Australian Journal of Science and Technology

ISSN Number (2208-6404)
Volume 3; Issue 4; December 2019

Original Article

A numerical compilation model for accuracy estimations on accident detection based prototypes: Perspective of a singular vehicular *ad hoc* network

Julkar Nayeen Mahi[1]*, Redwanul Islam[2]

[1]Department of Computer Science and Engineering, City University, Khagan, Ashulia, Dhaka -1216, Bangladesh, [2]Department of Physiotherapy, Bangladesh Health Professions Institute, Faculty of Medicine, University of Dhaka, Dhaka - 1000, Bangladesh

ABSTRACT

A small-scale controller module for vehicle's in particular wireless transceiver (nRF) can cleverly recognize mishaps in a solitary vehicular *ad hoc* network (VANET). In this work, the authors have proposed a realistic mathematical model arrangement for mischance identification phenomena and recommend the modeling framework for attaining proper accuracy under vehicular *ad hoc* communication mediums. Although, each VANET communication process relies on close by system working criteria or system manipulation designs, they frequently discover trouble in communication preparing to bring down system exactness under specific conditions. This paper depicts a novel scientific thought of achieving good approximated accuracy over smaller scale implanted systems on the social event information. These numerical or statistical (e.g., watershed models) measures work just in specific situations and does not bolster distinctive situations of system work handling. The fundamental insight of this work is, if the irregular data or information (e.g., large data set) underpins normal cumulative distribution they are polynomial in nature and if those are polynomial (e.g., Taylor expansion) they can be taken care of through linear regression for the best estimation of data accuracy and system reliability.

Keywords: Force sensing resistors sensor, global system for mobile communications and general packet radio service module, microcontroller module (Arduino Mega 2560), model fit, nRF, Universal Serial Bus nRF, vehicular *ad hoc* network

Submitted: 27-10-2019, **Accepted:** 11-11-2019, **Published:** 29-12-2019

INTRODUCTION

Vehicular *ad hoc* network (VANET) is a trendy expression of the present world. VANET involves a few versatile hubs and it is by itself an extraordinary sort of mobile *ad hoc* network that can work in decentralized, self-arranging or multi hops routing pattern. VANET combined with a few tiny wireless sensor hubs or microcontrollers that enable cars to connect inside a given particular range.[1] VANET can act in unfamiliar situations such as – securities and safeguards, potential hazardous conditions (e.g., mishap, fire, and traffic jams), and so on. VANET possesses *ad hoc* on-demand vector, Dynamic Source Routing, greedy-face-greedy, greedy other adaptive face routing, greedy perimeter stateless routing, and individual other routing component. Evidently, it keeps up a part of unconstrained chip-based systems administration proportion through inter-vehicle communication and release the pressure from Roadside Units (RSU's) or cellular networks.[2] Dynamic changes in topology are the key piece of VANET[3] and therefore optimization, imperatives, or constraints are the fundamental research issue in the eye of these day's analysts.

VANET's fundamental motivation behind work is to build up a remote network for sending and accepting information or data thinking about vehicles. As of late, there is a change in outlook in vehicular accident detection. To diminish the workload of human administrators, sensor and detector based computerized mishap monitoring frameworks[4] are progressively mainstream. These frameworks are to a great extent reliant on global system for mobile communications

Address for correspondence: Julkar Nayeen Mahi, Department of Computer Science and Engineering, City University, Khagan, Ashulia, Dhaka - 1216, Bangladesh. E-mail: mahi.1992@gmail.com

(GSM), general packet radio service (GPRS), and satellite navigation advancements. GSM, GPRS, and radio-frequency (RF) advancements[5] have singular advantages for distinctive purposes. For GSM-based communication, we need to maintain versatile administrator based cost systems. GSM additionally needs the confirmation of mobile administrators to build up communication. In any case, this happens delay for the particular demonstrating ground situation. The GSM communication design dependably depends on satellite routing frameworks.

In addition to, GPRS which is a redesign of GSM innovation,usually used for higher communication speeds while having a good deal of advantages. Both GSM and GPRS are reasonable for good communication design more than RF[6] however in, the event that where tower cannot be embedded, those communication setups are somewhat sketchy despite the fact that this setup costs more contrasted with the RF communication[7] technology. Being mainstream in the present patterns, these current advances have signal drop propensity and poor system[8] communication issues. An extreme spending plan is likewise an issue for low asset nations.

Whereas, RF innovation can be converged with internet of things (IoT) or other embedded stages. A portion of the advantages of RF navigations are that these communication examples can be added to GSM, GPRS based framework when there is a need of it. RF advancements can be off in the different employments of machine to machine (M2M) communications.[9] Besides, random access channel scheduling, optimal or suboptimal technique,[10] we can reap the energy of RF modules with the assistance of GSM. In various parts bursty requests, expanded the range over GSM, lessen gadget cost, associating with broadened gadgets, RF map-reading module[10] demonstrates excellent outcomes. RF can be utilized as a part of GSM, GPRS or LTE upgrades if there should arise an occurrence of M2M interchanging communication.[11] For a minimal effort module, RF dependably a superior decision for certain careful framework improvement.

Considering these issues in our past examination, a minimal effort mishap discovery system is made for single VANET.[12]

In this work, a reliable scientific modeling is made for single VANET based property[13] which can be determined through the estimation of exactness on assembling information[14] or data from the force detection based resistor values. Existing advances are affected by moderate reaction time of GPRS,[15] satellite routing, signal communication insufficiencies, and absence of accuracy for utilizing long distant equipment, while in our approach[16] the novel numerical thought proposes a reliable direction to model fitness both in accuracy and communication pattern on a vehicular network environment.

The paper is classified into seven segments individually. To start with is the introduction, the second is method, third is experimental results, fourth is work process mind map of the proposed modeling, fifth is related works, sixth is limitations and future works, and toward the end the seventh part is conclusion.

METHODS

Experimental Hypothesis

The general property of this work depends on for the most part two occasion premise. In this work, the null hypothesis is – probability[17] of happening no mishap, that is, server Arduino does not send data to the neighborhood server, regardless of whether the accident detection happens or not.

The alternative hypothesis is – probability of mishap event, that is, Arduino server and client both send the data, and the nearby server[18] accumulates the prepared data if the mishap happens or not.

