**SA_Dir**	**SA_Thre**	**SA_Qout**	**SA_Dec**	**SA_Narrtv**	**SV**
Average frequency in FR	0.070	0.002	0.008	0.038	0.038	0.318	0.545	0.024
Average frequency in TR	0.024	3.27E-05	0.002	0.014	0.040	0.572	0.348	0.005
P-value	*2.53E-05*	0.159	0.083	*0.003*	0.433	*9.92E-21q*	*2.30E-21*	*3.3.E-05*

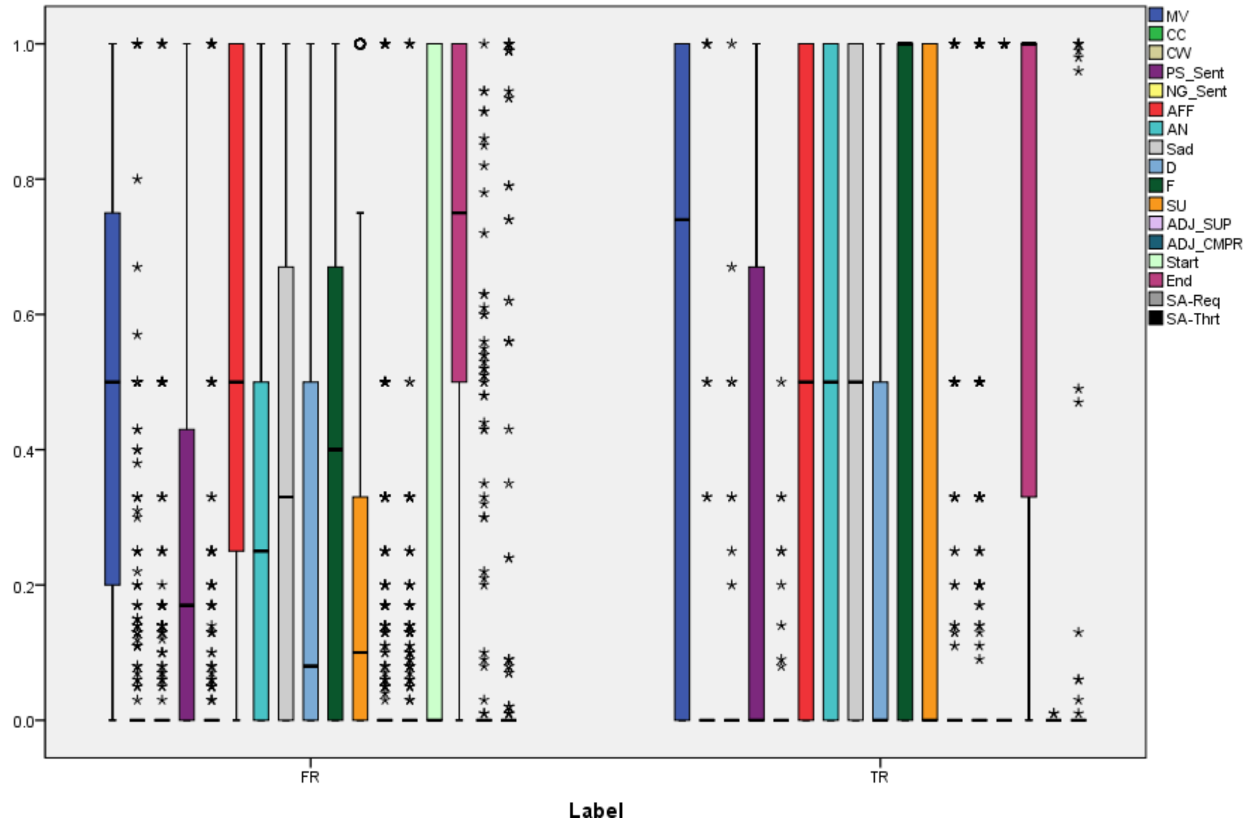

Figure 2: Graphically depicting the distribution of emotion-based features by boxplots in two classes of FR (0) and TR (1) to compute the importance score.

