ISSN 2319-2720
Volume 7, No.2, April – June 2018
International Journal of Computing, Communications and Networking
Available Online at http://www.warse.org/ijccn/static/pdf/file/ijccn43722018.pdf
https://doi.org/10.30534/ijccn/2018/43722018

SECURE ANALYSIS OF SENSOR DATA IN BIG-DATA PLATFORM FOR GEO-SOCIAL NETWORK

Vardhini HP[1], Prof. ChetanaSrinivas[2]
[1]Student EWIT, vardhinigowda@gmail.com
[2]Associate Professor, Dept. of computer Science &Engineering, EWIT, chetanasrinivas@ewit.edu

ABSTRACT:

Geosocial Network information can be provided as service for the users to take decisions for the events that occurs in the real-world like earth quakes, bombblast, and fire and also other events occurred in the different places in the world. After analyzing the data the user can able to take decisions for the real-time to predict for the future which are the places that will be affected by the particular events.We propose a system that should be able to handle the data generated by the different sensors in a secure manner the data sent by the sensors will be stored in the encrypted format in the Geo-social network. When the user is searching for the particular event that occurred the user need to decrypt the data before using the information**.** Twitter is analyzed using the proposed system to recognize the events like earth quakes, fire, Bomb related data. We are using the Hadoop system for analysis of the big-data generating by the different sensors.

(Keywords—secure,sensor,Big data,hadoop)

1. INTRODUCTION

Social networking is radically propelling their element day by day while making themselves from informal communities to Geosocial Systems. It engages individuals to make their substance open alongside their geological data. This has brought about an expansion in the utilization of Geosocial Networks by giving clients with the capacity to voice assessments, report occasions, and offer perspectives, outrage, or love while associating with others, which was unbelievable in the pre-Internet age. The data partook in any media is geosocial in light of the fact that: 1) the posts have broad content that speaks to land data with particular areas that are either entered unequivocally (with registration) or included verifiably (by Earth arranges, for example, scope or elevation), and 2) the perspectives shared via web-based networking media uncover social information and reinforce relationship and correspondence.

Utilizing geosocial arrange information isn't just valuable to governments, yet it can likewise majorly affect human life. Geosocial Network information can give advantages to ordinary nationals also, businessmen. Nonetheless, when reaping geosocial information from systems, for example, Twitter or Facebook, it ought to be noted that these systems have a huge number of clients who post thousands of tweets and statuses within 60 minutes. In this way, it can be effortlessly contemplated that every one of the clients of different informal organizations creates a noteworthy measure of information: such information may go in the terabytes inside minutes. Thusly, gathering such realtime geo-social information is an extremely difficult errand. We require an extraordinary computational condition and propelled processing procedures with savvy administration keeping in mind the end goal to give in time/continuous investigation. All the previously mentioned methods do not think about in excess of one informal community at any given moment, and their investigations are adaptable as far as information estimate. In this way, with a specific end goal to address these computational difficulties, in this paper, we propose a progressed geosocial information expository framework that not just forms disconnected information proficiently inside a period constrain yet in addition gives

continuous information investigation to different interpersonal organizations, counting Twitter, Flickr, Facebook, YouTube, and so on. The framework sends a Hadoop biological community for information handling and investigation We tried the framework by taking two interpersonal organizations, Twitter Whatever is left of the report portrays the proposed framework.

2. OBJECTIVE

The main objective of the project is to generate secure sensor data for generating of the particular events like fire, Bomb blast and Earth quake by taking the secure data the user can search for the particular events get the secure information.

3. PROBLEM STATEMENT

The data generated in the system is not stored in the secure way.Any un-authorized user can access the information present in the Geo-social network. To provide a security for the system we are using ECC algorithm for securing data.

4. LITERATURE REVIEW

Stefanidis proposes a method for Online networking created from numerous people is assuming a more prominent part in our everyday lives and gives a one of a kind chance to increase significant knowledge on data stream and informal communication inside a society. Through information gathering and examination of its content, it underpins a more prominent mapping and comprehension [1] of the developing human scene. The data scattered through such media speaks to a deviation from volunteered geology, as in it is not geographic data in essence. By and by, the message regularly has geographic impressions, for instance, as areas from where the tweets start, or references in their substance to geographic substances.