Sampling Technique

The random determination method is utilized to show that the samples (weights, ranges: 10–1000 g; pieces, $n = 20$) are secured over the sensor distance across to acquire the signal value within the framework.[19-21]

Defining Variables

In this work, the dependent variable is – response time in minutes and the independent variable is – mischance (sorted as a false alarm and detection. The false alarm implies that collision does not prompt mischance. On the other hand, detection suggests when a crash prompts an accident, over the sensor width).

Statistical Analysis

At first, assumptions of normality for the continuous variable (e.g., response time) are checked utilizing Kolmogorov–Smirnov test with Lilliefors correction. Likewise, independent sample t-test is utilized to test the hypothesis.[22] Besides, to check inequality, Lorenz's curve is used. Every single numerical test and analysis is performed utilizing the Statistical Package for the Social Sciences (IBM SPSS Statistics for Windows, Version 19.0. Armonk, NY: IBM Corp) and Matrix Laboratory Software Package (MATLAB for Windows, Version R2013 a. 3 Apple Hill Drive, Natick, MA: The MathWorks, Inc.) where necessary.[23] All P-values reported are based on two-tailed comparisons, where the most relevant and statistical level of significance are set at alpha 5% ($P < 0.05$).

EXPERIMENTAL RESULTS

The dependent variable (time) Y follows normality (normality approximation through cumulative distribution and Lorenz's curve inequality – Figure 1 (P = 0.000). Here, an alternative hypothesis is acknowledged in view of the independent t-test (P = 0.005). An examination demonstrates that our dynamic approach works viably for normality approximation,[24] that is, the reason the analysis is done in cumulative distribution to show approximated value in the normal distribution for attaining better reliability. In view of normality assumptions following equations are formulated within such examples.

$$\sum_0^n \{Sensor\}_{(high)} = \sum_0^n \{Vehicle\}_{(id)} + \sum_0^n \{Sensor\ Resistance\}_{(low)} + \sum_0^n \{Encryption\,Value\} \quad (1)$$

$$\sum_0^n \{Sensor\}_{(low)} = \sum_0^n \{Vehicle\}_{(id)} + \sum_0^n \{Sensor\ Resistance\}_{(high)} + \sum_0^n \{Encryption\,Value\} \quad (2)$$

$$\sum_0^n \{Vanet\ Server\}_{(output)} = \sum_0^n \{Sensor\}_{(high)} + \sum_0^n \{Vehicle\}_{(id)} + \sum_0^n \{Encrypt\}_{(output)} \quad (3)$$

$$\sum_0^n \{Vanet\ Server\}_{(output)} = \sum_0^n \{Sensor\}_{(low)} + \sum_0^n \{Vehicle\}_{(id)} + \sum_0^n \{Encrypt\}_{(output)} \quad (4)$$

Let us consider, for a probability density function (PDF) with finite mean = 0 and standard deviation (SD) σ = 1 the formula is,

$$f(x) = \frac{1}{2\pi} e^{-x^2/2}$$

The cumulative distribution function is an anti-derivative of PDF or PDF. Thus, the apparent formula can be written as –

$$f(x) = 1/2\pi \int_{-\infty}^{x} e^{-t/2}\ dt$$

For, PDF the formula relies on

$$(x) = \int_{-\infty}^{\infty} P(x)\ dx = 1$$

On this basis, the empirical property of our work relies on –

$$P(0 \le x \le 8) = \int_0^8 f(x)\,dx = P(x)\,|_0^8$$

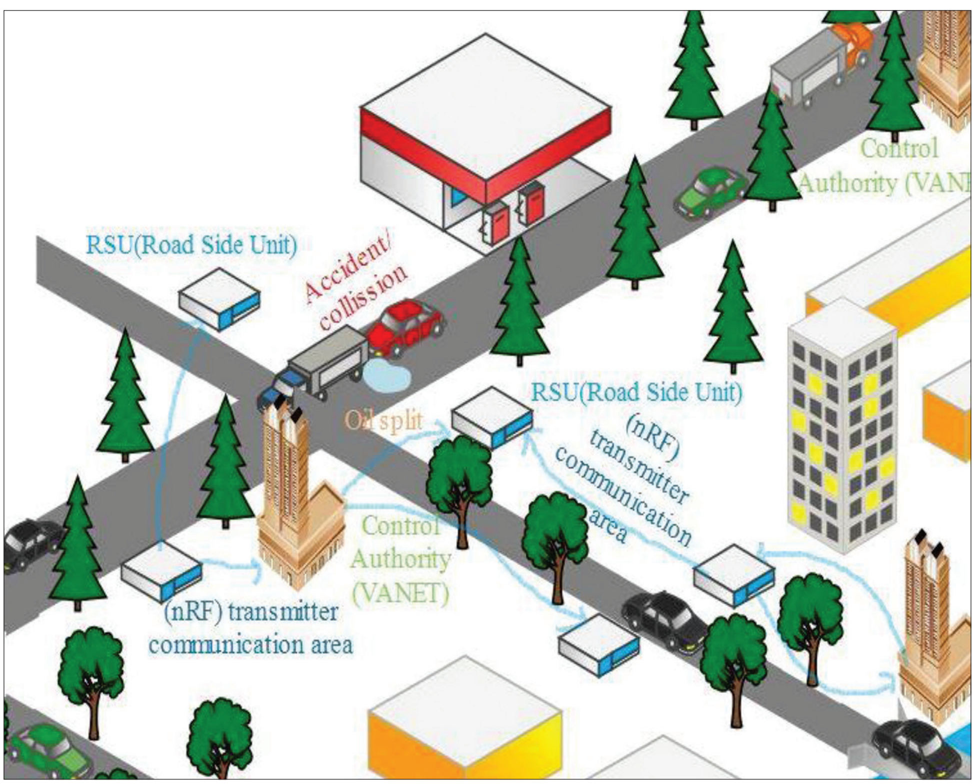

Figure 1: A typical vehicular *ad hoc* network communication pattern

$$= F(8) - F(0) = P(x \le 8) - P(x \le 0)$$

(Where, 0 and 8 is redirected as response time in the server in minutes).