Table 5 demonstrates the result of studentâĂŹs t-test for "Power" feature. Since, $p - value = 0.002 \leq 0.05$, the null hypothesis is rejected for SPR. Based on the t-test, the power benchmark can be used as a feature in the process of categorizing rumors since it is able to distinguish FRs from TRs. Also, the average of the spread power in FR texts is more than TR texts. So, it can be concluded that the spread power of FRs is more than TR texts.

5.2.2 Graphical representation of feature distribution

The distribution of introduced featured for computing the spread power of a rumor is displayed using the boxplots. The boxplot is a standardized way of displaying the distribution of data based on the five number summary: minimum, first quartile, median, third quartile, and maximum. Graphical representation of the distribution of features for FR and TR in three categories, "Emotional", "Newsworthy" and "Ambiguity" are shown in Figures 2, 3 and 4, respectively. In addition, Figure 5 illustrates the discriminative capacity of five factors of "Emotional", "Importance", "Ambiguity", and "Power" in two classes of FR and TR. As shown in the boxplot diagram, the three features of ambiguity, emotional, and power in the FRs have a high distribution than TRs.

5.2.3 The average of SPR on Twitter and Telegram

RSPMM is investigated on propagated FRs and TRs on Twitter and Telegram. Therefore, the spread power of 1566 FR and TR on Twitter and 1764 FR and TR on Telegram is calculated. The results of the statistical analysis (Table 6) on these two datasets (Twitter and Telegram) show that the average of the spread power of FRs is higher than TRs in both social networks. Therefore, it can be concluded that the characteristics of a fast-spreading rumor in FRs are more than TRs. As shown in Table 6, the average propagation power in the Twitter data set is lower than the Telegram data set. The reason for this is that Twitter tweets collected by Zamani et al. [12] are limited to 140 characters and the full text of the tweets is referenced in links. Thus, little textual information is available to analyse the textual features based on RSPMM.

Table 5: Independent T-test values for power

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	**Sig**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval of the Difference**	
									Lower	**Upper**
Power	Equal variances assumed	0.562	0.453	3.155	1206	**0.002**	.00764	0.0024201	0.0028869	
	Equal variances not assumed			3.157	1199.372	**0.002**	0.00764	0.0024187	0.0028896	0.0123804

