TOPIC MODELLING: AN ANALYSIS OF SOCIAL ISSUE THROUGH TWITTER

Elavarasi D[1]
AP/CSE, Mount Zion College of Engineering and Technology, Pudukkottai

***Abstract*—The prime focus is on "topic modeling" by extracting the tweets from Twitter about Sterlite Project. The extracted tweets are processed using LDA (Latent Dirichlet Allocation) and Gibbs Method. Using these methods we analyze locus standi on Sterlite Project from various dimensions. The centroid of the issue is the central government stand on extraction of copper by hydraulic without considering other environmental issues that hampers the livelihood of the inhabitants. Which in turn favors the corporate profit and by destroying nature. There was public outcry from various groups in Social Media (Twitter) against the Sterlite Project. The topic modeling clearly explores all opinions and options from various dimensions about Sterlite Project. In our study there were few topics that are been discussed in a large scale such as 'BanSterlite','SaveTuticorin','support','tamilnadu' and 'Sterlite protest' by which we can say that the sterlite project is a threat to the agriculture and destroys the sentiment of the people associated with farming. It is also observed that the emotions are changing with the respect to time and reactive policy changes from the government. A time series analysis with the topic modeling is carried out and studied for the better understanding of the sentiments.**

***Index Terms*—LDA, Gibbs Method, Sterlite Project, Topic Modeling, Social Media**

I. INTRODUCTION

Social media plays a vital role in every user around the world, such as sharing our pictures and current mood as post and tweets we use social media like "Facebook" and "Twitter" and more often we interact with our friends through messages in these media. Nowadays social media has gone a long way like education, shopping and find our life partner too.[1] The initial reason to create a platform for such social media is to improve communication and get us connected to people who are not near us. Many of us always depend on these social media to gather information, learn stuff that we often don't find in the books and know the culture and taboo of various places around the globe.

Bringing the media accessible in our devices such as our smart phones make us more attached to these devices; instead of spending our time with the real time people who is around us. There is also complaint that students are not spending enough time in studying or performing practical work rather they spend more time in social media and the devices related to it. There are couples who developed relationship and were able to identify their life partners through social media; at the same time there are fake accounts to distract youngster who are not matured enough, misguiding them to get themselves trapped as a victim.

Other than all the criteria given above social media is basically to connect people, and the sites are created in a user friendly way to break our personal fence(privacy) and fly high to see others in the world beyond the sky. More time we are able to identify our own character about what is our role in the drama and find the perfect character that we want our self to perform the way we want and not the way others expect us to do. Before two decades an individual was not able to raise voice for his own rights, he was not identified, he was not heard, he had no means to convey his necessity and he was hidden from the individuals around him. Nevertheless now raising our voice for our rights, [2] for social problem, and to have our cultural heritage and even to save our environment the social media is helpful.

These reverberations are not only to be heard by the government and its official instead it is now carefully attended by each and every individual in the social media.

With such voices there had been protest that is been organized for some sort of problems without a strong basement and got success with the help of social media. Such as the cases in "Iran '09", "Egypt '11" and "2017 pro-jallikattu protests" [3]. Those who have accounts in any of the social medias are having all rights to create, share and comment any article into the media spreading them around the world. Often government is using social media to spread awareness about the endangering species and many others send facts that is useful for daily life or facts that is been forgotten in our daily life. Even though there is a fun side in sharing all these stuff there is also the other side where the false facts is being published and again creating confused story behind us. It is mandatory for us to identify the truth behind the scene about the article that is been shared, and to find the loyalty and the truth about it.

There are streams that are been reviewed and judged with the help of social media, as there are always positive and negative views for a particular object. The movie review, political parties with more supports, organizes social activities and so on. When it comes to review there is always two faces, sometimes more than that; let us not worry the number of faces instead concentrate on positive and negative faces.

There are number of studies focusing on how a social media can be a tool in shaping social movements for both offline and online at a global level. Social media such as Facebook, Twitter, YouTube and the various online blogs have given their voices of support for many individuals otherwise they would have never been heard. The first biggest revolution started with the help of social media is known as 'Arab Spring'

which is commonly known as 'Egyptian Revolution', 'Facebook revolution' and Twitter revolution', [4] why does this movement has its name after Facebook and twitter, it is only because of the role of social media in this movement. There has been study on the movement of how a social media has played a major role and the changer it has bought to the society. There were blogs written and published in the internet against the Mubarak regime, Egyptian government failed to block or privacy the internet user which lead to share instant message through Facebook and the information to form a group in the Tahrir Square to protest against the government. Ultimately in the end they had succeeded their victory. [5]

In order to identify the related topics that are been analyzed using social media and R we use a statistical method called "Topic Modeling". It is used to group the topics and phrases that is been mainly used by the public and draw a conclusion from the response of the user and plot them as a graph. The tweets that are being extracted are completely analyzed and broken into separate words and categorize them into different set, by the maximum number of usage of these words.[13] The most used words are formed into word cloud with all the topic related to the subject that we are analyzing. To complete this statistical method we use some techniques such as LDA (Latent Dirichlet Allocation) and Gibbs method.