Therefore, the probabilistic property simplifies on the range for the specific system about,

$$P(x) = \int_{-\infty}^{x} P(t)\,dt \begin{cases} 0\, if\ x < 0 \\ \dfrac{x}{8}\, if\ 0 \le x \le 8 \\ 1\, if\ x > 8 \end{cases} \quad (5)$$

In Figure 2, the X-axis cumulative % of false alarm probability, Y-axis cumulative % of detection probability, and in the Z-axis cumulative % of response time in minutes are appeared. In the three-dimensional (3D) graph – Lorenz inequality and cumulative normality approximation support the tendency to keep the dataset in linear regression of maximum – likelihood estimators even the population size is changed. The maximum – likelihood estimators always require an assumption of full distribution. In our situation, the arbitrary errors take after a normal distribution.

In our work, we initialize the system accuracy up to 0.5–1 probabilistic values. For accuracy, we can define normally as,

Accuracy = (Correctly predicted class/Total testing class)×100%

Total accuracy of the system = 15/20×100%=75%

This is only a standard-essential calculation, yet it cannot demonstrate the unadulterated estimation of framework[25] accuracy or reliability. An accuracy is a kind of estimation that characterizes how close an estimation is contrasted

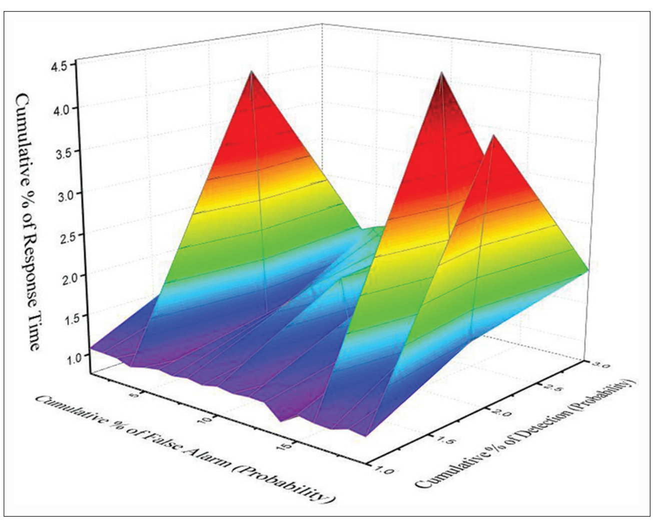

Figure 2: Dataset inequality tends to the normal distribution

with the genuine estimation amount of a specific entity (e.g., framework, machines), and so on. In this situation, the accuracy of our framework accomplishes both arbitrary or random and deliberate or systematic errors.

In systematic error, the property relies on –

i. Effects on the system in blunders about poor information procurement
ii. Movement consequences of the framework brought about a blunder and yields that are blended with fake information
iii. Weight factor blunder of tests over the sensor distance across for having a smaller region to execute yields.

As the above depictions told about a specialized model prototyping for reliable communication, in that case, systematic error is an exceptionally concerned issue for accuracy. Meanwhile, the framework information[26] is assembled by fluctuating arbitrary measuring yields. In this relative case here, random error is additionally thought about the issue. Since the information assembled from our demonstrated framework support standard distribution, the framework has achieved normality in both false alarms and in detection case. In any case, this is a trial experimental[27] research work. In an experiment, unique estimations are driven by unusual changes and random error happens. The progressions are made because of various ecological conditions; they are –

i. For estimating the watched value if the framework[28] deciphers a value continuously, it cannot recognize for misleading of values
ii. Irregular changes in data acquisition on system model movements.

As the data-acquisition property follows normal or Gaussian distribution, it can be said that the system has a random error in a good proportion. The framework achieves the likelihood estimation by changing times and detection criteria of the mishap where the incorrect (false) alarm demonstrates 00, and detection indicates 11 for output view to the light-emitting diode display. This output states the minimum and maximum values of the independent variable X with respect to the dependent variable, time that is Y. As indicated by the SD which is got from the data index is 0.503, standard error of mean 0.112, variance is 0.253. Whereas, false alarm and detection frequency are 8 and 12, respectively. The tested samples are 20. Since from the SD it can be seen that the measured SD value is 0.503 which means; there is 50% chance of uncertainty with respect to the framework accuracy to get detection individually. On the off chance that the SD value stays nearer with the mean value for detection (variable) 0.2000 and response time (variable) 1.8350, at that point one might say that the framework achieves more accuracy. Maybe, our cases are diverse, as indicated by the measured data index. The individual SD of variables detection and response time is 1.00525 and 1.08967, respectively, somewhat a long way from the mean.

The standard error estimation is $s \div \sqrt{n} = 0.0503 \div \sqrt{20} = 0.0112$; from the Cronbach alpha analysis, we got the value 0.707; that implies the standard Cronbach alpha on items are 0.707. The variables, detection, and response time are considered items. From the Cronbach alpha value, it has been seen that the framework has been reliable for recognizing events yet not as exact as a proper model.[29] The probability value like 0.707 is characterized as adequate if there should be an occurrence of reliability in quality of framework statistics.

For a systemic random error, the formula can be stated as –

$$\begin{aligned} S &= \sqrt{1 \div (n-1)}\ \Sigma\ (x_i - \bar{x}) \\ &= \sqrt{1 \div (20-1)}\ \Sigma\ (1.13 - 0.0565)^2 \\ &= 0.246 \\ &= 0.25 \\ &= 25\%\ \text{(approximately)} \end{aligned}$$

The random error is an error of fast vacillations of information assembling inside a framework. The error can be minimized if the population size increments, while the systematic error defines the operation error of a procedure.[30] The operation error essentially demonstrates the framework inefficiency under a certain amount of process criteria or properties.

