International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 3, March 2016

Social Networking for City Safety - Abilities' And Unattended Challenges In Techniques Foreseeing Safety Value

[1]C.S.Saradha, [2]Dr.P.Arul
[1][2] Associate Professor
[1]School Of Computer Science [2] Department Of Computer Scicence
[1]CMS College of Science & Commerce
[2]Govt Arts College Trichy
Coimbatore
[1]sara.kumarasamy@gmail.com [2] phdarul2004@yahoo.co.in

Abstract: **"Power of Social Conversation" have moved beyond just a passing fed. People spend a staggering 700bn minutes per month on face book, twitter and other social networking sties. This social media has become part of how we lead our lives and the everyday events that are part of it. The maze of pictures, comments, events and social gaming could be our window into a gold mine of insights to predict catastrophic events especially around public security and disaster management.**

Main responsibility government agencies and municipalities to leverage the power of broadband connectivity to embrace the safe city approach, with centralized emergency response and management in the face of threats, attack, natural disasters, crime or industrial-scale accidents. In other words, effective communications is the groundwork of protecting lives and mitigating damage.

This paper studies abilities and unattended challenges of various techniques which forecasts smart, safe cities, by combining the use of personal mobile devices and social networks to make users aware of the safety of their surroundings. And hence Proceeds towards Limitations in the existing techniques and its Research challenges.

***Index Terms*—ARIMA,, PredPol software, NoVA**

I. INTRODUCTION

Metro cities combine technology and human resources which improves the life's quality and reduces expenditures. One of the main open problem which remains unattended is ensuring safety of city.

Public security remains one of the biggest problems facing today's federal and local governments. Traditional method of policing have failed as criminals and terrorists have adopted technology and social channels with great poise. 66% of all US government agencies currently use some form of social media monitoring - from blogs and wikis to instant messaging and discussion boards, according to a recent study.

Police agencies could use big data mining with text analytics and decision engines to rank and filter targeted information regarding social behaviour. Social media can serve the role of citizen communication on important events or crimes by communicating across channels. [1]

We study on various suite of techniques for defining the safety of locations and users.

II. EXISTING SOCIAL NETWORKS

a. Face book: Face book is a social networking service and website with more than 845 million active users. The site allows users to post messages, photographs, and other files for other users to consume, and it supports user-defined controls over the visibility of these items. It uses Regional Records Management

Systems RMS interoperability set-up, which has several free marketing tools that enhance project communication. Face book could be used to highlight major project milestones, such as end users' completing an onsite product demonstration assessment: e.g. "The selected RMS vendor's system offered all the functionality we needed in the field, and it was easy to use." [2]

b. YouTube: YouTube is a video-sharing website on which users can upload, view, and share videos.5 YouTube makes it possible for anyone with an Internet connection to post a video that a worldwide audience could watch within a few minutes. According to YouTube, the site receives more than 3 billion views per day. In case of RMS setup, YouTube could be used to post just-in-time training. For example, several new software which illustrates geographic mapping of crime could be recorded, narrated, and posted to YouTube in support of in-service training. [3]

c. LinkedIn: LinkedIn is a business-related social networking site with more than 120 million registered users in more than 200 countries. On LinkedIn, users invite other users to "connect," based on some common interest or professional relationship. Once a user accepts an invitation, the two users are connected, which allows them to receive information about one another, send one another messages, and recommend one another to other users. In case of RMS setup, the members of team could post their professional profiles to LinkedIn, noting their professional experience and abilities so that other team members could see how their skill sets fit within their assigned project roles and responsibilities. The project manager could start a private ("members only") group for the project, invite all the stakeholders to join, and then use the group for discussions and dissemination of project news. [4][5]

d. Twitter: Twitter is an online social networking and micro blogging service that enables users to send and read text-based posts of up to 140 characters, informally known as "tweets." As of 2012, Twitter has more than 300 million users, generating more than 300 million tweets and handling more than 1.6 billion search queries per day. In case of RMS setup, implementation has been completed, Twitter messages could be sent out to operational personnel to remind them of key user functions. This tool can be used to market the ongoing value and utility of the RMS to operational users. [6]

