GIS with Crowdsourcing as a Service

Nantanapak Saranpattarawet, Yossapol Jearamanytwesin, Wanida Putthividhya
Thammasat University, Thailand

INTRODUCTION

GIS with Crowdsourcing as a Service (GCaaS) provides a service for creating and maintaining Geographic Information System based web site (GIS-based web site) without requiring any specific knowledge. The GIS-based web sites created by GCaaS (known as deployments) can be used to gather data from various sources, e.g. remote sensors, specific-purpose mobile applications, social media network Twitter, etc., and display the gathered data in a meaningful manner on maps. Filtering data from Twitter uses hashtags specified a priori by deployments' administrators. Deployments created by GCaaS version 1 and 2 have two major limitations. They merely support gathering data relating to disaster events. Besides, tweets attached with erroneous hashtags will not be filtered to process further at all.

PURPOSE

GCaaS version 3 aims at improving GCaaS's module which is mainly responsible for filtering data from social media networks. With the improved data filtering module, GCaaS will be able to support both disaster and regular events of interest. In addition, the improved data filtering module will be able to handle erroneous hashtags up to a certain degree. As a result, we can increase an opportunity to filter valuable information into deployments.

RELATED WORK

GCaaS version 1 and 2 support two hashtag patterns of disaster event as follows:

*"#GCaaS #<deployment's name> **#HELP** #<help needed> #<help no.>" and*
*"#GCaaS #<deployment's name> **#ROAD** #<road condition(good,damage,severe)>".*

METHODOLOGY

Each deployment created by GCaaS contains a module which is responsible for filtering data from social media network Twitter. This module filters tweets relating to deployments based on hashtags specified by deployments' operators. Then, the information associated with tweets are extracted and displayed on maps. GCaaS version 3 enhances this data filtering module in two aspects.

Firstly, the data filtering module in GCaaS version 3 allows deployments' operators to specify hashtags as desired. As a result, GCaaS version 3 can create GIS-based web sites for both disaster and regular events of interest. The variety of events GCaaS can support enhances the impact of GCaaS.

Second, the data filtering module has been enhanced in GCaaS version 3 to support erroneous hashtags up to a certain degree. Specifically, deployments created by GCaaS version 3 would be able to filter tweets with erroneous hashtags. Tweets associated with correct hashtags will be filtered into deployments Figure 2 illustrates the algorithms used to handle erroneous hashtags in various scenarios.

The first algorithm is executed when the main hashtag is meaningful such as #Donate, #University, #Advise, etc. The second algorithm is used when the main hashtag is meaningless; however, its substrings are meaningful such as #gunshot, #planecrash, #BikeForMom, etc. The third algorithm is applied when both main hashtag and its substrings are meaningless #LAX, #Y2K2000, #Kaokonlakao, etc.