From the root mean square error (RMSE), it is seen that RMSE likewise utilized in deciding the model exactness. From our model, we get by inferring the RMSE or RMSD value from X and Y axis is Ŷ (Ŷ is the predicted value in regression). The $R^2 = 0.372$; adjusted R squared is 0.337, standard error of estimate 0.887 (mean square error). The mean square at the regression line is 8.386, residual mean square is 0.787, and the RMSE is 0.942. Consequently, our samples are just 20 so this is not a decent measurable forecast in this phenomenon for claiming decent accuracy in information procurement. In this situation, 0.94 RMSE esteem is not guaranteed to be a flat out accuracy estimation or fit value for the model.[31]

Assessing the regulatory and statistical returns (RSR) values, we conclude that, RSR accuracy of model fit on our work event is –

$$RSR = RMSE \div STDEV_{obs}$$

$$\begin{aligned} &[\sqrt{\sum_{i=1}^{n}\left(Y_i\,obs - Y_i\,sim\right)^2}] \div [\sqrt{\sum_{i=1}^{n}\left(Y_i\,obs - Y\,mean\right)^2}] \\ &= \sqrt{\{20(22.5204-5)^2\}} \div \sqrt{\{20(22.5204-1.13)^2\}} \\ &= 78.35 \div 95.70 \\ &= 0.81 \\ &= 81\%\ \text{(approximately)} \end{aligned}$$

Where, *obs* is characterized as the total of the observed value from the informational collection; *sim* is defined as the simulated value (e.g., total response time) and mean are characterized as the total mean value of the samples inside the dataset. From the RSR calculation, we see that, it works[32] exceptionally well when the sample size is smaller. For large samples, RSR estimation will not have this sort of accuracy for model fitness. From Nash-Sutcliffe model fitness and our framework information we can determine that,

Efficiency,

$$\begin{aligned} E &= 1 - [\{\sum_{i=1}^{n}\left(X_i\,obs - X_i\,model\right)^2\} \div \\ &\quad \{\sum_{i=1}^{n}\left(X_i\,obs - X_{mean}\,obs\right)^2 \\ &= 1 - [\{20(22.5204-5)^2\} \div \{20(22.5204-1.13)^2\}] \\ &= 1 - [\{20(307)\} \div \{20(458)\}] \\ &= 1 - [6140 \div 9160] \\ &= 1 - 0.670 \\ &= 0.33 \\ &= 33\%\ \text{error (approximately)} \end{aligned}$$

In this estimation, as far as possible ranges from $-\infty$ to 1. It is certainly a decent measurement in the probability of statistics. This efficiency estimation does not rely on extensive or little sample sizes, yet clearly demonstrates a decent indicator that exists in sample sizes. Here, the exactness lies at 77% generally. The Nash-Sutcliffe model proficiency portrays how exactly the model[33] is looking at information acquisitions.

A productivity of 0 or 0.33 is a long way behind 1 of every a probability value. For 1 or closer to 1, it depicts the model fitness about an ideal match between the model and perceptions.[34] Although, an effectiveness of 0 or 0.33 shows the model depiction about expectations; those are as relative and precise as the mean of the observed data. Negative efficiency bodes well when observed mean is a superior predictor than the model.

From the percent bias (PBIAS) calculation we get that,

PBIAS efficiency of accuracy

$$\begin{aligned} PBIAS &= [\{\sum_{i=1}^{n}\left(Y_i\,obs - Y_i\,sim\right)*\left(100\right)\} \div \{\sum_{i=1}^{n}\left(Y_i\,obs\right)\}] \\ &= [\{20*(22.5204-5)*(100)\} \div \{20*(22.5204)\}] \\ &= 17.5204 \div 22.5204 \\ &= 0.7779 \\ &= 77.80\%\ \text{(approximately)} \end{aligned}$$

The precise estimations to check the wellness of a VANET[35] framework modeling varies under different calculation strategy determinations. The over four statistical equations imply the exactness of model fit through our framework which depends on 75–81% in an approximation. That is the reason, not only a basic computation characterizes a decent precision but also a decent statistical measure works exceptionally well to characterize legitimate accuracy. For little samples, RMSE shows and supports maximum guess of information transmitting accuracy. For this situation we select, the RSR estimate is a decent, accurate estimation for this model fit foundation. However,

Table 1: Data gathered from accident detection and false alarm phenomenon

X (resistance value)	Y time (approx.)	XY (multiplicative value)	X^2 (square value of X)	Y^2 (square value of Y)
1.071	1.6 (2.40 min)	1.7136	1.14	2.56
1.074	1.82	1.9546	1.15	3.31
1.0675	1.53	1.6332	1.13	2.34
1.058	0.87	0.920	1.11	0.75
1.001	4.25	4.254	1.00	18.06
1.0478	0.86 (1.26 min)	0.900	1.09	0.737
1.0789	1.35	1.4565	1.16	1.82
1.0542	1.50	1.5834	1.11	2.256
1.0845	1.69	1.8349	1.17	2.862
1.0561	1.29	1.3718	1.11	1.687
1.0974	1.89	2.0740	1.20	3.57
1.069	0.9	0.9621	1.14	0.81
1.089	1.70	1.8523	1.18	2.893
1.082	1.34	1.4520	1.17	1.800
0.9147	0.76	0.6960	0.83	0.579
1.0562	1.63	1.7194	1.11	2.650
1.0235	4.52 (7.32 min)	4.529	1.00	20.51
1.013	1.82	1.8436	1.02	3.31
1.0791	[1]1.47 (2.27 min)	1.5959	1.16	2.187
1.0529	3.92	3.921	1.00	15.37
ΣX=20.926	ΣY=36.743	ΣXY=38.267	$ΣX^2$=20.81	$ΣY^2$=90.06

Source: Gathered data, from test model scenario (academic testing purpose only)

for large samples, the RSR may not be a superior predictor for characterizing accuracy in a model fit of this type of framework. In our proposed framework modeling shows, we utilize Taylor linear approximation for the estimation of framework model fitness. Taylor's approximation is a decent paradigm for determining the model wellness in case of substantial continuing samples while checking or estimating measurements for accuracy. The outlines are delineated in the "Appendix" segment.[11]

Defining Samples for Standard Accuracy under Linear Taylor Polynomial Approximation

A linear regression formulates an equation, which is like Y = aX+b, where X is the independent variable, and Y is the dependent variable. The slope of the line or tangent is b and a, defines the intercept that means the value of Y when the value of X is 0. Table 1 illustrates the data set we have gotten from a testbed module. The property fulfills the reason to choose linear regression as an analytical modeling. We have defined the variable X in two category options, and Y is interpreted as a fully dependent variable for the phenomena. This also defines the modeled sector in a statistical equation that can tell the overall criteria for large distribution of a dataset if it follows or not.

1 (N.B): The authors take the "time" (approximate) as a numeric decimal value to make the probability calculation easier on log chart.