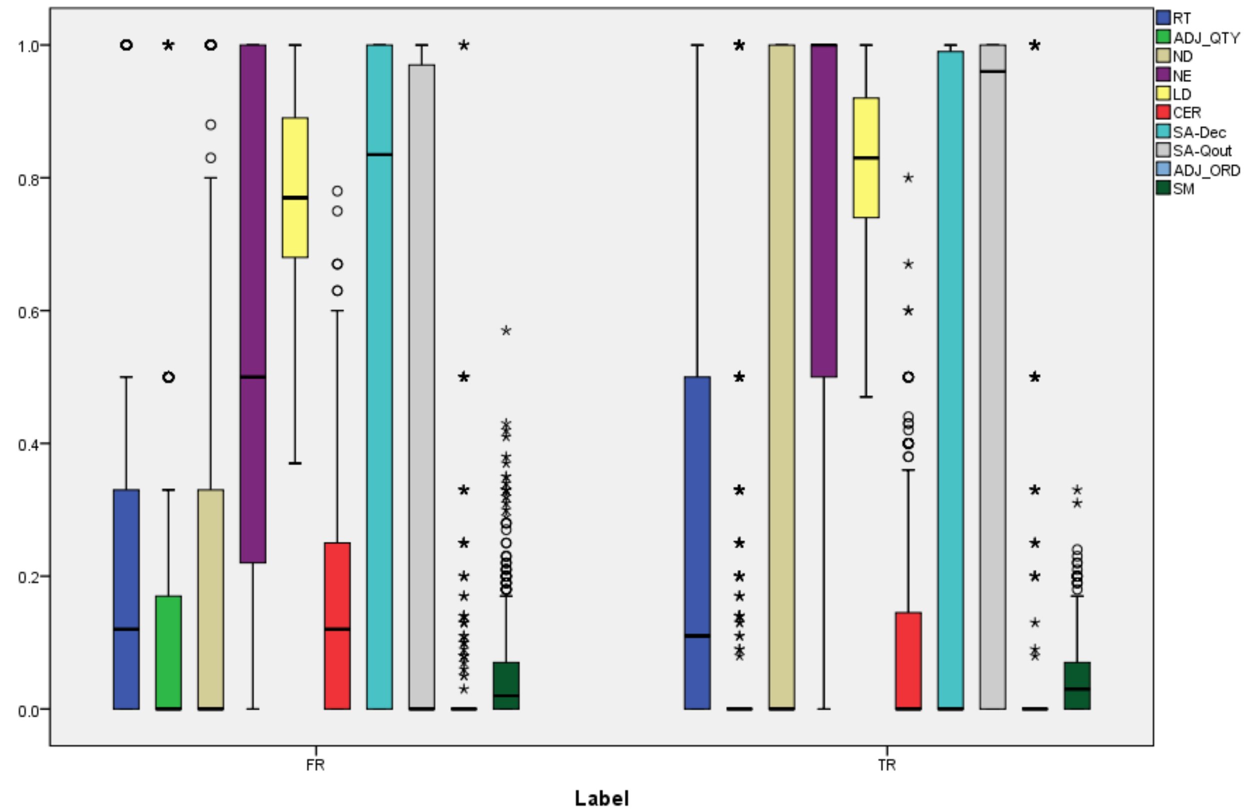

Figure 3: Graphically depicting the distribution of newsworthy-based features by boxplots in two classes of FR (0) and TR (1) to compute the importance score.