ARCHITECTURE

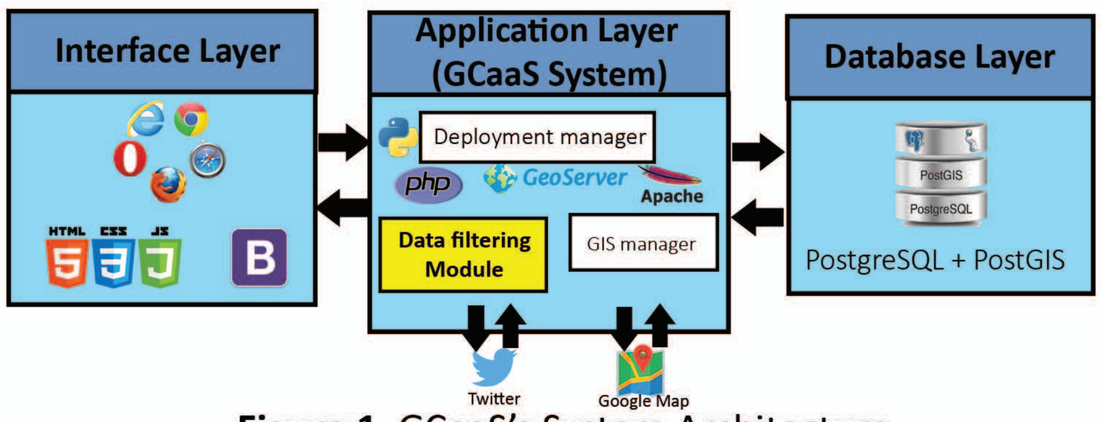

Figure 1 GCaaS's System Architecture

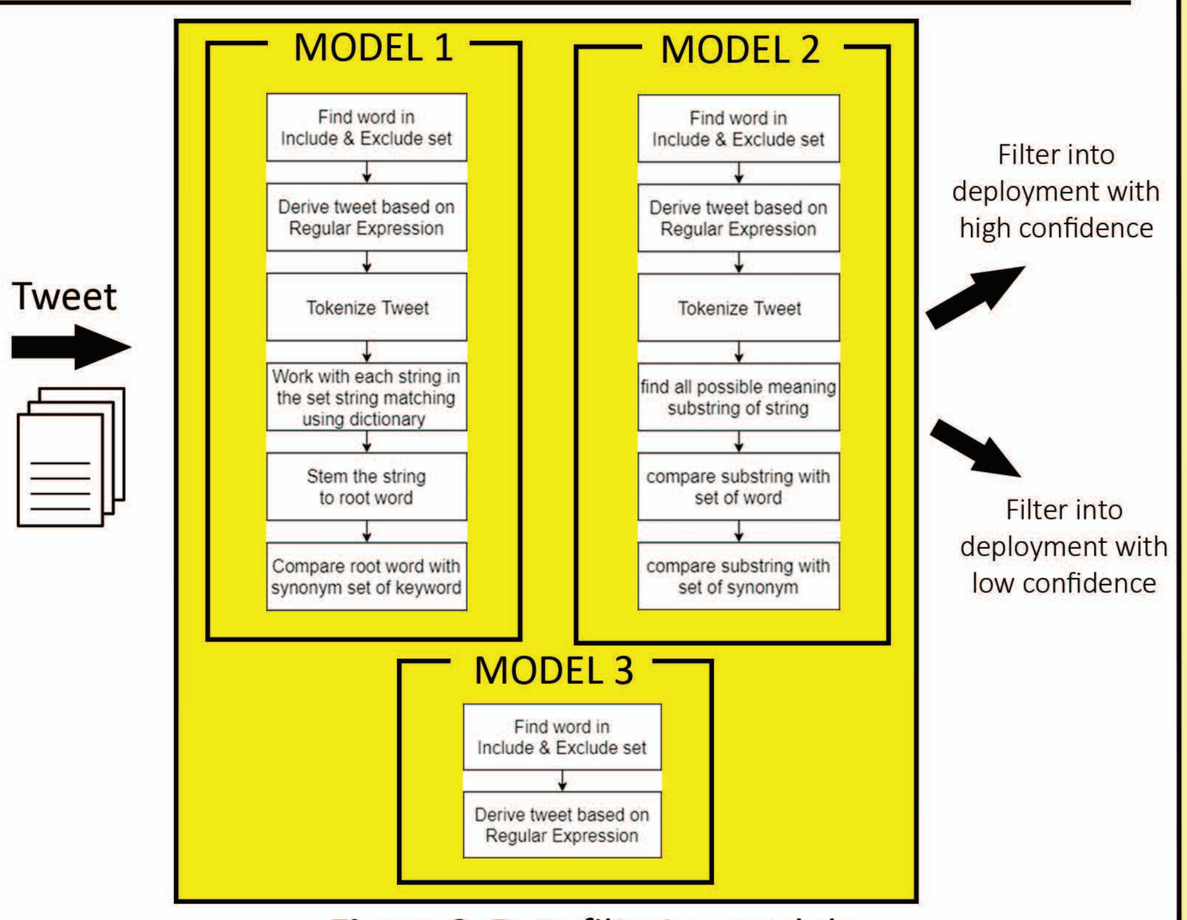

Figure 2 Data filtering module

RESULTS

Main hashtag of deployment	Actual hashtag in tweet	No. of tweets	Numbers of tweets that each version supported	
			V.1 & V.2	V.3
HELP	#HELP	20	20	20
	Part of misspell hashtag such as #Help, #help, #hepl, #need	70	0	70
	Other hashtag such as #haha	10	0	0
ROAD	#ROAD	20	20	20
	Part of misspell hashtag such as #Road, #road, #road, #way	70	0	70
	Other hashtag such as #red	10	0	0
gunshot	#gunshot	20	0	20
	Part of misspell hashtag such as #Gunsoht, #gunshoter, #gunshoot, #Shooter, #gunshooter, #shooting	70	0	70
	Other hashtag such as #shotshot	10	0	0
Total		300	40	270

The results showed that data filter module of GCaaS version 3 outperformed its predecessor in GCaaS version 1 and 2. The improved data filtering module support erroneous patterns and can filter more tweets relating to deployments. This results in rising up an opportunity to filter more valuable information into deployments.

CONCLUSION

GIS with Crowdsourcing as a Service (GCaaS) version 3 enhanced the data filtering module to handle erroneous hashtags. The improved module is mainly responsible for filtering data from social media networks into deployments. Our results showed that the new data filtering module can filter more tweets relating to doployments compared to the previous version in GCaaS version 1 and 2.

REFERENCES

Techin Maneesorn and Oraphan Itritmeechai, **GIS with Crowdsourcing as a Service (GCaaS)**, Thammasat University

Warittha Ruangjarus and Noppawan Joysa, **GIS with Crowdsourcing as a Service Version 2**, Thammasat University