From the dataset we get the value of

$$a = \{(\Sigma\, Y)*(\Sigma\, X^2) - (\Sigma\, X)*(\Sigma\, XY)\} \div \{n*(\Sigma\, X^2) - (\Sigma\, X)^2\}$$
$$= \{(764.621 - 800.8121)\} \div \{(416.2 - 437.930)\}$$
$$= 36.1911 \div 21.73$$
$$= 1.6654$$

And,

$$b = \{n*(\Sigma\, XY) - (\Sigma\, X) * (\Sigma\, Y)\} \div \{n*(\Sigma\, X^2) - (\Sigma\, X)^2\}$$
$$= \{(765.346 - 768.913)\} \div \{(416.2) - (437.930)\}$$
$$= \{3.567\} \div \{21.73\}$$
$$= 0.164$$

After evaluating the linear regression intercept and slope value, we get the following equation that the system supports is –

$$Y = aX+b$$
$$Y = 1.6654X+0.164 \quad (6)$$

We get an equation from our simulated values that have been stated above, which can be defined as –

$$Y = 1.6654X+0.164, \text{ where } a = 1.6654;\ b = 0.164$$

And the above-stated equation is for a linear regression property. Varying this equation our system model supports linear regression.

Nevertheless, a linear approximation of values is expected to indicate better accuracy in framework model wellness. In Figure 3, the log chart information[36] speaks to that in intense response time; there happened a genuine mishap of vehicles, which additionally implies that high measures of force are infused over the sensor width and the bar descends in scale ranges on probability measurements. On the side of detection, the amounts are appearing in numerical measures (samples) that involve the data communication pattern across the sensor and neighborhood server processes. In that circumstance, inferior probability redirects fewer casualties. In other words, it can be informed that fewer casualties meant a lower force infusion as well as the rising nature of bars ranges from zero to one in likelihood measures on the chart. The regression line supports round trip or gross probabilistic measurement property, which counterparts the sensor-based framework modules.[37] In reality, the regression line diverts top to bottom of general strategies. In spite of the fact that, a downfall in the bars demonstrates a better detection probability agreeing than the reciprocal force sensor molding activity as well as extravagantly accident occurrences with a high response time at the nearby server backbone.

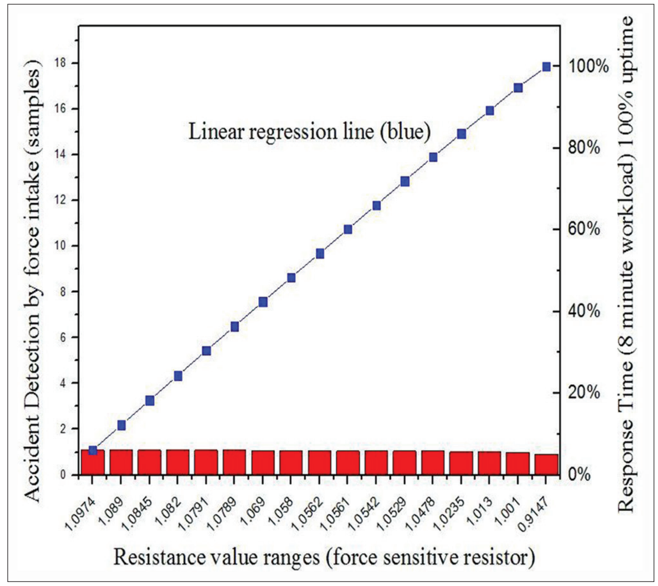

Figure 3: Accident detection property (logarithm plot for 20 samples)

In Figure 4, the linear curve fitting graph shows that the detection of data residuals and outliers (probability) with respect to response time (minutes) is measured in addition to the assumed regression line sighted in the middle. The blue concatenated data points (blue) demonstrate the random nature of data set considering response time (minutes) in the even states of information control. The linear curve fitting line for 0 shows the approach of attaining maximum accuracy under Taylor polynomial approximation modeling. Figure 4 states that maximum data accuracy can attain through this modeling.[38] Concatenated data also show the maximum outcomes in probability metrics under linear fitting curve statistics. This also redirects our above experimental hypothesis as true for this type of prediction approach.

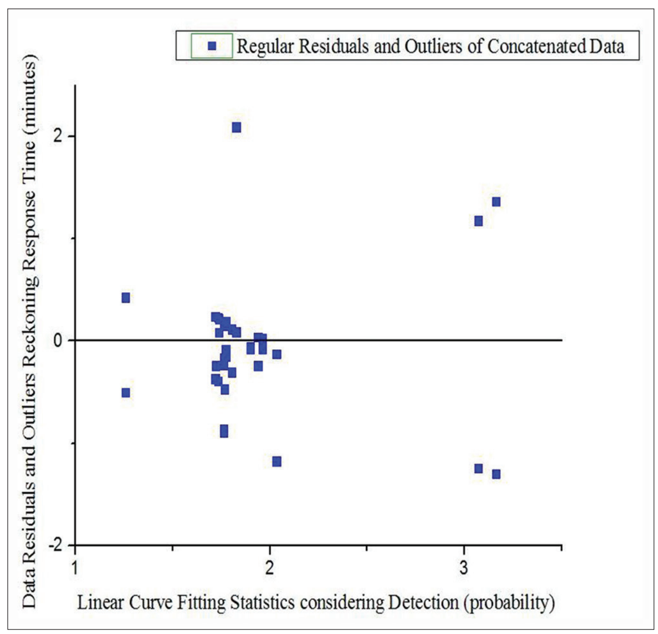

Figure 4: Linear curve fitting statistics (Dealing with detection and response time of data points)

Figure 5 depicts polynomial linear approximations considering large data sets (orange) measured through detection (probability)[39] with respect to response time (minutes). The illustration above shows approximations tends to very similar to the odd and even pairs of data points. We can see that the odd interval polynomials of data 3, 7, and 9 in order have very similar position changes within the distribution fitness. The even ones 10, 16 orders are also following the same. Thus, we can conclude that polynomial rule can check the above data distribution in the way of utmost approximation model fitness. In addition to the measurement which is taken from 0 to 45th polynomial series of the data points also follows the linear regressive estimation rule. On an assuming basis, in terms of probability we can see that maximum residual data outstretched within the linear regression line on 0 in a relative closely coupled manner. Moreover, the data points accompany the ranges between 0 and 1 measuring probability metrics. The data points alongside with the regression line support our hypothesis as true for the proposed scientific modeling.