II. IMPACT ON SOCIAL MEDIA

1. Correlation of policy changes with respect to environmental issues.
2. Productivity or loss from farming land.
3. Corporate must think and act eco-friendly.
4. Sustainable policy changes towards non eco-friendly project.

.

Theoretical Background

Social media and its effects on individual and social system completely gives the idea of how the global communication is been improved and to make sure that even the civilians in the deserted area are able to express their views and get it exposed to the world. It mainly consider between virtual interaction of global communication and individual communication which has improved our connection with others by making a strong bonding with those who are away from us.

Social media has gone a long way from individual communication to global communication by expressing their views of good and bad of the public cry or an activity performed in the society. The role of each individual they act through social media has their impact to the society like usually the users are mainly fond in streaming news feed and the profile pictures and so on,[17] some focus on blogs and forum to develop their knowledge. Whereas now the number of individuals taking part in activism through social media has increased, in order to have a strong movement in the field of political science and social movements the effect will always depend upon the communication between the individuals and the decision maker.

There are number of studies focusing on how a social media can be a tool in shaping social movements for both offline and online at a global level. Social media such as Facebook, Twitter, YouTube and the various online blogs have given their voices of support for many individuals otherwise they would have never been heard. The first biggest revolution started with the help of social media is known as 'Arab Spring' which is commonly known as 'Egyptian Revolution', 'Facebook revolution' and Twitter revolution', why does this movement has its name after Facebook and twitter,[18] it is only because of the role of social media in this movement. There has been study on the movement of how a social media has played a major role and the changer it has bought to the society. There were blogs written and published in the internet against the Mubarak regime, Egyptian government failed to block or privacy the internet user which lead to share instant message through Facebook and the information to form a group in the Tahrir Square to protest against the government.

Ultimately in the end they had succeeded their victory.

A survey in data mining technique is a survey that deals with the different type of mining techniques done in social media for the past years and up to date. With the help of this survey we can understand the techniques that are been followed in the data mining done in social media,[21] it has a significance of how the data is being used in order to share our view critic an event or an individual. According to the graph theory the major components are nodes and links that is to get the followers and get the link through them. There may be different data mining techniques used in social network analysis which totally depends on the supervision that is been conducted on by retrieving information and contents of the data generated. The algorithm or a statistical method that is been used in the topic modeling is LDA (Latent Dirichlet Allocation) it is a generative statistical model that allows set of observation that is been defined by the group of unobserved and find the similarity between them. The major role is to find the set of description that are collected from the large collection of a document.LDA is fully constructed with the mathematical equation and statistic method, initially LDA was a graphical method for topic discovery that the major topics discussed are found and plotted in a graph. LDA is a collection of discrete data by a flexible generative probabilistic model. It is an easy method for identify the topics that are mainly discussed in a large scale

III. PROPOSED SYSTEM

Our current focus is on 'Topic modeling' on the current issue based on 'Sterlite Project' in Tuticorn Dt., Tamil Nadu, India. The tweets are been extracted for several interval period of time for my analysis. The high frequently used topics that are been discussed about the Hydro-Carbon Project are been collected in different period are identified and are been compared. The frequency is been set by the user using the DTM(Document Term Matrix) where the topics are tabulated in a matrix form where rows represent topics and columns represent the number of term.