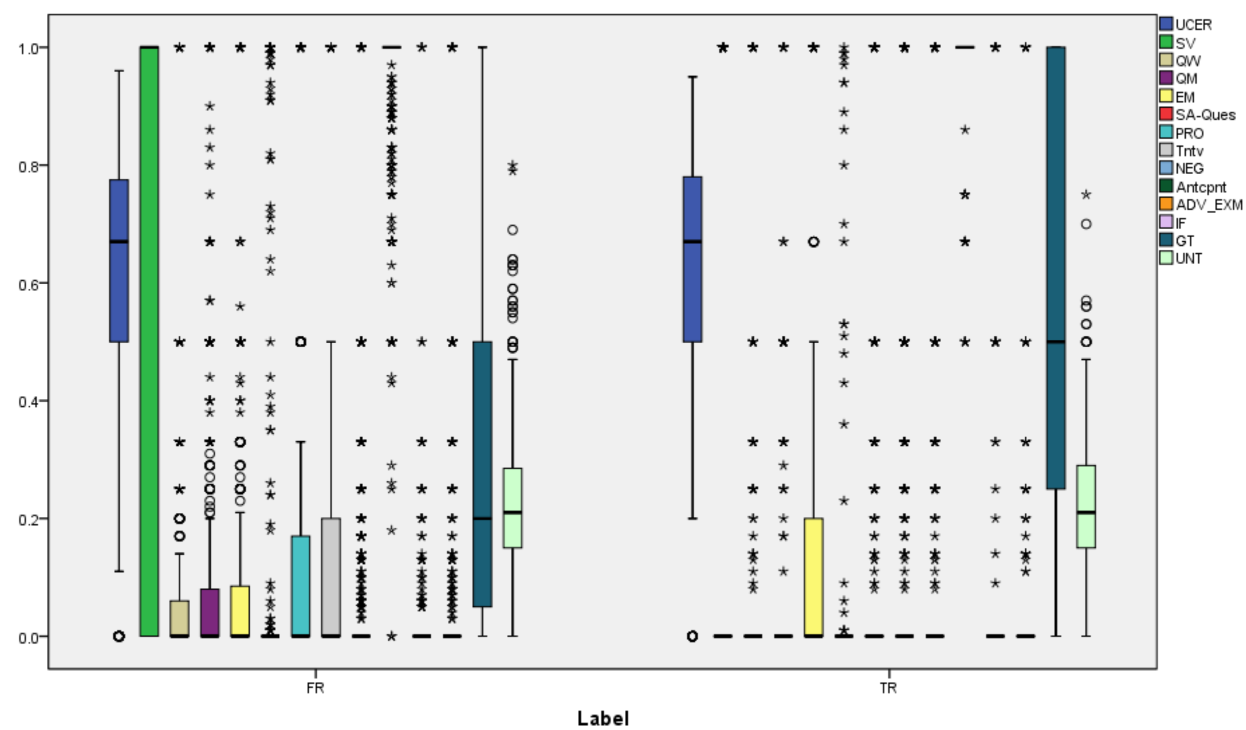

Figure 4: Graphically depicting the distribution of ambiguity-based features by boxplots in two classes of FR (0) and TR (1) to compute the ambiguity score.

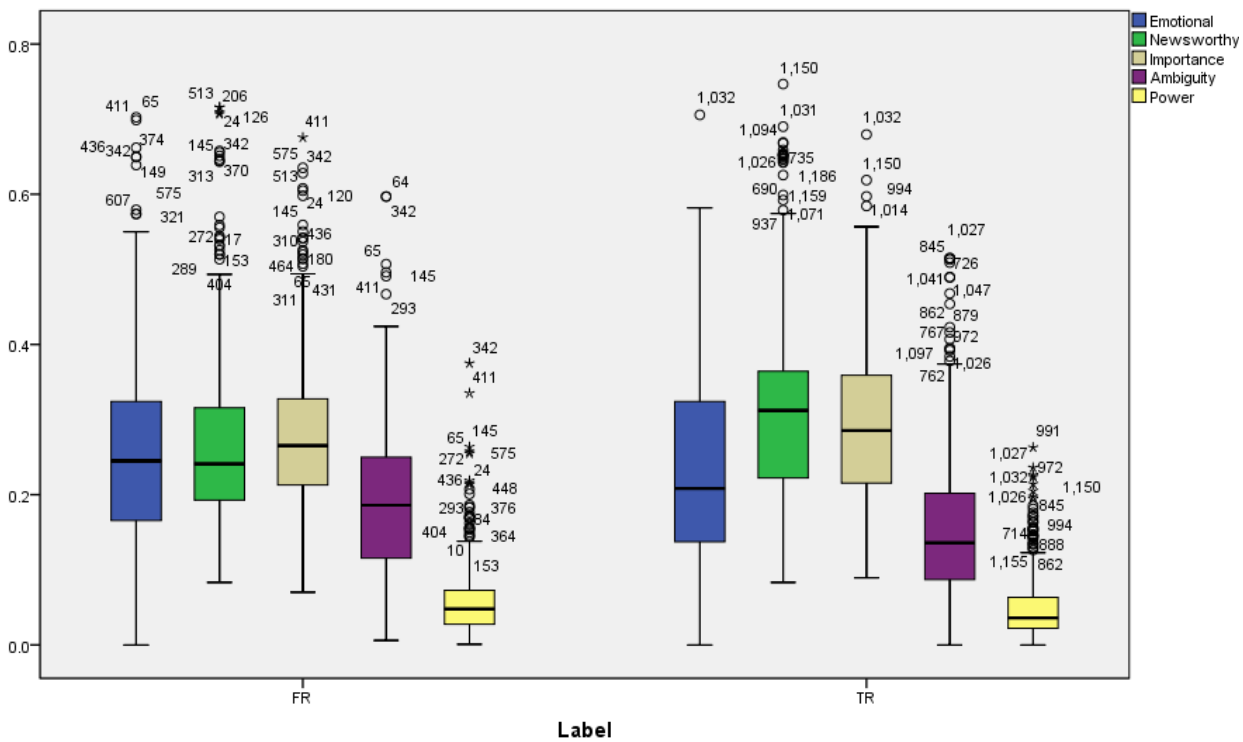

Figure 5: Graphically depicting the distribution of five informative features to compute the power score in two classes of FR (0) and TR (1) by box plots.

Table 6: Comparing the the average values of spread power in FRs and TRs on Twitter and Telegram.