In Figure 6, a lag graph of the dataset is shown. The lag graph usually supports the criteria of being a distribution of dataset is in normal or randomly distributed in the form. From the distribution, we can see that detection (probability) with respect to response time (minutes) supports random in nature, and most of the residual data points range from 0 to 1 probability values.

From the dataset, we can also see that data points (magenta) are closely coupled and have relatively similar position changes within ranges from 0.7 to 0.9 and almost near about 1. This redirects that data point's follows random or normal distribution and follows polynomial ordered approximation. Moreover, then the data points accompany the ranges between 0 and 1 measuring probability metrics. Hence, the data points alongside with the regression line support polynomial approximation fitting as well as our established hypothesis as true for single VANET based prediction approaches to attain maximum system modeling[40] accuracy.

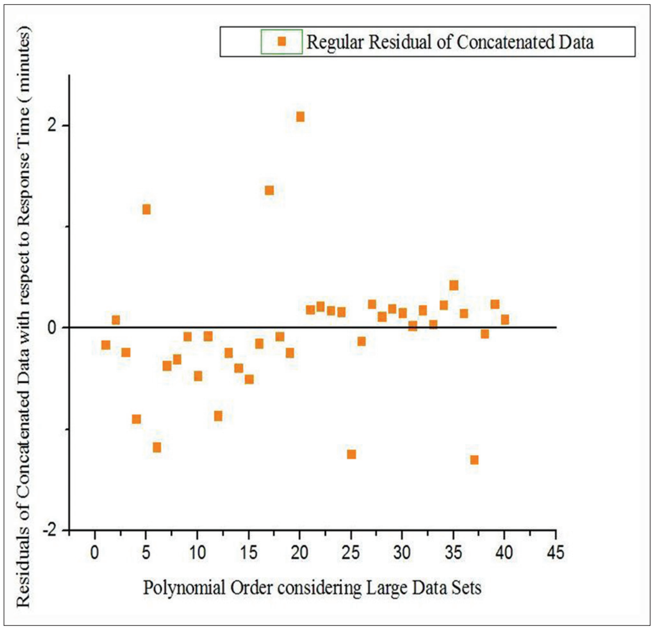

Figure 5: Polynomial orders of datasets (Related with large distribution of data points)

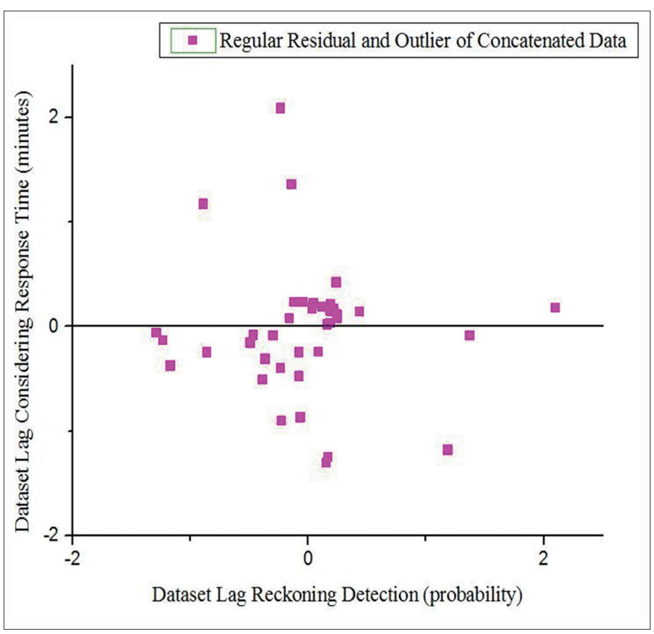

Figure 6: Lag graph test distribution of datasets (Dealing with response time, regression line, and detection of data points)

Figure 7 portrays independent variable detection (probability) with respect to dependent variable response time (minutes) in probability metrics. The illustration above shows approximations tend to very similar on the odd and even pairs of data points (dark green). The concatenated residual data points follow regressive polynomial approximation in a probability metric from 0 to 1, except for some outliers. The maximum probability ranges from 0.9 to 1. We can also see that proliferated data points in 0, the assuming linear regression line supports maximum attainable accuracy for the system modeling. Most of the concatenated residual data points rely on the determined probability ranges for linear curve fitting statistics. Furthermore, the data points alongside the regression line support polynomial approximation fitting also as well as our established hypothesis as true for the proposed scientific modeling which claims to attain system[41] accuracy at a maximum level.

For example, if we supposed that for a parabolic polynomial equation where,

$f(x) = x^2 – 5x + 7$, and that is to be linearized at $x = 5$

Therefore, $f(5) = 5^2 – 5 \times 5 + 7$
$= 25 – 25 + 7$
$= 7$

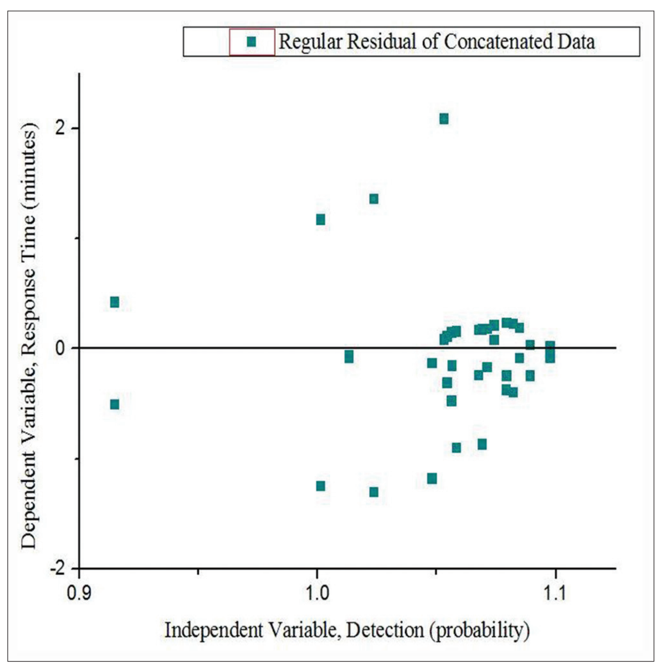

Figure 7: Probability of approximations in accuracy (Dealing with response time, regression line, detection, and polynomial relation of concatenated data points)

$f'(x) = 2x–5$
$= 2×5–5$
$= 5$

Therefore, $f(x) = f(5)+f'(5)(x–5)$
$= 7+5x–25$
$= 5x–18$

Therefore, $g(x) = 5x–18$

If x is nearer at a point a, where a = 4, then g (x) = 2. Thus, x = 2 is the point when the expression can have its tangent or it can be said that $g(x) = 5x–18$ is the equation of tangent to curve $f(x) = x^2–5x+7$ at x = 5.