There are streams that are been reviewed and judged with the help of social media, as there are always positive and negative views for a particular object. The movie review, political parties with more supports, organizes social activities and so on. When it comes to review there is always two faces, sometimes more than that; let us not worry the number of faces

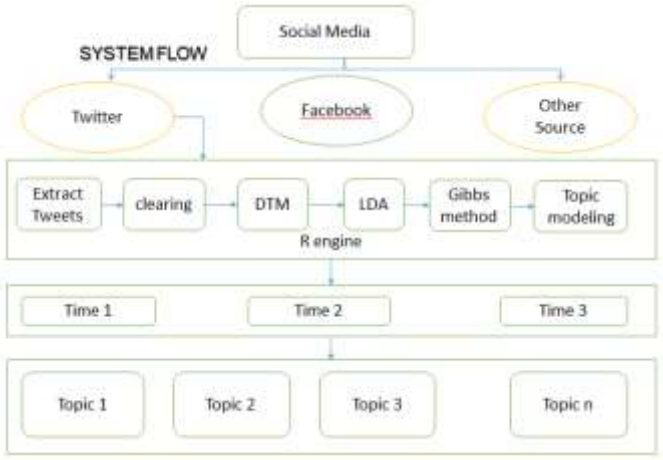
instead concentrate on positive and negative faces. In order to find such cases we must run an assessment to analysis and get an output on how the people has shared their views about the subject that we are going to analyze. In this paper with the help of an open source programming language 'R', the post and tweets are extracted from Facebook and twitter respectively, analyzing and representing them in graphical form and obtain an conclusion. There are number of tweets/post been extracted using R in a several period of time and find the mentality of the public in that time, plotting in graphical structure.

In order to identify the related topics that are been analyzed using social media and R we use a statistical method called "Topic Modeling". It is used to group the topics and phrases that is been mainly used by the public and draw a conclusion from the response of the user and plot them as a graph. The tweets that are being extracted are completely analyzed and broken into separate words and categorize them into different set, by the maximum number of usage of these words. The most used words are formed into word cloud with all the topic related to the subject that we are analyzing. To complete this statistical method we use some techniques such as LDA (Latent Dirichlet Allocation) and Gibbs method.

Since we look forward to a point that has clear definition of the subject that we are going to analyze, but we extract number of tweets and post which are collected randomly in the period of time. These post and tweets are identified with the help of "metadata tag" which is commonly known as "hash tag (#)"; post are often uses this symbol and type the subject that are being posted or tweeted to categorize and to make sure that they are referring to the particular subject in the posts. So when collect all these post we get a lot of posts referred to the subject and we use topic modeling in order find related and most used topics about that subject; here is where we use LDA to study these post and find how many percentage of this post is referred to the subject that we are searching about. Many times we share many subject into single post which may complicate the analyze by introducing foreign words into our subject, which may affect our topic modeling to avoid such issues we follow LDA method into topic modeling to make it simpler and easier.

The algorithm or a statistical method that is been used in the topic modeling is LDA (Latent Dirichlet Allocation) it is a generative statistical model that allows set of observation that is been defined by the group of unobserved and find the similarity between them. The major role is to find the set of description that are collected from the large collection of a document.LDA is fully constructed with the mathematical equation and statistic method, initially LDA was a graphical method for topic discovery that the major topics discussed are found and plotted in a graph. LDA is a collection of discrete data by a flexible generative probabilistic model. It is an easy method for identify the topics that are mainly discussed in a large scale.

IV. FRAMEWORK



The social media is a group of social sites like Twitter, Facebook and other such sources in which the post and views for all related topics are been shared and discussed. Initially my study is done in twitter, where the tweets related #SterliteProject are been posted and these tweets are extracted into the R engine with the access the twitter apps has provided to our account. After which the tweets are saved in R engine. The extracted tweets are then processed into cleaning tweets in which the conjunctions, punctuations, links, numbers and symbols are removed such as only the topics that are been discussed is obtained.

Since lots of topics are discussed some may be discussed for a small time and some for a longer period. In order to get the most discussed topics a frequency is been fixed and all the topics that fits into the frequency is represented as DTM. A document-term matrix or term-document matrix is a mathematical matrix that describes the frequency of terms that occur in a collection of documents. In a document-term matrix, rows correspond to documents in the collection and columns correspond to terms. There are various schemes for determining the value that each entry in the matrix should take. One such scheme istf-idf. They are useful in the field of natural language processing.

The "topics" produced by topic modeling techniques are clusters of similar words. A topic model captures this intuition in a mathematical framework, which allows examining a set of post and tweets, by discovering, based on the statistics of the words in each, what the topics might be and analyze the result by plotting them in graph.

The algorithm or a statistical method that is been used in the topic modeling is LDA (Latent Dirichlet Allocation) it is a generative statistical model that allows set of observation that is been defined by the group of unobserved and find the similarity between them. The major role is to find the set of description that are collected from the large collection of a document.LDA is fully constructed with the mathematical equation and statistic method, initially LDA was a graphical method for topic discovery that the major topics discussed are found and plotted in a graph. LDA is a collection of discrete data by a flexible generative probabilistic model. It is an easy method for identify the topics that are mainly discussed in a large scale.