III. REGISTERED MISDEED DATA

A historical database of more crime incidents are reported and labelled with a crime type (e.g., homicide, larceny, robbery, etc), the time and the geographic location where it has occurred. These data's are briefly documented for pre-processing. First, since records come from different Police departments, the crime type labels are non-uniform, (e.g., *murder* in Miami Beach vs. *homicide* in North Miami). Second, crime reports include many minor incidents (e.g., fire alarms issues), resulting in variety of crime types.
The ambiguities are eliminated by mapping them to different labels such as Murder, Forcible Rape,

Aggravated Assault, Robbery, Larceny/Theft, Burglary/Arson, Motor Vehicle Theft. Manual mapping is infeasible due to large number of records, hence machine learning techniques are used namely Naive-Bayes (NB) classifier and the Decision Trees (DT) classifier. [7]

A crime type classifier such as NLTK library [8] is used for classification that predicts the correct crime in which the NB classifier achieved an accuracy of 91% and the DT classifier an accuracy of 98%. Thus, the outcome of the DT classifier is used.

IV. COMPARING VARIOUS FORECASTING METHODS

a. ARIMA Model: Autoregressive Integrated Moving Average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data for three main reasons. (i) to better understand the data (ii) to predict future points in the series (forecasting) and (iii)to apply where data show evidence of non-stationary. This model uses parameters *p*, *d*, and *q* which are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. The ARIMA forecasting procedure consists of four steps, (1) identifying the ARIMA(p, d, q) structure, (2) estimating the unknown parameters, (3) fitting tests on the estimated residuals and (4) forecasting future outcomes based on historical data [9]. The Generalized model of ARIMA is given by,

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t = \delta + \left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t$$

where L is the lag operator, α_i are the parameters of the autoregressive part of the model, θ_i are the parameters of the moving average part and they ε_t are error terms. The error terms ε_t are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.[9.1]

This defines an ARIMA(p,d,q) process with **drift** $\delta/(1-\Sigma\varphi_i)$.

b. Exponential Smoothing: Exponential smoothing is a technique that can be applied to time series data, either to produce smoothed data for presentation, or to make forecasts. The time series data themselves are a sequence of observations. The observed phenomenon may be an essentially random process, or it may be an orderly, but noisy, process. Whereas in the simple moving average the past observations are weighted equally, exponential smoothing assigns exponentially decreasing weights over time. The raw data sequence is often represented by $\{x_t\}$, and the output of the exponential smoothing algorithm is commonly written as $\{s_t\}$, which may be regarded as a best estimate of what the next value of x will be. When the sequence of observations begins at time $t = 0$, the simplest form of exponential smoothing is given by the formulae:

$$s_0 = x_0$$
$$s_t = \alpha x_t + (1-\alpha)s_{t-1}, \ t > 0$$

Where α is the *smoothing factor*, and $0 < \alpha < 1$.[10]

c. Artificial Neural Network (ANN): ANNs are data-driven self-adaptive methods that learn and generalize from experience and capture subtle functional relationships among the empirical data even if the inherent relationships are unknown or difficult to describe. further, the multilayer perceptrons (MLP) ANN model, which is particularly suitable for forecasting, due to its ability for input-output mapping. The input layer (of the same size as the input vector), two layers of hidden nodes and an output layer providing the forecast value. The input data is normalized to a (−1, 1) range; following the prediction step the output is mapped back to the initial range. For the training phase, multilayer feed forward network trained using back propagation and the Levenberg-Marquardt algorithm to perform function fitting (nonlinear regression). [9]

