Volume 5, Issue 3, July 2018 ISSN: 2348-8565 (Online)
International Journal of Modern Engineering and Research Technology
Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

Real-Time Monitor RTMS System over Text Streams

D. Bharath Kumar
PG Research Scholar
Sri Indu College of Engineering & Technology
Ibrahimptnam (T.S.) [INDIA]
Email: bharathkumar9224@gmail.com

K. Shiva Krishna
PG Research Scholar
Sri Indu College of Engineering & Technology
Ibrahimptnam (T.S.) [INDIA]
Email: shivakirshna924@gmail.com

K. Manirathnam
PG Research Scholar
Sri Indu College of Engineering & Technology
Ibrahimptnam (T.S.) [INDIA]
Email: kurumanani1234@gmail.com

Ch. GVN Prasad
HOD & Professor
Department of Computer Science & Engineering
Sri Indu College of Engineering & Technology
Ibrahimptnam (T.S.) [INDIA]
Email: prasadch204@gmail.com

Abstract

Internet web platforms have been extremely popular in the big data era due to its real-time diffusion of information. It's important to know what data streaming events are trending on the social network and be able to monitor their evolution and find related data stream. In this paper we demonstrate RTMS real time monitor RTMS system over internet web text streams. RTMS integrates our efforts on both emerging data streaming monitor RTMS research and system research. From the data streaming monitor RTMS perspective, RTMS proposes a data stream analytic approach such that it able to detect earlier methods and discover correlations in monitoring the system with efficiency. The entire text stream with linear and horizontal scalability is proposed clearly and advantages over existing system and application monitoring in periodic times.

Keywords: *— RTMS, Big data, DataStream*

I. Introduction

The short and noisy nature of internet web text stream makes traditional topic detection methods and their derivatives inappropriate for the task. Latent space topic models are unsuitable for the fast changing online event monitor RTMS scenario. The model takes a long time to train and the topics are fixed. While the online event monitor RTMS scenario demonstrate fast changing set of topics and require high throughput of data processing. The value and application of real time monitor RTMS system over such platforms are many-fold. For instance, it can detect and respond to emergency events in a timely manner, track event evolution in an orderly fashion.

A wide variety of data streaming events would emerge from such platforms, ranging from political or daily affairs to natural disasters or public security menace. These platforms have many times been the first reporter of significant events, such as earthquakes and accidents, or even the major hosting venue of significant events, such as a presidential election campaign. A data stream space topic model is an well unsuitable for the fast changing online event monitor RTMS scenario. The model takes a long time to train and the topics are fixed. While the online event monitor RTMS scenario demonstrate fast changing set of topics and require high throughput of data processing.

The method would also tend to generate data streaming events that only contain a single keyword as description, which is hard to comprehend for users. None of the above methods provide the horizontal scalability with distributed implementations of their algorithms, nor do they investigate system optimizations for their applications.

Despite the challenges, there are many desired features when building an emerging data streaming monitor RTMS system over short text stream. From an emerging data streaming monitor RTMS perspective, we require the following steps:

1. The detection of emerging data streaming for all data stream.
2. High view of each sub-event as data streaming sub division.
3. Calculating the efficiency in processing system and computing the workload of each data stream detection and monitor RTMS.

II. Implementation

In the RTMS system, emerging data streaming monitor RTMS includes early detection, correlation analysis and temporal evolution tracking of data streaming events. Early detection would capture emerging events before they go viral. Correlation analysis would automatically reveal multiple aspects of the data streaming event, or the causality of data streaming events, or categorical structure of related data stream.

Data streaming events would emerge from different time granularities, e.g., a publicly concerned long trial or an overnight pop concert. The temporal evolution tracking of events could recover the evolution process of an data streaming event, to trace its origin and get the big picture. Such monitor RTMS happens in real-time and provides valuable intelligence for government agencies, news groups and marketing agencies, etc to process big data, the design of RTMS data stream model provides the following efficiency optimization opportunities. RTMS has a distributed data stream processing engine specifically optimized for our data streaming detection method. Algorithms are implemented to have linear horizontal scalability to handle big data, i.e., full stream of Web or Twitter data.

We designed a data stream stream model for data streaming event detection over the short and noisy text in micro blog services. To represent texts with a data stream stream model, we consider keywords as nodes and their co-occurrence relationships in each tweet as edges. For each incoming tweet, we generate a binary clique data stream over its keywords and retrieve the edge set.

Detecting the data streaming method has following steps:

Trending keyword detection and community detection over keywords. Our intuition is that the essential keywords about an event would have similar trends and show a burst in usage compared to their own history.

Trending keywords are detected as a data streaming with abrupt usage increase aspects of events, causality of events or categorical structure of events could be revealed through the correlation analysis.

Proposed method is naturally used to noisy keywords of more features, such as words about trivial things of mood, food etc., since these usage are relatively stable in the big picture. We note that the more the data, the better the chance to acquire accurate keywords with similar trends, the more meaningful the results would be. Keywords of the same event would occur and be more evently linked some internal with keywords from a different event. Following are the steps considered further to follow modularity based or between centrality based community detection methods for event detection.

The following are the drawback:

1. Density changes in parts of the data stream would affect the overall result for community detection.
2. A keyword cannot appear in different events
3. Keywords in different contexts would likely to be assigned to the same event due to lack of explicit definition of keyword community.
4. Hierarchical correlation analysis is absent for such methods.
5. Events would naturally consist of different aspects which formulate a hierarchical structure.
6. We would introduce our method for event detection and hierarchical sub-event correlation analysis.