With the use of LDA and Gibbs sampling the topic modeling is carried out by identifying the most used topics in

the extracted tweets in different sets and then the each set are compared with each other and the results are obtained.

A. *Phase-1: Data Extraction*

Data Extraction will be used to pull the tweets from R code using Oauth facility by Submitting Twitter credentials. Similarly, Facebook Posts will be pulled from R code using FB oauth facility by Submitting Facebook credentials. Corpus Cleaning will be used to clean present the extracted data to R engine.

Sparsity

Sparsity of data is the collection of words that occurs occasionally in the given set of documents. This means that sparse data is the collection of rare words from the set of documents.

Data Cleaning

The punctuations, conjunctions, links, pictures, numbers and symbols are cleared from all the extracted tweets from twitter. This helps in identifying the proper and meaningful topics that are been discussed.

Frequent Words

Frequent Words can be identified from the Term Document Matrix. Frequent Word and all its associations can be finding through this data sets. These frequent words and all of its associations are projected as a graph with two axes like the count and words. These are visualized in a graph with the words on the mentioned frequency.

B. *Term Document Matrix*

A term-document matrix is also known as document-term matrix which is also shortly called as TDM. In the collection of documents, document-term matrix is a mathematical matrix which describes the frequency of terms. The collection of rows corresponds to documents and columns correspond to terms in the document term matrix. To determine the value there are many schemes for to enter in each matrix. One such scheme istf-idf. They are useful in the field of natural language processing.

C. *Topic Modeling*

Topic modeling is used to group the topics and phrases that is been mainly used by the public and draw a conclusion from the response of the user and plot them as a graph. The tweets that are being extracted are completely analyzed and broken into separate words and categorize them into different pits, by the maximum number of usage of these words. The most used words are formed into word cloud with all the topic related to the subject that we are analyzing. To complete this statistical method we use some techniques such as LDA (Latent Dirichlet Allocation) and Gibbs method.

D. *Report and Generation*

In order to discover the abstract "Topics" in collection of documents, statistical model is used. A text-mining tool for discovery of hidden semantic structures in a text body by the topic modeling. Instinctively, given that a document is about a particular topic.

V. METHODOLOGY

Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) is the classical method to perform topic modeling. In order to do topic modeling LDA and Gibbs sampling is been used and it is a statistical method that create a set of observed topics from a large document. In natural language processing LDA is a generative statistical model that allows sets of observation to be explained by unobserved groups that explain why some parts of the data are similar. Now suppose we would like to predict the number of times each class will occur over a series of *N* trials. If we have two classes, we can model this with a binomial distribution, as we would normally do in a coin flip experiment. For *K* classes, the binomial distribution generalizes to the multinomial distribution, where the probability of each class, *pi*, is fixed and the sum of all instances of *pi* equals one. Now, suppose that we wanted to model the random selection of a particular multinomial distribution with *K* categories. The Dirichlet distribution achieves just that. LDA is constructed with substantial amount of mathematics, but it is important to understand the work of this model and how it is gathering the topics. Given that the mathematical equation to perform LDA is

$$(\bar{x}|a_k) = \prod (K\sum Kk=f1(aakk)) \; K \;\; xk \; ak-1$$

$$k=1 \; k=1$$

Where,
K is the count of the occurrence of the each topic
P is the probability of the occurrence that may occur
$\bar{x}$ is the vector with K component
x_k is a multinomial distribution
a is component which contains K parameters

This equation seems complex, but if we break it down to its constituent parts and label the symbols used, we will then be able to have a better understanding. To begin with, the *x* term is a vector with *K* components, *xk*, representing a particular multinomial distribution. The vector _ is also a *K* component vector containing the *K* parameters, *αk*, of the Dirichlet distribution. Thus, we are computing the probability of selecting a particular multinomial distribution given a particular parameter combination. Notice that we provide the Dirichlet distribution with a parameter vector whose length is the same as the number of classes of the multinomial distribution that it will return.

The fraction before the large product on the right-hand side of the equation is a normalizing constant, which depends only on the values of the Dirichlet parameters and is expressed in terms of the gamma function. For completeness, the gamma function, a generalization of the factorial function, is given by the following:

$$\infty$$

$$(t) = (x^{t-1}e^{-x)}$$
0

Binomial Distribution

The value of αk is changed periodically and the changes that are made in the final graph are plotted sequent and their behavior is being compared. Lastly, in the final product, we see that every parameter, αk, is paired with the corresponding component of the multinomial distribution, xk, in forming the terms of the product. The important point to remember about this distribution is that by modifying the αkparameters, we are modifying the probabilities of the different multinomial distributions that we can draw.