Dataset	Category	Avg. Importance	Avg. Ambiguity	Avg. Spread power
Telegram	FR	0.277	0.190	0.056
	TR	0.293	0.156	0.049
Twitter	FR	0.183	0.126	0.039
	TR	0.212	0.092	0.024

5.3 Feature Weighting

Two techniques of Particle Swarm Optimization (PSO) [40] and Forest Optimization Algorithm (FOA) [42] is utilized to find optimal weights for each feature. Feature weighting is a continuous space search problem and attempts to influence the distance function by giving different weights to different features. The output of weighting algorithms is the weighted value of each feature of the dataset. These weights are used in the process of calculating SPR. Thereby; high-weight features will be more effective in results of classification. Based on the evaluation of the weights obtained from each of these two methods of weighting and calculating the spread power of TR and FR, the PSO technique is selected as the best weighting method.

After determining the weight of each feature and calculating SPR, the performance of the RSPMM is evaluated separately using RF classifier based on the rumors published in Telegram and Twitter in two separate experiments (Table 7). The first time, without weighting features and the second time with weighing features is calculated SPR. Table 7 represents the optimal impact of weighted features in the process of calculating SPR. Despite the limited length of Twitter tweets, the results of the evaluation of RSPMM on tweets were very close to the results of RSPMM on Telegram posts.

Table 7 represents the optimal impact of weighted features in the process of calculating SPR. Despite the limited length of Twitter tweets, the results of the evaluation of RSPMM on tweets were very close to the results of RSPMM on Telegram posts.

Table 7: The precision, recall, and f-score values for RF classifier based on proposed features, in calculating the spread power of a text during rumor detection on Telegram and Twitter (with and without feature weighting).

Dataset	Category	Recal (with/without)	Precision (with/without)	F-measure (with/without)
Telegram	FR	0.814 / 0.781	0.791 / 0.742	0.802 / 0.760
	TR	0.803 / 0.729	0.825 / 0.768	0.814 / 0.751
	Avg.	0.808 / 0.755	0.808 / 0.755	0.808 / 0.755
Twitter	FR	0.746 / 0.712	0.772 / 0.750	0.759 / 0.730
	TR	0.780 / 0.763	0.754 / 0.726	0.766 / 0.743
	Avg.	0.763 / 0.738	0.763 / 0.738	0.763 / 0.737

Table 8: The effect of different features in rumor detection in Telegram using RF classifier.

	TP Rate	FP Rate	Precision	Recall	F-Measure
(1) Previous features					
FR	0.753	0.230	0.766	0.753	0.759
TR	0.770	0.247	0.757	0.770	0.764
Avg.	0.762	0.238	0.762	0.762	0.762
(2) New proposed features					
FR	0.790	0.208	0.792	0.790	0.791
TR	0.792	0.210	0.791	0.792	0.792
Avg.	0.791	0.209	0.791	0.791	0.791
(3) New features + "power of spread" factor					
FR	0.791	0.175	0.814	0.791	0.802
TR	0.825	0.209	0.803	0.825	0.814
Avg.	0.808	0.192	0.808	0.808	0.808
(4) Previous features + New features + "power of spread" factor					
FR	0.802	0.145	0.802	0.846	0.824
TR	0.855	0.198	0.855	0.812	0.833
Avg.	0.828	0.172	0.828	0.829	**0.828**

5.3.1 Application of spread power in identifying rumors

The study also aims to demonstrate the application of the spread power of the text as a feature to identify rumors. Also, the impact of new features that are used to calculate the power of release is shown in the identification of rumors. These evaluations have been performed on the Telegram dataset. Four types of experiments are carried out to assess the effect of the spread power score and new proposed features in the process of rumors identification. In the first experiment, the classification of two classes of FR and TR is performed based on all the previous features introduced by other researchers that are listed in Table 1.
In the second experiment, only a set of new features was used to identify rumors, so that the effectiveness of these features in distinguishing FRs from TRs is determined. In the third experiment, the "power of spread" factor is added to the new proposed feature set, and the process of rumors classification is done using the RF classifier. Finally, in the fourth experiment, a combination of three feature sets (previous features + new proposed features + power of spread) is used to identify rumors.