d. qualitative and quantitative: There are two general approaches to forecasting: qualitative and quantitative. Qualitative methods consist mainly of subjective inputs, which often rebel precise numerical description. Quantitative methods involve either the projection of historical data or the development of associative models that attempt to utilize *causal (explanatory) variables* to make a forecast. Qualitative techniques permit inclusion of *soft* information (e.g., human factors, personal opinions, hunches) in the forecasting process. Those factors are often omitted or downplayed when quantitative techniques are used because they are difficult or impossible to quantify. Quantitative techniques consist mainly of analyzing objective, or *hard,* data. They usually avoid personal biases that sometimes contaminate qualitative methods. In practice, either approach or a combination of both approaches might be used to develop a forecast. [11]

e. NoVA: NoVA is a anti-crime No Violence Alliance which considers Criminals as which are represented as dots on a computer screen and further uses police intelligence to create social networks that link suspects and their associates. City police and federal agents, state and federal prosecutors, social service providers and others in NoVA use the data to target certain people, either for prison or for help.

A NoVA network starts with suspects in murders, shootings or other serious assaults. Other police reports link them to other people and those other people to yet more people. Normally, police go “down the rabbit hole” to find such links after a crime but the network data allows more.

f. Judgmental forecasts rely on analysis of subjective inputs obtained from various sources,
such as consumer surveys, the sales staff, managers and executives, and panels of experts. Quite frequently, these sources provide insights that are not otherwise available.

Time-series forecasts simply attempt to project past experience into the future. These techniques use historical data with the assumption that the future will be like the past. Some models merely attempt to smooth out random variations in historical data; others attempt to identify specific patterns in the data and project or extrapolate those patterns into the future, without trying to identify causes of the patterns.

Associative models use equations that consist of one or more *explanatory* variables that can be used to predict demand. For example, demand for paint might be related to variables such as the price per gallon and the amount spent on advertising, as well as to specific characteristics of the paint (e.g., drying time, ease of cleanup).

Naive Methods A simple but widely used approach to forecasting is the naive approach. A naive forecast uses a single previous value of a time series as the basis of a forecast. The naive approach can be used with a stable series (variations around an average), with seasonal variations, or with trend. With a stable series, the last data point becomes the forecast for the next period. Thus, if demand for a product last week was 20 cases, the forecast for this week is 20 cases.

With seasonal variations, the forecast for this “season” is equal to the value of the series last
“season.” For example, the forecast for demand for turkeys this Thanksgiving season is equal to demand for turkeys last Thanksgiving; the forecast of the number of checks cashed at a bank on the first day of the month next month is equal to the number of checks cashed on the first day of this month; and the forecast for highway traffic volume this Friday is equal to the highway traffic volume last Friday.

For data with trend, the forecast is equal to the last value of the series plus or minus the difference between the last two values of the series. For example, suppose the last two values were 50 and 53. The next forecast would be 56:

Focus Forecasting. Some companies use forecasts based on a "best current performance" basis. This approach, called focus forecasting, was developed by Bernard T. Smith, and is described in several of his books. 2 It involves the use of several forecasting methods (e.g.,moving average, weighted average, and exponential smoothing) all being applied to the last few months of historical data after any irregular variations have been removed. The method that has the highest accuracy is then used to make the forecast for the next month. This process is used for each product or service, and is repeated monthly. Example 4 illustrates this kind of comparison.

Diffusion Models. When new products or services are introduced, historical data are not generally available on which to base forecasts. Instead, predictions are based on rates of product adoption and usage spread from other established products, using mathematical diffusion models. These models take into account such factors as market potential, attention from mass media, and word of mouth. Although the details are beyond the scope of this text, it is important to point out that diffusion models are widely used in marketing and to assess the merits of investing in new technologies.

PredPol software: PredPol's methodology utilizes existing crime data, advanced mathematics, computer learning, and cloud computing. PredPol analyzes data through complex computer techniques and applies theory about crime and gun violence. Then it maps "hot spots "for future crimes by analyzing times, dates, and places of recent crimes. The information is accessible from any tech device or on paper. For security, it is run on a secure, cloud-based software-as-a-service (SaaS) platform. It doesn't collect personal information about victims, offenders, or law enforcement. The PredPol tool was developed over the course of six years by PhD mathematicians and social scientists at UCLA, Santa Clara University, and UC Irvine in collaboration with crime analysts and officers from LAPD and Santa Cruz PD.