From the prototypes, we have got the dataset. If a particular data from the dataset can support parabolic polynomial function rules then we may say that the performance module of the prototypes does rely on linear regression and, that is, the point from which we define accuracy in our proposed work.

Suppose that $f(x) = x^2–5x+7$ is to be linearized at x = 1.502 instead of 1.0478 which is a value of X (from the gathered dataset).

Therefore, $f(1.502) = (1.502)^2–5×(1.502)+7$
$= 2.256004–7.51+7$
$= 1.746004$
$f'(x) = 2x–5$
$f'(1.502) = 2×(1.502)–5$
$= 1.004$
1.004 is closely related to 1.0478.

Therefore, $f(x) = f(1.502)+f'(1.502)(x–1.502)$
$= 1.746004+1.004x–2.25$
$= −0.51+1.004x$
$= 1.004x–0.51$

Therefore, $g(x) = 1.004x–0.51$

Then, it can be said that the expression can have its tangent at $g(x) = 1.004x–0.51$ is the equation of tangent to curve, $f(x) = x^2–5x+7$ at x = 1.502

Viewing Figures 1-7 and the derivatives of the polynomial data points, it can be told that the fourth-order derivative also has the likely value of x = a measured from the second one at specific multiplicative orders. The odd derivatives also match relatively with the odd ones. Thus, we can state that most of the derivative values are relatively matched to the adjacent ones at a previous multiplicative order. Apart from it, whatever the value of x changes rapidly in the derivatives the result of gathering data in our proposed system will rely on the property of linear regression.

Work Process Mind Map of the Proposed Modeling

Figure 8 represents the work process mind guide of the proposed display. At to start with, the data assembled from the force sensor and bolstered into the nRF client terminal for the data confirmation process. However, after verification, the data are sent from the Arduino client terminal downright to the nRF terminal to the Arduino server.[5] The processed data are accumulated and verified for sending them to the RSU's Universal Serial Bus nRF module. Subsequent to preparing the data gathered from the nRF Arduino server on board the RSU performs adjustment of the information by making them institutionalized through encryption coupling and comparing. After the verification and comparing process in RSU on board unit the data are sent to the main authority server (MAS) for checking the encryption standard method.[41] After sending data into the MAS a procedure verification is made through providing regression, log base knowledge rule checking. After regression modeling, the information is embedded in the database to perform intermittent session-based record input or output task logs. Moreover, subsequent to finishing the procedure on MAS, the data are sent toward another server to check authenticity and reliability, which is called data authentication and system reliabilitymechanism (DASR) unit.[40] In DASR on board unit the system property (e.g., functional programmed server) is designed with some knowledge-based functional programming rules to perform data accuracy and reliability countermeasures. After check approval from the MAS, the data are embedded into the DASR on board unit to compare linear regression polynomial features. If the polynomial approximation rule is not verified properly the running system process then checks again for additional data authenticity and accuracy through our configured rule based validation and verification. Unsupported data[39] are marked as an outlier and sent to the database log through issuing less relation to concatenated conjugate data. If the polynomial approximation of maximum accuracy rule is supported, then the data are marked as residuals and send to the database log issuing maximum relation compared to the concatenated conjugate data.[22] In relation to both unsupported and supported rule-based data for accuracy measures, reliable percentile-based log files are made in the database of DASR after approximation counter measures and sent it back to the MAS for observing error and usable benchmarks to take better exactness endeavors for vehicular communication pattern.

Related Works

To the best of our knowledge, this airbag, pervasive fall detection, cellular detection, on star native systems, and TEDAS are the systems that have been made so far. Likewise, airbag system[17] is a welfare guarded frontal; side attached system in the vehicle having a low-cost higher reliability and an eject based functional system with a view to showing the capacity to slow down the internal object speed to zero, whereas it does not have the proper acquisition system to