Crooks approaches method for Online networking sustains are quickly developing as a novel road for the commitment and dispersal of data that is regularly geographic. Their substance regularly incorporates references to occasions happening at, or influencing particular areas. Inside this article we investigate the spatial and fleeting qualities of the twitter[2] channel action reacting to a 5.8 greatness quake which happened on the East Coast of the United States (US) on August 23, 2011. We contend that these bolsters speak to a crossover type of a sensor framework that takes into account the recognizable proof and limitation of the effect zone of the occasion. By standing out this from the amount substance gathered through the devoted crowdsourcing 'Did You Feel It?' (DYFI) site of the U.S. Topographical Survey we evaluate the capability of the utilization of collected online networking content for occasion observing.

M.Zook said that The paper plots the manners by which data advances (ITs) were utilized as a part of the Haiti help exertion, particularly as for online mapping administrations. Despite the fact that there were various manners by which this occurred, this paper centers around four specifically: Crisis Camp Haiti, OpenStreetMap, Ushahidi, and GeoCommons. This examination shows that ITs where a key means through which[3] people could have an unmistakable effect in crafted by alleviation and help offices without as a matter of fact being physically present in Haiti. While not without issues, this exertion all things considered speaks to a wonderful case of the power and crowdsourced on the web mapping and the potential for new roads of connection between physically removed places that differ hugely.

Frenchman proposes technique that innovation for deciding the geographic area of phones and other handheld gadgets is winding up progressively accessible. It is opening the path to an extensive variety of uses, on the whole alluded to as area based administrations (LBS) that are principally gone for singular clients. In any case, if sent to recover totaled information in urban communities, LBS could turn into a capable instrument for urban[4] investigation. In this paper we plan to survey and present the capability of this innovation to the urban arranging group. Moreover, we show the `Mobile Landscapes' venture: an application in the metropolitan territory of Milan, Italy, in view of the geological mapping of wireless utilization at various circumstances of the day.

Programming design inquire about researches techniques for deciding how best to parcel a framework, how parts distinguish and speak with each other, how data is imparted, how components of a framework can develop freely, and how the greater part of the above can be depicted utilizing formal and casual documentations. My work is spurred by the want to comprehend and assess the engineering outline of network based application programming[5] through principled utilization of engineering requirements, in this manner acquiring the practical, execution, and social properties wanted of a design. An engineering style is a named, composed arrangement of building requirements

5. EXISTING SYSTEM

In the existing system the sensors sensed data will be stored in the Geo-social network.The data will be stored in the form of normal text.The un-authorized users also searching for the in which particular area where the events like Earth quake, Bomb blasting, fire will be occurred. The user can also get the information easily.There is no security for the data access.

6. PROPOSED SYSTEM:

In the proposed System the sensor sensing the values like in which particular region the events occurred the data can be encrypted by using the particular sensor public key.Only the authorized user can search in the Tweets where the particular event occurred. By using the particular private key the user can able to decrypt the data can be used for future decisions.ECC (elliptic curve cryptography) is a public key cryptography that will be used for encrypting and decrypting the data present in the Geo-social network.Twitter data can be analyzed by using the Proposed architecture.we are generating here secure Big-data to handle that we are using Hadoop system.

7. ARCHITECTURE:

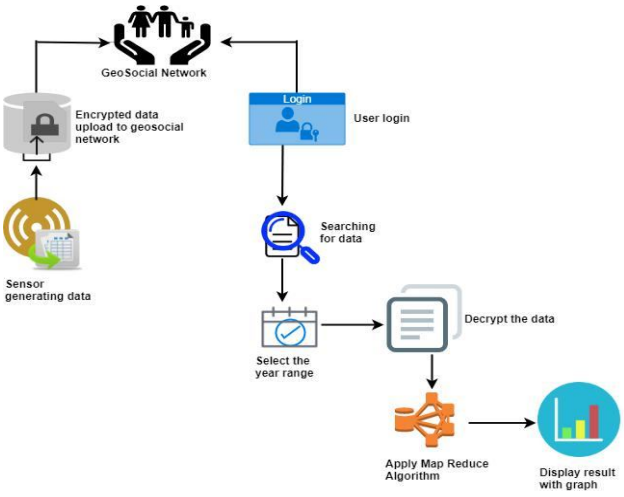

Fig1: Overall architecture of the Proposed system

In Fig1 it shows the architecture of the proposed System .It contains the below elements.