Table 8 demonstrates the result of the evaluated metrics of Precision, Recall, and F-measure to evaluate different features and their impact on validating rumors. As shown in Table 8, the set of new proposed feature more ability to categorize rumors than previous features. Also, SPR factor as a new feature has been instrumental in categorizing rumors. On the other hand, the classification accuracy using the combination of these previous and new features is improved from 0.762 to 0.828 using RF classifier. This result demonstrates the positive impact of new proposed features and "power of spread" factor on improving the classification accuracy.

As the results of the Table 8 shows, the spread power criterion has a significant effect on the process of categorizing two classes of FR and TR. The result of this analysis was an average F-measure of 0.828. In this study, the average F-measure is improved from 0.762 (first experiment) to 0.828 (fourth experiment) by focusing only on content features.

Table 9: Comparison of the proposed method with similar methods in Persian language.

	Dataset	Results			Features		
	Social Media	Precision	Recal	F-measure	User	Content	Structural
Zamani et al. 2017	Twitter	0.819	0.819	0.817	No	Yes	Yes
Proposed methode	Telegram	0.828	0.829	**0.828**	No	Yes	No
	Twitter	0.763	0.763	0.763	No	Yes	No

5.3.2 Comparison

The proposed method is compared to the only published work on the Persian language by Zamani et al. [12] on the same dataset of rumor tweets. The average F-measure of our model to recognize Twitter rumors is 0.763. The result of this assessment was satisfactory, because RSPMM is only based on textual information of rumors, and no other information on the user level and the propagation network is used. But, Zamani et al. [12] addressed rumor detection on Persian Twitter by analyzing two categories of rumor features: Structural and Content-based features. Their experiments yielded about 70% precision just based on structural (user graph) features and more than 80% based on both categories of features. However, the RSPMM in this study achieved a satisfactory result only by focusing on textual features. Table 9 shows the results of comparison of the proposed RSPMM with only published work by Zamani et al. [12].

6 Conclusion

The use of strong, emotional and affective expressions in the content of a text has a significant impact on the speed of its publication. This rule is also used to spread FRs quickly so that FRs are published with higher power and speed than TRs. Determining the power of spread of information available on online media is an unaddressed and new task in the field of rumors analysis. The study proposed a Rumor Spread Power Measurement Model (RSPMM) to compute SPR in Social Media. The problem of SPR has been proposed only theoretically by Allport and Postman. This study considered to be the first work to formulate the SPR based on contextual features of rumors. The study addressed this problem on Persian Telegram posts and tweets of Twitter. First, a dataset of Persian FRs and TRs is collected and labelled. Next, based on the analysis of different categories of contextual features, a set of textual features is introduced that imply the importance and ambiguity of a text. In order to better understand the impact of proposed features on calculating SPR, data analysis such as StudentâĂŹs t-test on these features is performed.

Four types of experiments have been performed to illustrate (1) the performance of previous features, (2) the performance of the proposed new features on the distinction between FRs and TRs, (3) the impact of the power score as an effective feature in the rumors classification, and (4) the combination of features used in three experiments and the use of these features to verify rumors. The results of these experiments were the answer to our research questions. The response to the main research question is that SPR score of a text can be calculated based on contextual features that indicate the importance and ambiguity of the text. Moreover, the response to the question "Is there any difference between the spread power of FRs and TRs?" the answer is yes. The null hypothesis is defined as "the spread power of FRs is equal to the TRs". Since, $p - value = 0.002 \leq 0.05$, so the null hypothesis is rejected for the spread power. Also, since the average of the spread power in FR is more than TR, it can be concluded that the spread power of FRs is usually more than TRs.

Based on Allport and Postman's theory, two factors of importance and ambiguity as effective factors in SPR are introduced. Thereby, two categories of features are introduced to calculate the importance and ambiguity of the text, where the importance factor is the combination of emotional and newsworthy and to compute the ambiguity of the text, a set of textual-based features are extracted that cause ambiguity in the content of the text.