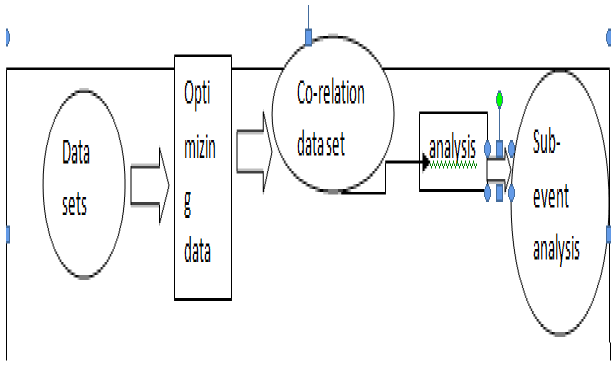

Figure 1: Representation of data streaming

Algorithm:

Data stream Construction Require and its maximal cliques, parameter to ensure

1. Each branches of data stream in to sub events
2. Put data stream in each in descending order according to their data set
3. Build a frequent pattern with each descending values of set
4. Check under each set of data stream the collection of event occurs in form of required matrix.
5. If the number of occurring set in pair of data stream is even then share the data set with efficiency.
6. If the number of occurring set in pair of data stream is other than even then frame the data set with considering the previous level of low efficiency.

The following are the effectiveness of data streaming detection

1. Early Detection Analysis
2. Effects of Multi-Scale Data streaming Detection
3. Keywords Coherence

Table1 : Data Set with Various Characteristics of Streaming

	Data1	data2	data 3	data 4
data set	numeri-cal	categorical	categorical	none
cate-gory	grouping	classifica-tion	structured	unordered
stream-ing	data1to data2	data2to-data3	data3to-data4	data4to-data5
analy-sis	data-set1tores ult	dataset2to-dataset1&re sult	dataset3to-dataset2&r esult	dataset4to-dataset3&r esult

Table 2: Data Sampling e with Datasets

sample data analy-sis	D1	D2	D3	D4
data1	12	45	67	93
data2	15	67	13	78
data3	34	67	24	79
data4	90	56	45	23

Table 3: Dataset to Result for all Analysis

result analysis to datasets where each data set mapped with result of data stream					Data stream sum	analy-sis ratio to result
data-set1toresult	12	45	67	93	217	0.06034
data-set2toresult	15	67	13	78	173	0.2521
data-set3toresult	34	67	24	79	204	1
data-set4toresult	90	56	45	23	214	null

Table 4: Analysis of Datasets

observations:	
analysis ratio in dataset1	0.6 to 1
analysis ratio in dataset2	0.25 to1
analysis ratio in dataset3	1
analysis ratio in dataset4	null

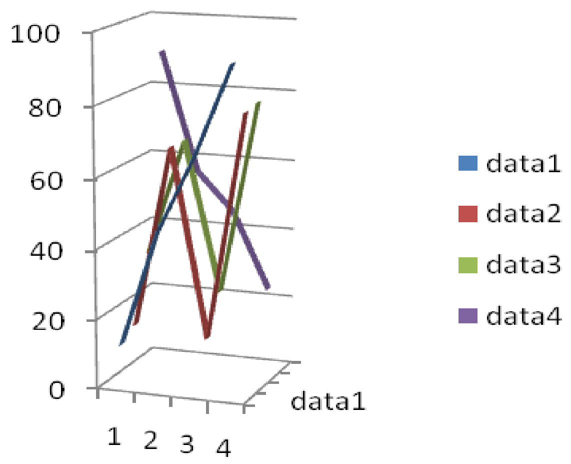

Figure 2: Data Analysis and Co-Relation of Datasets

We take the dataset of all category detection and emerging event detection using keyword co-occurrence. It has long been recognized that modeling topics or events based on keyword co-occurrence is an effective approach. Co-occurrence information has been used for term cluster RTMS and keyword extraction from documents.

Discovered that a word's contexts of different meanings could be represented as overlapping communities in the word co-occurrence data stream. A short text topic model that directly models the generation of word co-occurrence pattern has been proposed addressed event detection from news articles and performed topic detection for large and noisy social media collections .

III. Conclusions

RTMS a real-time emerging event monitor RTMS system over data platforms that integrate our efforts on both emerging event monitor RTMS research and system research. RTMS is able to monitor emerging event as to detect emerging events, build event correlations and trace event evolutions. Further, RTMS's infrastructure is equipped with customized optimization on its full-text search engine and distributed data stream processing engine to perform event monitor RTMS more efficiently. What's more RTMS supports event and text queries and much other functionality to assist the analysis of emerging events, as demonstrated in the user interface.

REFERENCES:

[1] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in WWW, 2010.

[2] R. McCreadie, C. Macdonald, I.

Ounis, M. Osborne, and S. Petrovic, “Scalable distributed event detection for twitter,” in IEEE BigData, 2013.

[3] Y. Chen, H. Amiri, Z. Li, and T.-S. Chua, “Emerging topic detection for organizations from data s,” in SIGIR, 2013.

[4] W. Xie, F. Zhu, J. Jiang, E.-P. Lim, and K. Wang, “Topicsketch: Realtime bursty topic detection from twitter,” in ICDM, 2013.

[5] E. Schubert, M. Weiler, and H.-P. Kriegel, “Signitrend: scalable detection of emerging topics in textual streams by hashed significance thresholds,” in KDD, 2014. F. Wei, S. Liu, Y. Song, S. Pan, M. X. Zhou, W. Qian, L. Shi, L. Tan, and Q. Zhang, “Tiara: a visual exploratory text analytic system,” in KDD, 2010.

[6] P. Lee, L. V. Lakshmanan, and E. E. Milios, “Incremental cluster evolution tracking from highly dynamic network data,” in IEEE International Conference on Data EngineeRTMS (ICDE), 2014, pp. 3–14.

[7] C. Li, A. Sun, and A. Datta, “Twevent: segment-based event detection from tweets

* * * * *