

Finding Real Time Contextual Tweets on Twitter

Punam Sanjay Sawale¹, Prof. Vijayshri A. Injamuri²

¹M.Tech Final Year Student, Department of Computer Science and Engineering, Government College of Engineering, Aurangabad, Maharashtra, India

²Associate Professor, Department of Computer Science and Engineering, Government College of Engineering, Aurangabad, Maharashtra, India

Abstract - Twitter microblogging system is a pivotal part of web 2.0 eras and has become the momentous in almost everyone's life. Twitter streams information pertaining to emerging real world events. Over 500 million tweets uploaded on twitter per day and 200 billion tweets per year posted on twitter by twitter users. Tweets are restricted to 280 characters; because of this tweets are become ambiguous by nature. Due to this and fast pace at which tweets are generated, it is often difficult for users come across to the most apropos ambient tweet. We present a tweet search system "Finding Real Time Contextual Tweets on Twitter" to filtered the tweets on dynamic data streams. Our technique is to use RESTful rules and proper filtered stream endpoint which enables us to filter the real time stream of public tweets. It also allows users to listen for specific topics and events in real time.

Key Words: Microblogging, RESTful, Web 2.0, Bearer Token, Expansion, Fields.

1. INTRODUCTION

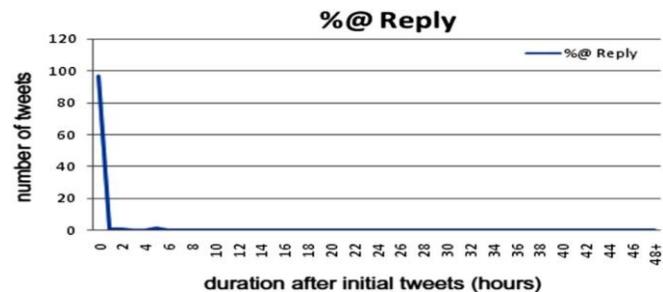
Microblogging is an amalgamation of conventional blogging and instant short messaging to be posted and shared with an online audience globally or with a specific community. This concept has been around since the web 2.0 eras that is social networking gets more popular. Microblogging is a communication medium which is differ from traditional blogs and bounded by 140 characters only, hence its main benefits is too spent less time developing contents.

Twitter is the most popular example of microblogs and has become the world wide phenomenon. It is mobile conventional platform that provide an easy way to share news, time sensitive information, announcements, opinions, real life events etc. Twitter is the more direct medium to communicate with the people globally. Twitter is the pathfinder for the concept of 'Followers', Tweets are more valuable and carry more social currency if more people see them. This encourages people to create large followings by inviting family, friends, colleagues and acquaintances. They turn join the platform and do the same. This invite follower loop has catalyzed explosive growth of twitter.

Twitter microblog service used to find out real time events as and when they are happening. Twitter is an important medium to track the real-world event as and when it is happening. This got the researchers thinking

about making a general-purpose system to identify all important events using Twitter messages [3].

Graph 1 demonstrate the growth of the Twitter conversations, y axis is the number of tweets and xaxis is the relative temporal distance from the original tweets, measured by hours. Given that Twitter is a real-time service, overall, about 97.87% of replies are generated



Graph 1: Number of tweets over an hour

within the first hour, while an additional 0.98% of replies happen in the second hour, which shows that Twitter can propagate information quite fast and a meaningful context tree can be formed very quickly [2]. This paper introduces simple and easy approach that extract basic context information from unstructured tweets in real time by providing the search query.

1.1 Existing System

The existing system uses unsupervised machine learning techniques that extract complete identical summary for the tweets and naturally identifies multiple facets of the live data stream in a scalable and effective manner [1] [2]. The work is proposed by the author named Manoj K Agrawal [1]. In [6], Author Shou et al. present a method to sum up a twitter data stream, filtered in the context of a given user query. In [9], authors demonstrates a method to summarize a given event topic. Their methodology is applicable for structured and recurring events such as sports events and need the prior knowledge of similar events.

1.2 Proposed System

In proposed system, we used node Js. Node Js is a JavaScript runtime; it's neither a framework nor a library or anything like that, the language is only JavaScript. Inside of JavaScript

running in the browser environment, it actually running on your machine as server runtime. It does by V8 JavaScript engine. Node Js is best for anything that is not CPU intensive.