1. Sensor: To generate the particular event related to the disasters the sensor should sense the data in which particular area the event occurred.

2. Upload data: After sensor generating the data the data sent to the geosocial network in the form of encryption the data will be stored.For encryption we are applying ECC algorithm

3. Geo-social network:In the geo-social network all the sensed data will be stored.

4.User: The user should login to the system then user can able to search for the particular data for which the event occurred.After login the user can search for the data in year range.In which particular year the event in which place the particular event occurred the user can decrypt the data.

5.Map reduce:After user decrypt the data we can use the map reduce function and will get the count of each particular event in which countries occurred.The results will be display in the form of graph.

8. ALGORITHM

ECC (elliptic curve cryptography technique will be used for encryption and decryption of data.

The following are symbols we are used,

E-->Elliptic curve

P-->Point on the curve

n-->/Maximum limit(prime number)

- **Generation of keys:**

Keys will be used for the encryption and decryption.Here we are using public key for encryption and private key for decryption.We have to choose a number **'s'** with in range of **'n'**.Using following formula we can generate the public key

→W=s*p

Where s=The random number selected within the range(1 to n-1) P is point on curve.

→**'W'** is public key and**'s'** is private key.

- **Encryption**

Assume **'x'** is the data that sensed by the sensor sent to the Geo-social network.Represent this data on curve.Consider**'x'** as point **'M'** on the curve **'E'**.Randomly select **'k'** from[1-(n-1)].Two cipher texts will be generated let be m1 and m2

M1=k*p

M2=M+k*W

- **Decryption:**

We have to decrypt the data send by senor

X=M2-s*M1;

Where x is the original message.

- **Proof:**

X=M2-s*M1

'x' can represent as **'**M2-s*M1**'**

M2-s*M1=(x+k*Q-s*(K*P) (M2=x+K*Q & M1=K*p)

=X+k*s*P-S*K*p (cancel k*s*p)

=x(original message)

9. RESULT AND ANALYSIS

To analyze the proposed system we are using the Twitter dataset which can contains the various diseaster information like earth quake, fire, Bomb event we are using dataset from year 2000 to 2010.We are analyzing the all Twitter tweets. The analysis is conducted based on the user searching for which particular event. Here we are also showing results for graph of Encryption time and decryption Time taken for particular event data.

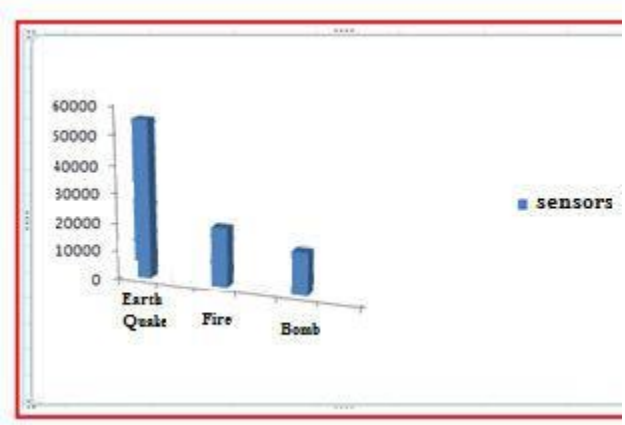

Fig2:Shows the number of tweets of each events.the tweet contains #Earth quake,#Fire,#bomb.

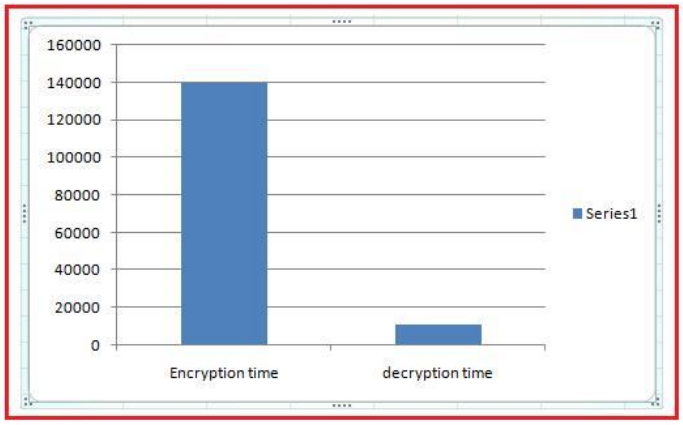

Fig3:Shows the data security Time taken for the encryption and decryption .Here the encryption time is more because all the sensor need to sense and send the data.Decryption Time is less because after searching the user can download the data.

