Khoj

Discovering for Local Events



Tribhuvan University

Department of computer science and information technology

Submitted By Submitted to

Khoj Team:

Aayush Basnet

Jitendra Khadka

Nischal Kunwar

Utshav Paudel

[Teacher Name]

**Acknowledgement:**

We extend our heartfelt gratitude to Madan Bhandari Memorial College for providing us with the platform and resources to undertake this project.

I am deeply indebted to our esteemed faculty advisor, [Teacher Name], whose guidance, expertise, and unwavering support were instrumental in shaping our project's direction and success. His dedication to our growth and development has been invaluable, and we are truly grateful for his mentorship.

We also wish to express our appreciation to our fellow students at Madan Bhandari Memorial College for their collaboration, insightful discussions, and constructive feedback, which enriched our project and contributed to its refinement.

Lastly, we commend the commitment and support of each individual. Their dedication, support, and perseverance were vital in overcoming challenges and achieving our goals.

This project stands as a testament to the collective effort and support of everyone involved, and we are deeply appreciative of their contributions.

Khoj

**Contents**

1 Introduction….……………………………………………………

2 Related Work ……………………………………………………

3 The proposed Local Event Schema...…………………………

4 Experiment ……………………………………………………….

5 Discussion………………………………………………………..

6 Conclusion………………………………………………………..



Khoj Team

Abstract

In this paper, we propose a local event detection scheme by analyzing relevant documents in social networks to improve the accuracy of event detection. To detect local events by using geographical data, the proposed scheme embeds them using a geographical data dictionary and generates a weighted keyword graph using social network characteristics. The data left by users in social networks include not only postings but also related documents such as comments and threads. In this way, the proposed scheme detects a local event based on a keyword graph that is constructed through the analysis of the relevant documents. This can improve the accuracy of local event detection by analyzing relevant documents embedded with region-related information using a geographical data dictionary, without requiring users to tag geographic data. In order to verify the superiority of the proposed scheme, we compare it with the existing event detection schemes through various performance evaluations.

**Personal Foreward by Khoj Team for Readers**

Our local event detection scheme can be used in various applications such as traffic flow control services, event location finding services, intrusion detection services, and disaster prevention services. In these applications, our scheme can be used to find the local events in the real world, including accidents, state and county fairs, city festivals, circuses, protests, sport games, flea markets, other public gathering events, and natural disasters.

Life is no brief candle for me. It is a sort of splendid torch which I have got hold of for the moment, and I want to make it burn as brightly as possible before handling it on to future generations.

George Bernard Shaw

Introduction

Recently, with the popularization of smart devices, social network services (SNSs) have been widely used to communicate and share information among users. SNSs have been used not only to make personal connections but also to rapidly deliver information when a local event occurs. A local event means an event in the real world, including state and county fairs, city festivals, circuses, protests, sport games, flea markets, and other public gathering events, or naturally occurring incidents, such as disasters. SNS users upload information and posts in real time to share meaningful information when a local event occurs at a particular time and place. When Hurricane Sandy hit the eastern part of the U.S. in 2012, a refueling fiasco arose to secure oil in order to operate power generators for each household due to the lack of electric power supply caused by damaged transmission towers. At that time, people shared conditions, contact numbers, and waiting times of gas stations through SNSs that allow users to share information in real time. In the situation of a disaster, SNSs have merits that can allow individuals to utilize highly reliable collective intelligence .

Event detection schemes based on SNSs have been studied extensively. The posts written by users have various information, such as hashtags, situations, time, and locations. Text-based event detection schemes have been proposed for considering the importance and frequency of keywords by analyzing and extracting meaningful words on hashtags, contents, and comments. Text-based event detection schemes utilize term frequency–inverse document frequency (TF-IDF) to extract keywords. Recently, some studies utilized machine learning algorithms such as Word2Vec. Graph-based event detection schemes, which involve making a graph from extracted keywords and conducting clustering algorithms to find events, have also been proposed. The authors generated a graph by adapting vector space models; the graph can detect events from clusters that are built based on weights assigned to the graph considering the frequency of co-occurrences.

Recently, local event detection schemes using geo-tags in SNSs have been studied [[**17**](https://www.mdpi.com/2076-3417/11/2/577#B17-applsci-11-00577),[**18**](https://www.mdpi.com/2076-3417/11/2/577#B18-applsci-11-00577),[**19**](https://www.mdpi.com/2076-3417/11/2/577#B19-applsci-11-00577)]. Zhang et al. [[**17**](https://www.mdpi.com/2076-3417/11/2/577#B17-applsci-11-00577)] proposed a local event detection scheme that uses time-stamps and geographical information included in the posts. Choi et al. [[**18**](https://www.mdpi.com/2076-3417/11/2/577#B18-applsci-11-00577)] proposed a geo-tag based local event detection scheme. However, since the proposed scheme in [[**17**](https://www.mdpi.com/2076-3417/11/2/577#B17-applsci-11-00577)] performs text-based event detection, there is a problem when little information is obtained in the event of a short posting and hence less noise in the process of extracting keywords related to local events, resulting in poor accuracy. In addition, since the actual majority of social network data do not have geo-tagging, the event detection scheme using only the geo-tagging approach proposed has a problem with poor accuracy. In this paper, we propose a local event finder by analyzing relevant documents in social networks. The proposed scheme creates a keyword graph by extracting keywords in the social network data. After that, we assign weights to vertices and edges of the keyword graph by considering social network characteristics. We can easily understand the relationships between an event and keywords by using the keyword graph. In order to detect local events by using geographical information, the proposed scheme utilizes a geographical dictionary and geo-tag information to classify the geographical keywords from the keyword graph. The main contributions of this paper are as follows:

1. The proposed scheme uses a geographical dictionary to solve the limitations of the existing schemes that provide and detect local events, considering that geo-tags appear very sparsely with posts. The geographical dictionary refers to a database that consists of location data mapped to a noun that represents a particular region or area.
2. The proposed scheme analyzes the related documents such as comments and threads to improve the accuracy of the event detection scheme. After the analysis, we apply a clustering algorithm to the keyword graph according to weight values.
3. The proposed scheme extracts and provides the local events by merging and dividing clusters by using an edge weight of in-cluster and out-cluster.
4. We compare the proposed scheme with the existing event detection schemes through various performance evaluations in order to verify the superiority of the proposed scheme.

The proposed Local Event Finder

In this paper, we propose a local event detection scheme by analyzing relevant documents to solve the problems of the existing schemes such as the geo-tag and text mining based schemes. We embed geographical information to the detected events by using text mining techniques and a geographical dictionary if the events do not have geo-tags. In order to complement sparse geographical data, we exploit relevant documents that have geo-tags or geographical words