A.REST API

REST API is an architectonic figure for creating the web services. A RESTful Service is one which implements that figure. It swivels around the resource where resource is accessed by a common interface using the HTTP protocol.

RESTful rules we used in our system are get, set, and delete which defines a URI service that provide resource representation JSON and set of HTTP methods.

We make a request to HTTP server which then issues the request to the twitter's API. After receives request from server, twitter issues API response back to the server and server turn the data into view for user.

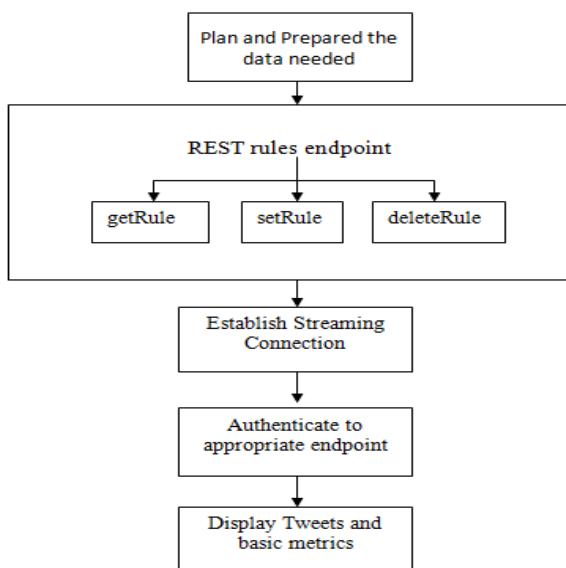


Fig -1: Architecture of Finding Tweets in Real Time

B.STREAMING CONNECTION

To unfold the twitter live data stream, to deliver the tweets according to our specific search query (Keyword for instance, 'MPSC'). We create a streaming connection and send it to the twitter streaming API. After receiving the request, twitter accept the request and when tweet streamed as they occur related to search query, twitter send them to streaming connection process. Streaming connection received streamed tweets, process them and store it as a result. This result forwarded to HTTP server which then

pulls the tweets from data store and rendered it into view for user.

C.AUTHENTICATION

Twitter handle mountainous amount of data (6,000 tweets in each moment). Authentication is the way to ensure that this data is secure.

1. OAuth2 Bearer token

OAuth2 bearer token is one of the authentication methods that twitter provides. This method allows application to access information publically behalf of twitter account. For this we must have a bearer token which is nothing but the array of strings, integers and special characters that can be generated by providing consumer key and secrete key through the POST OAuth2/token endpoint.

Sample bearer token:

"AAAAAAAAAAAAAAABBBBMLheAAAAAA0%2BuSeid%2BULvsea4JtiGRiSDSJSI%3DEUifiRBkKG5E2XzMDjRfl7ZC9Ub0wnz4XsNiRVBChTYbjcE3F"

How it works?

After successfully registering on twitter API, it generates consumer tokens and secret keys which directs user to verify credentials. After verification, application requests for basic fields and public metrics of tweets that contains username, name, id, text and retweet count using the bearer token. Twitter then responds with requested information to the authenticate application.

D.METHODOLOGY

We used twitter real time streaming API so that we can gets tweets in real time and apply RESTful rules which are basically like queries with the aim that we get tweets related to our specific search query. We create a server and used package called needle to make request with twitter API.

1.WORKING

In the first place, we must have a Twitter developer account and activated developer portal to access the twitter API and proper endpoints. We create a Node Js web server by using Express framework and used package needle to make request with twitter API. To set a search keyword we generate an array to hold the rules we add, it is basically an object with a value for instance 'MPSC' because of this we will have tweets which contains that keyword only.