Finally, as to whether SPR factor can be used to verify rumors, the results shown that SPR as a distinctive feature between FRs and TRs improves the accuracy of rumor verification from 0.791 (2nd experiment) to 0.808 (3rd experiment).

Therefore, the study demonstrated that the application and the positive impact of the SPR score is an effective criterion to verify rumors.

7 Acknowledgements

This project is supported by a research grant of the University of Tabriz (number S/806).

References

[1] Nicholas DiFonzo and Prashant Bordia. *Rumor psychology: Social and organizational approaches.* American Psychological Association, Washington, 2007.

[2] GORDON W. Allport and Leo Postman. *The psychology of rumor*, volume 257. John Wiley & Sons, Ltd, oct 1947.

[3] Jayson Harsin. The rumour bomb: Theorising the convergence of new and old trends in mediated us politics. *Southern Review: Communication, Politics & Culture*, 39(1):84–110, 2006.

[4] K P Krishna Kumar and G Geethakumari. Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences*, 4(1):14, dec 2014.

[5] Xichen Zhang and Ali A. Ghorbani. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, page 102025, 2019.

[6] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web - WWW '11*, page 675, New York, New York, USA, 2011. ACM Press.

[7] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. Prominent features of rumor propagation in online social media. In *13th IEEE International Conference on Data Mining (ICDM'2013), Dallas, Texas, U.S.A.* IEEE, dec 2013.

[8] Vahed Qazvinian, Emily Rosengren, Dragomir R. Radev, and Qiaozhu Mei. Rumor has it: identifying misinformation in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1589–1599. Association for Computational Linguistics, 2011.

[9] Fan Yang, Yang Liu, Xiaohui Yu, and Min Yang. Automatic detection of rumor on sina weibo. In *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics - MDS '12*, pages 1–7, New York, New York, USA, 2012. ACM Press.

[10] Ke Wu, Song Yang, and Kenny Q. Zhu. False rumors detection on Sina Weibo by propagation structures. In *2015 IEEE 31st International Conference on Data Engineering*, pages 651–662. IEEE, apr 2015.

[11] Soroush Vosoughi. *Automatic detection and verification of rumors on Twitter*. PhD thesis, Massachusetts Institute of Technology, 2015.

[12] Somayeh Zamani, Masoud Asadpour, and Dara Moazzami. Rumor detection for persian tweets. In *2017 Iranian Conference on Electrical Engineering (ICEE)*, pages 1532–1536. IEEE, may 2017.

[13] Sarah A. Alkhodair, Steven H.H. Ding, Benjamin C.M. Fung, and Junqiang Liu. Detecting breaking news rumors of emerging topics in social media. *Information Processing & Management*, page 102018, 2019.

[14] Yahui Liu, Xiaolong Jin, and Huawei Shen. Towards early identification of online rumors based on long short-term memory networks. *Information Processing & Management*, 56(4):1457 – 1467, 2019.

[15] F. Xing and C. Guo. Mining semantic information in rumor detection via a deep visual perception based recurrent neural networks. In *2019 IEEE International Congress on Big Data (BigDataCongress)*, pages 17–23, July 2019.

[16] Zhe Zhao, Paul Resnick, and Qiaozhu Mei. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, pages 1395–1405, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee.

[17] Arkaitz Zubiaga, Maria Liakata, and Rob Procter. Exploiting context for rumour detection in social media. In *International Conference on Social Informatics*, pages 109–123. Springer, 2017.

[18] S. Wang and T. Terano. Detecting rumor patterns in streaming social media. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2709–2715, Oct 2015.

[19] Sejeong Kwon, Meeyoung Cha, and Kyomin Jung. Rumor detection over varying time windows. *PLOS ONE*, 12(1):1–19, 01 2017.

[20] Ahmad Yahya M. Floos. Arabic rumours identification by measuring the credibility of arabic tweet content. *International Journal of Knowledge Society Research*, 7(2):72âĂŞ83, 2016.