We are especially interested in the total sum of the αk parameters as well as the relative proportions among them. A large total for the αk parameters tends to produce a smoother distribution involving a mix of many topics and this distribution is more likely to follow the pattern of alpha parameters in their relative proportions.

A special case of the Dirichlet distribution is the Symmetrical Dirichlet distribution, in which all the αkparameters have an identical value. When the αk parameters are identical and large in value, we are likely to sample a multinomial distribution that is close to being uniform. Thus, the symmetrical Dirichlet distribution is used when we have no information about a preference over a particular topic distribution and we consider all topics to be equally likely.
Similarly, suppose we had a skewed vector of αk parameters with large absolute values. For example, we might have a vector in which one of the αk parameters was much higher than the others, indicating a preference for selecting one of the topics. If we used this as an input to the Dirchlet distribution, we would likely sample a multinomial distribution in which the aforementioned topic was very probable.

Gibbs Sampling Method

Gibbs sampling method is used after LDA to get a clear understanding between the topics that are been obtained through LDA method, by comparing the each set of topic with the rest of the sets the final output is obtained. With Gibbs sampling method the topics that are probably expected to predict the relation with all the topics that are been observed using LDA method previously. It is a randomized algorithm that predicts randomly which can be fixed by an approximate count.

Gibbs sampling is often suggested for such applications where it is well suited to coping with incomplete information. However, this makes computational cost for many applications Nevertheless, understanding of this Gibbs sampling method which provides valuable insights into the problems of statistical inference.

It is a statistical inference, especially Bayesian inference. It is also a randomized algorithm and is an another choice to deterministic algorithms which is also used for statistical inference such as the expectation-maximization algorithm.

$$(x_1, \dots, x_n)$$
$$(x_j/x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = \frac{}{(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)} \propto (x_1, \dots, x_n)$$

Normalization Function

"Proportional to" in this case means that the denominator is not a function of and thus is the same for all values of; it forms part of the normalization constant for the distribution over. In practice, to determine the nature of the conditional distribution of a factor, it is easiest to factor the joint distribution according to the individual conditional distributions defined by the graphical model over the variables, ignore all factors that are not functions of (all of which, together with the denominator above, constitute the normalization constant), and then reinstate the normalization constant at the end, as necessary. In practice, this means doing one of three things:

1. If the distribution is discrete, the individual probabilities of all possible values of are computed, and then summed to find the normalization constant.
2. If the distribution is continuous and of a known form, the normalization constant will also be known.
3. In other cases, the normalization constant can usually be ignored, as most sampling methods do not require it.

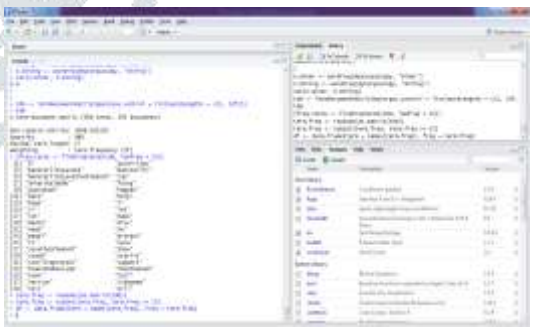


Fig. Topic Modeling

VI. CONCLUSION

In the end of our study with the help of the technique followed (Topic modeling) there are eight sets of topics that are mainly discussed in a large scale, the featured topics are ‘farmer’, ‘savefarm’,’support’,’tamilnadu’ and ‘student protest’. By analyzing these topics we can come to an conclusion that ‘Hydro Carbon Project’ must be reconsidered by the government for the welfare of the farming land and the sentiment of farmers, according to the tweets from twitter it is against the ‘Hydro Carbon Project’ in Neduvasal. So with my study result the government should reconsider the project that is been initiated in Neduvasal and make necessary changes in the policy for the welfare of the farmers and farming land in

Tamil Nadu. The sentiments are taken in the month of March 2018 which clearly portraits strong wave against government stand in this project.