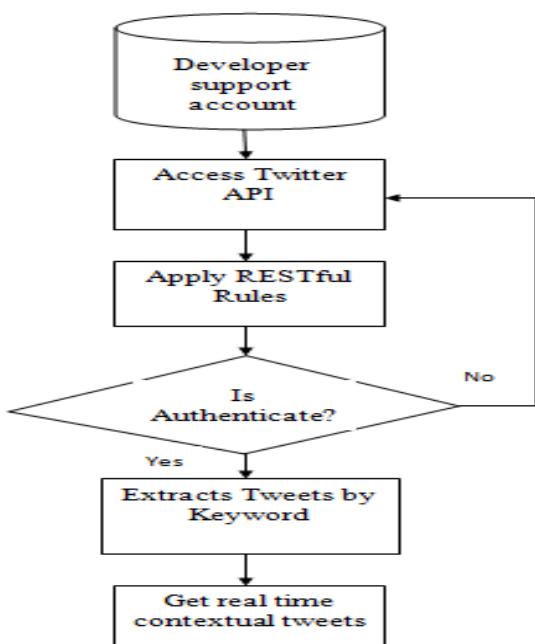


Fig -2: Flow Chart of the system

We generate three different functions `getRules()`, `setRules()` and `deleteRules()` for RESTful rules. In `getRules()`, we make a request to get the response. It is an `async()` function so that we could make request with needle and it will go to return a promise.

In `setRules()` function, we used the POST request to set the rules. We add data after the endpoint since we sends the data in header with 'Content-type = Application/JSON'. We have to delete rules before adding the new rules, for this we create `deleteRules()` function.

For streaming the filtered tweets from twitter we used function `streamTweets()` in which we make a request to stream URL.

2. RESULT

Our twitter search system returns the lists of tweets related to the keyword with the option 'Tweet', by clicking on it we can directly jump on the tweet of twitter home page.

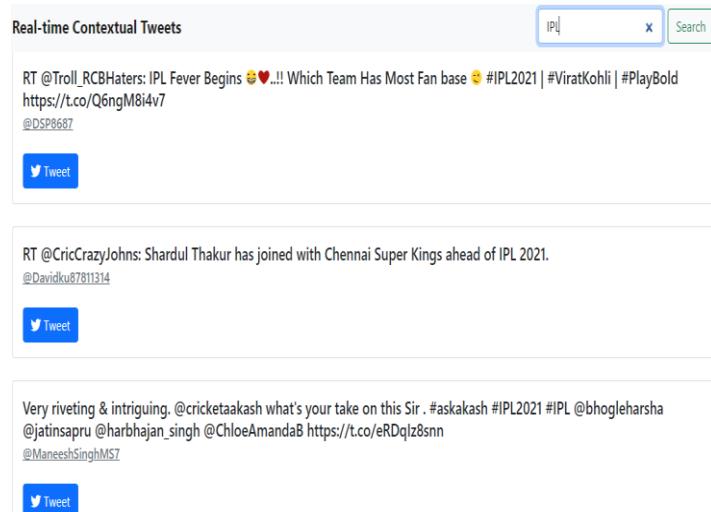


Fig -3: Screenshot of Result

Our system demonstrates:

1. We discovered the most ambient tweets over the live data stream for a specific query.
2. It also displayed the basic metrics, fields and expansions about the tweet.
3. We can jump to the user profile.

In future, will present the demographical summary of the live tweets related to search query together with number of related tweets posted for that specific keyword. We will also present that the tweet summary created by our system is complete.

3. CONCLUSION

Twitter is prime approach to trace the real world events when they appear. Innumerable people post tweets on twitter, making this treasure of information. We present the real time tweet search system "Finding real Time Contextual Tweets on Twitter" on vital tweets streams, to extract specific keyword related tweets from live events in real time on twitter. It extracts basic fields and metrics of the tweet and also displays the user profile who post that tweet.

REFERENCES

- [1] M. K Agarwal, K. Ramamritham, M. Bhide "Identifying Real World Events in Highly Dynamic Environments", in VLDB 2012.
- [2] M. Mathioudakis, N. Koudas, "Twitter Monitor Trend", in SIGMOD 2010.
- [3] Ramamritham, M. Bhide "Finding Contextual Messages in Real Time" 2019
- [4] Rim Faiz "Conversational based method for tweet contextualization" Vietnam J Computer sci,2017
- [5] Patrick Jacob ,Alexandra L. Uitdenbogerd "Readability of Twitter Tweets for Second Language Learners"

2016.

- [6] ENZA ANTENOS-CONFORTI "Microblogging on Twitter: social Networking in Intermediate Italian Classes" Montclair State University
- [7] L. Shou, Z. Wang, K. Chen, G. Chen, "Sumblr: Continuous Summarization of Evolving Tweet Streams", in SIGIR 2013.
- [8] Fernando Rosell-Aguilar "Twitter as languagelearning tool: from potential to evidence"2017
- [9] D. Chakrabarti, K. Punera, "Event Summarization using Tweets", in. ICSWM 2011