10. CONCLUSION

Geosocial Networks can be an advantage for governments in terms of giving offices . Additionally, such systems can profit to normal nationals by giving prescribed frameworks, transport security, social insurance, and so forth., and to business visionaries for propelling new items in different zones by observing the geosocial information of a specific zone. In any case, such advantages must be inferred with better examination that utilizes a lot of information produced from different Geosocial Networks. This is conceivable with propelled innovation and better examination, and a framework with high processing capacities. Along these lines, in this paper, we proposed a framework that utilizations geosocial information for better arranging, wellbeing from fiascos, and appropriate administration, mindfulness, and so forth., in light of different geolocations. The framework not exclusively can reap a lot of information at fast from Geosocial Networks, however, it can likewise process, investigate, and settle on choices continuously. We investigated Twitter information for different occasions utilizing the proposed framework. The framework was produced utilizing a Hadoop biological community with Spark. The framework was more proficient when preparing a part of datasets, and demonstrated the benefit of expanded throughput with an expansion in information volume.

REFRENCES:

1. Stefanidis, A.crooks, and J.Radzikowski,**"**Harvesting ambient geospatial information from social media feeds,**"**GeoJournal, pp.1-20, 2012.

2.Crooks, A.Croitoru, A.stefanidis, and J.Radzikowski,**"**#Earthquake: Twitter as a distributed sensor system,**"**Transactions in GIS, 17(1), pp.124-147, 2012. https://doi.org/10.1111/j.1467-9671.2012.01359.x

3. M. Zook, M. Graham, T. Shelton, and S. Gorman, **"**Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake,**"** World Medical & Health Policy, vol. 2, no. 2, pp. 7–33, 2010. https://doi.org/10.2202/1948-4682.1069

4. Ratti, S.Williams, D.Frenchman, andR.pulselli,**"**Mobile landscapes using location data from cell phones for urban analysis,**"**Envoirnment and planning and Design.

5. Fielding.RT.,**"**Architecture atyles and design of networks nased software architectures, **"**Thesis. University of California.

6. R. Lan, M. D. Lieberman, and H. Samet, **"**The picture of health: mapbased,collaborativespatio-temporaldiseasetracking,**"** HealthGIS'12, pages 27-35, Redondo Beach, CA, 2012. https://doi.org/10.1145/2452516.2452522

7. L. A. Waller and C. A. Got way, **"**Applied Spatial Statistics for Public Health Data,**"** Book: John Wiley & Sons, Hoboken, New Jersey, volume368. 2004.

8. Rathore, M. Mazhar, et al. "Real-time Medical Emergency Response System: Exploiting IoT and Big Data for Public Health." Journal of medical systems 40.12 (2016): 283.

9. Ahmad, Awais, et al. "An Efficient Multidimensional Big Data Fusion Approach in Machine-to-Machine Communication." ACM Transactions on Embedded Computing Systems (TECS) 15.2 (2016): 39.

10. Rathore, M. Mazhar, et al. "Exploiting encrypted and tunneled multimedia calls in high-speed big data environment." Multimedia Tools and Applications (2017): 1-26.

11. Fielding, R.T., **"**Architectural styles and the design of network-based software architectures,**"** Thesis (PhD). University of California, Irvine, CA, 2000.

12.Eriksson, Brian, Paul Barford, Joel Sommers, and Robert Nowak, "A learning-based approach for IP geolocation," In International Conference on Passive and Active Network Measurement, pp. 171-180. Springe Berlin Heidelberg, 2010. https://doi.org/10.1007/978-3-642-12334-4_18

13.Poese I, Uhlig S, Kaafar M A, Donnet B, and Gueye B, **"**IP Geolocation databases: Unreliable?**"** Computer Communication Review 4 (2): 53–56,2011. https://doi.org/10.1145/1971162.1971171

14. Twitter stream, **"**Archive Team: The Twitter Stream Grab,**"**https://archive.org/details/twitterstream, [Accessed on 19/02/2016].

15. TWEET MAP, **"**MAPD Twitter,**"** http://tweetmap.mapd.com, [Accessed on 19/02/2016].