V. UNTOLD BENEFITS

Social media tools allow emergency managers to broadcast information to wider audiences, interact with the public, screen social media networks to get a better sense of what's happening on the ground during a crisis, get better situational awareness, and improve teamwork for sharing information during an emergency and sharing of best practices and lessons learned.

VI. TECHNIQUES FOR MEASURING INACCURACY

There are many techniques for measuring errors in accuracy. In which, a. (RSME) Root Mean Squared Error Technique which tends toward more objective measure in implicit magnitude. b. (MAPE) Mean Absolute Percent Error which is likely affected by the magnitude of series. But MAPE does provide data about the relative magnitude of forecast error.[12]

VII. RELATED WORK

Dynamic safety practices acting on social networks and GPS mobile phones have been introduced in [13] to create a system for personalized safety awareness. The multimodal sensing capabilities of devices enable a broad range of applications that leverage collected data from participants, sensed from their surroundings.

A variety of univariate and multivariate methods have been used to predict crime. Univariate methods range from simple random walk to more sophisticated models like exponential smoothing. While exponential smoothing offers greater accuracy to forecast "small to medium-level" changes in crime. The goal of all forecasting technique is not intrinsically crime forecasting. Instead, it incorporates crime forecasting techniques for safety metrics, in an attempt to provide to participating users a dynamic framework for safety awareness.[12]

VIII. LIMITATIONS OF FORECASTING TOOLS - RESEARCH CHALLENGES

Along with the benefits of using social media tools for forecasting, there are several limitations. These include:

1. Time needed to use these tools
2. Primarily used as marketing tools
3. Posted information available for all to see
4. Public safety security policies

IX. CONCLUSION

New emerging technologies are playing an increasingly crucial role in the daily work of frontline police officers, equipping them with enforcement and investigative tools that have the potential to make them better informed and more effective.

Law enforcement use of computer technology has expanded substantially over the past two decades. Given the increased power and diminishing costs of technology, the extensive growth in mobile communications infrastructure, and the expansion of innovative applications available, computer usage continues to increase in law enforcement agencies throughout the World.

The above mentioned challenges outlined in this study must have made a **Gateway for the Research Scholars.**

REFERENCES

[1] Using Social Insights for making our cities safer - DATAQUEST, THE BUSINESS OF INFOTECH VOLXXXI NO 20.

[2]"Fact Sheet," http://newsroom.fb.com/content/default.aspx?NewsAreaId=22.

[3] "Press Room," accessed Dec 4, 2013, www.youtube.com/t/press.

[4] "What is LinkedIn?" http://learn.linkedin.com/what-is-linkedin/.

[5] "LinkedIn," Wikipedia, last modified April 19, 2012, http://en.wikipedia.org/wiki/LinkedIn.

[6] "Twitter," Wikipedia, last modified April 27, 2012, http://en.wikipedia.org/wiki/Twitter."

[7] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to Data Mining*. Addison Wesley, 1 edition, May 2005.

[8] NLTK Project. Natural Language Toolkit. http://nltk.org/.

[9] Towards Safe Cities: A Mobile and Social Networking Approach Jaime Ballesteros, Bogdan Carbunar, Mahmudur Rahman, Naphtali Rishe, S.S. Iyengar

[9.1]http://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average

[10]http://en.wikipedia.org/wiki/Exponential_smoothing#Background.

[11]http://highered.mcgraw-hill.com/sites/dl/free/0073525251/886181/stevenson11e_sample_ch03.pdf.

[12]http://users.cis.fiu.edu/~iyengar/publication/UnderReview/J-(2013) -Towards Safe Cities - A Mobile and Social Networking Approach - IEEE-TPDS.pdf

[13] http://users.cis.fiu.edu/~carbunar/safety.pdf