VII. FURTHER WORK

The technique followed (Topic modeling) contain eight sets of topics that are mainly discussed in a large scale. By analyzing these topics we can come to a conclusion must be reconsidered by the government for the welfare. So we are using the time series analysis to study the variation of people mind set using the social media. Model topics without taking into account 'time' will confound the topic discovery. In the second category, paper has discussed the topic evolution models, considering time. Several papers have used different methods of model topic evolution. Some of them have used discretizing time, continuous-time model, or citation relationship as well as time discretization. Hence this is the future work of this project.

VIII. REFERENCE

[1] *NataschaZeitel-Bank. Social Media and Its Effects on Individuals and Social Systems in Management Center Innsbruck, Austria on June, 2014.*

[2] *Bart Cammaerts. Social Media and Activism in Mansell, R., Hwa, P., The International Encyclopedia of Digital Communication and Society. Oxford, UK: Wiley-Blackwell on 2015, pp. 1027-1034.*

[3] *Amandha Rohr Lopes. The Impact of Social Media on Social Movements: The New Opportunity and Mobilizing Structure in Creighton University on 2014.*

[4] *Caroline S. Sheedy. Social Media for Social Change: A Case Study of Social Media Use in the 2011 Egyptian Revolution, Presented to the Faculty of the School of Communication on April, 2011.*

[5] *Graham Cormode & S. Muthukrishnan. An Improved Data Stream Summary: The Count-Min Sketch and its Applications, submitted to Elsevier Science on December, 2003.*

[6] *Mariam Adedoyin-Olowe, Mohamed MedhatGaber& Frederic Stahl. A Survey of Data Mining Techniques for Social Network Analysis in World Wide Web.*

[7] *Gurmeet Singh Manku & Rajeev Motwani. Approximate Frequency Counts over Data Streams in Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002.*

[8] *David M. Blei, Andrew Y. Ng & Michael I. Jordan. Latent Dirichlet Allocation in Journal of Machine Learning Research 3 (2003) 993-1022.*

[9] *ZeynepTufekci& Christopher Wilson. Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square in Journal of Communication ISSN 0021-9916 on 2012.*

[10] *C. Wang, D. Blei & D. Heckerman. Continuous Time Dynamic Topic Models in Uncertainty in Artificial Intelligence, Helsinki, Finland, on July 2008.*

[11] *Sebastia´n Valenzuela, Arturo Arriagada& Andre´ s Scherman. The Social Media Basis of Youth Protest Behavior: The Case of Chile in Journal of Communication ISSN 0021-9916 on 2012.*

[12] *Tim Markham. Social Media, Protest Cultures and Political Subjectivities of the Arab Spring in Media, Culture & Society 36(1): 89-104 on 2014*

[13] *Takeshi Sakaki, Makoto Okazaki & Yutaka Matsuo. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors in World Wide Web on April,2010*

[14] *Lei Shi, NeerajAgarwal, AnkurAgrawal, Rahul Garg& Jacob Spoelstra. Predicting US Primary Elections with Twitter in World Wide Web.*

[15] *Rui Miguel Forte. Mastering Predictive Analytics with R, Pages 317-323.*

[16] *J. Allan, R. Papka, and V. Lavrenko. On-line new event detection and tracking. In SIGIR, pages 37–45, 1998.*

[17] *L. AlSumait, D. Barbar´a, and C. Domeniconi. On-line lda: adaptive topic modelsforminingtextstreamswithapplicationstotopic detectionandtracking. In ICDM, 2008.*

[18] *T. Brants, F. Chen, and A. Farahat. A system for new event detection. In SIGIR, pages 330–337, 2003.*

[19] *K. R. Canini, L. Shi, and T. L. Griffiths. Online inference of topics with latent dirichlet allocation. In Proceedings of the International Conference on Artificial Intelligence and Statistics, volume 5, pages 65–72, 2009.*

[20] *G. Cormode and S. Muthukrishnan. What's hot and what's not: tracking most frequent items dynamically. In Proceedings of the twenty-second ACM SIGMODSIGACT-SIGART symposium on Principles of database systems, pages 296–306, 2003.*

[21] *Q. Diao, J. Jiang, F. Zhu, and E.-P. Lim. Finding bursty topics from microblogs. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1, pages 536–544, 2012.*

[22] *T. Griffiths and M. Steyvers. Finding scientific topics. Proceedings of the National Academy of Sciences of the United States of America, 101(Suppl 1):5228–5235, 2004.*

[23] *M. D. Hoffman, D. M. Blei, and F. Bach. Online learning for latent dirichlet allocation. Advances in Neural Information Processing Systems, 23:856–864, 2010.*