[21] Y. Liu and S. Xu. Detecting rumors through modeling information propagation networks in a social media environment. *IEEE Transactions on Computational Social Systems*, 3(2):46–62, June 2016.

[22] T. Takahashi and N. Igata. Rumor detection on twitter. In *The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, pages 452–457, Nov 2012.

[23] Cheng Chang, Yihong Zhang, Claudia Szabo, and Quan Z. Sheng. Extreme user and political rumor detection on twitter. *Advanced Data Mining and Applications Lecture Notes in Computer Science*, page 751âĂŞ763, 2016.

[24] W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee. Behavior deviation: An anomaly detection view of rumor preemption. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 1–7, Oct 2016.

[25] S. Jain, V. Sharma, and R. Kaushal. Towards automated real-time detection of misinformation on twitter. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2015–2020, Sep. 2016.

[26] Li Zeng, Kate Starbird, and Emma S. Spiro. Rumors at the speed of light? modeling the rate of rumor transmission during crisis. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 1969–1978. IEEE, jan 2016.

[27] Benjamin Doer, Mahmoud Fouz, and Tobias Friedrich. Why rumors spread so quickly in social networks. *Communications of the ACM*, 55(6):70, jun 2012.

[28] Lina Zhou, Judee K. Burgoon, Jay F. Nunamaker, and Doug Twitchell. Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation*, 13(1):81–106, jan 2004.

[29] Sardar Hamidian and Mona T Diab. Rumor detection and classification for twitter data, 2019.

[30] Saif M. Mohammad and Peter D. Turney. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465, aug 2013.

[31] Arsalan Golfam, Azita Afrashi, and Ghazale Moghadam. Conceptualization of the persian simple verbs of motion: a cognitive approach. *Journal of Language and Western Iranian Dialects*, 1(3):103–122, jan 2014.

[32] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. Aspects of Rumor Spreading on a Microblog Network. In *International Conference on Social Informatics*, pages 299–308. Springer, Cham, nov 2013.

[33] Cass R. Sunstein. *On rumors : how falsehoods spread, why we believe them, and what can be done*. Princeton University Press, 2014.

[34] Wenbin Zhang and Steven Skiena. Trading Strategies to Exploit Blog and News Sentiment. In *Fourth International Conference on Weblogs and Social Media, ICWSM 2010*, Washington, DC, USA, 2010.

[35] Zoleikha Jahanbakhsh-Nagadeh, Mohammad-Reza Feizi-Derakhshi, and Arash Sharifi. A speech act classifier for persian texts and its application in identify speech act of rumors. jan 2019.

[36] Zili Zhang, Ziqiong Zhang, and Hengyun Li. Predictors of the authenticity of internet health rumours. *Health Information & Libraries Journal*, 32(3):195–205, sep 2015.

[37] Hamed Moradi, Farid Ahmadi, and Mohammad-Reza Feizi-Derakhshi. A hybrid approach for persian named entity recognition. *Iranian Journal of Science and Technology, Transactions A: Science*, 41(1):215–222, mar 2017.

[38] Hamid Mohammadi and Seyed Hossein Khasteh. A machine learning approach to persian text readability assessment using a crowdsourced dataset. oct 2018.

[39] Nwokocha K. U. Nkpa. Rumor mongering in war time. *The Journal of Social Psychology*, 96(1):27–35, jun 1975.

[40] James Kennedy. Particle swarm optimization. In *Encyclopedia of Machine Learning*, pages 760–766. Springer US, Boston, MA, 2011.

[41] Ali-Reza Feizi-Derakhshi, Mohammad-Reza Feizi-Derakhshi, Mehrdad Ranjbar-Khadivi, Narjes NikzadâĂŞKhasmakhi, Majid Ramezani, Taymaz Rahkar Farshi, Elnaz Zafarani-Moattar, Meysam Asgari-Chenaghlu, and Zoleikha Jahanbakhsh-Nagadeh. Sepehr_RumTel01. jan 2019.

[42] Manizheh Ghaemi and Mohammad-Reza Feizi-Derakhshi. Forest optimization algorithm. *Expert Systems with Applications*, 41(15):6676–6687, nov 2014.