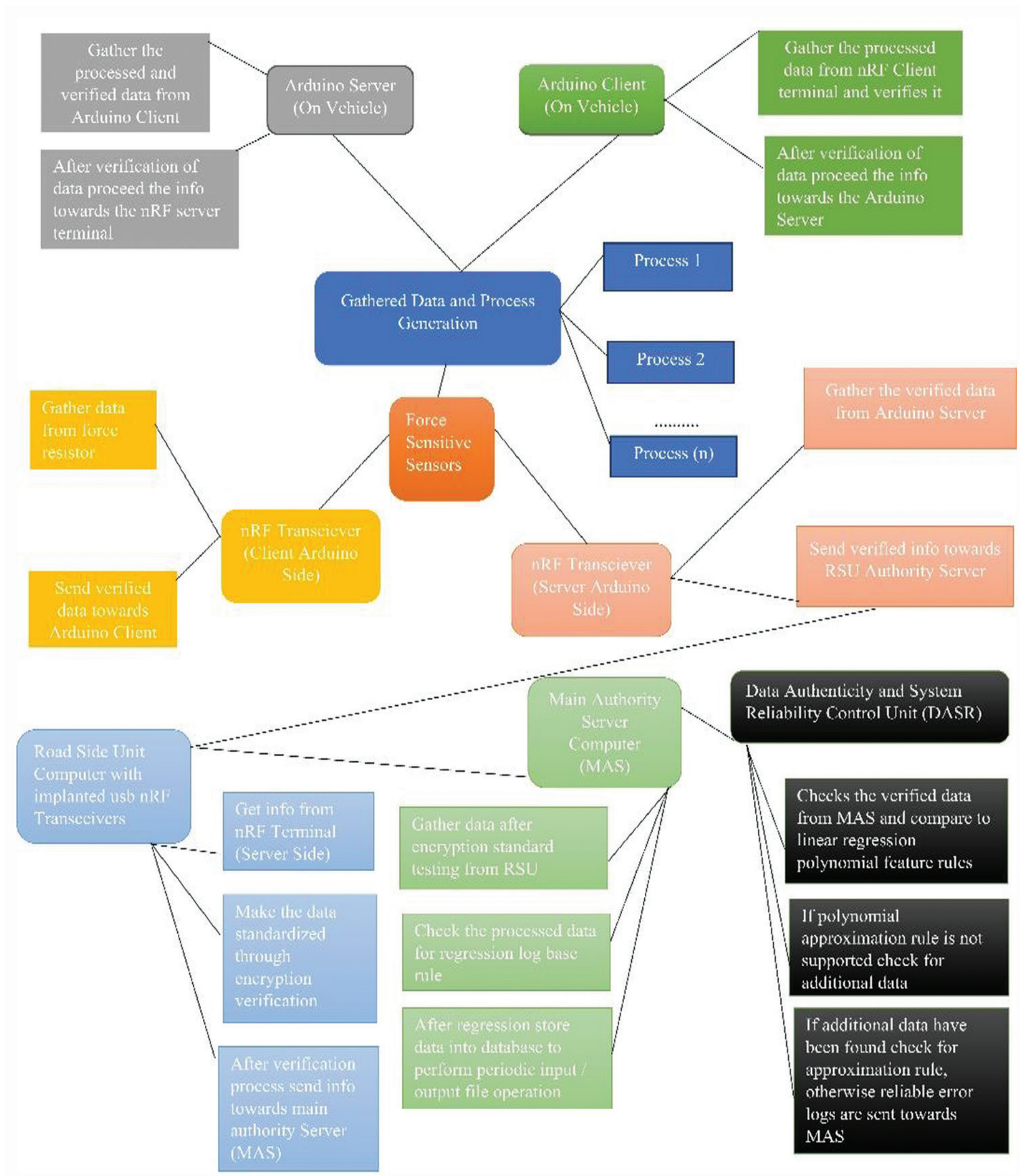

Figure 8: Workflow mind map of the process description modeling

protect the vehicle from the right and left side. However, our proposed numerical modeling can detect accident with reliable significance in the real-time event occurrences. The approximation process of the modeling is a primary key for detecting events acutely and performs different rules of distributed data, which is a criterion of showing effective workflow of a model, but these above systems cannot identify event occurrence rules with accurate measures.

In pervasive fall detection,[18] a 3D approach is used to show the reading through an accelerometer by employing a specific threshold value, but in this case, the distance must have to be limited in comparing the base and the sensor. The finer distance can cause higher budget. In fact, the signal ranges to work at larger distances are a bit questionable while having no transmitter implanted. GSM and GPRS[12] can work fine over RF for energy harvesting and speedy communication, but GSM and GPRS cause a good amount of delay because of the far located satellites.[13] This far communication technique causes a good quantity of equipment and huge cost to thrive and work through its processes for the betterment of people's lives. Nevertheless, our proposed modeling has an effective technique of data distribution that performs the utmost machine readability which is a far-reaching consequence that fall detection systems do not have.

The mobile[11] based movable accident detection system must need an operating system and also good network coverage for fixed locations while our proposed numerical modeling property does not need a changeable grounded communication because it is a learning rule based mishap detection approach.

The on-star corporation native system[17,18] is a U.S.A based company, which works in the detection system. The major demerit of the on-star system is country, dependency, 1-time installment on vehicles, having more options such as the theft identifier and auto accident detection. None of these above described systems show proper accuracy for the accident detection phenomena. On the contrary, these properties show inferiority in service and also expensive because of the shortage of assistance and attachment of different items. Even so, our proposed mathematical modeling approach with reliability (e.g., accuracy) is a learning rule based modifier characteristic that shows outputs considering residual concatenated relationships with reusable data. The more related to data shows better results of accuracy that also redirects the proper event (e.g., accident) detection in a periodic fashion. Regardless of the fact that the on-star system is incompetent of accuracy estimation which does not support learning-based rule approximation of data management.

TEDAS: A twitter-based event detection and analysis system that works mainly on three event basics and they are – detecting new events, analyze the spatial, temporal pattern and identify the importance of events. Moreover, the system works with GSM, GPS, and satellite[31] navigated data interchanging through detecting new events and analyzing patterns.[42-44] The importance of events can be detected through twitter messages using internet protocol and longitude, latitude from GSM points. The system only works with a twitter-based approach and does not perform system accuracy on event detection. On the contrary, our proposed numerical modeling ensures better accuracy of a system model and also it defines machine learning-based functional knowledge rule property by which an operator can know about the system model lags. Seeing the tweets of a TEDAS system no operator can ensure which system fails or not. TEDAS also does not have a log file restoring database format which is a big issue for event detection based system models.

Limitations and Future Works

Our proposed work has some definitive constraints. The nRF transceivers at both ends run through microcontroller stages. The span of the nRF transceiver is low around 80–300 m on average.[6] The sensor is utilized as a part of a small diameter and embedded in the cap of the vehicle for identifying the pressure efficiently. Expanding the range of the nRF transceivers might be a decent alternative later on to get the most extreme scope.

The authors have utilized structure query language (MySQL) database server to guarantee information credibility or authenticity and accuracy, demonstrating in a functional programming measurement type. Using MySQL database MSA and DASR are additionally made. In any case, these servers are uniquely designed and go through Intel family chipsets and processors, not Xeon or Xeon-phi server processors. As a reason, the processing verification and functional polynomial approximation rule-based checking are a bit tedious, which comes about delays in making database log records as well as percentile measures for proper accuracy estimation that can be characterized as server process[25] slack.

The authors have also influenced an automated input, output log record deletion process through periodic everyday perception at particular circumstances. Hence, the model[16] can diminish the information redundancy issues, yet redundant information occurs delays in processor caches, which demonstrates confinements in custom execution modeling. Maybe up-degree of gadgets can guarantee better outcomes for this type work purpose.

Maneuvering MySQL as a database server is a long haul impediment of this work, but cloud alternatives might be superior criteria for these kinds of works[8] in the current future. The force resistor has some lag on handling values when counter measures are made subsequently. Lessening the data gathering lag endeavors can get a decent work impact on the near future.

The authors have tested the simulation criteria for this work considering 20 samples which are not an established choice to obtain desired accurate measures through these little samples modeling. Using more samples may change the polynomial data attributes in a different way still above simulations supports the approximation in accuracy process for large distribution, but a real-time analysis of large data will mark a distinguishable change in the future in the sector of data analytics.

The accuracy estimation model is made considering a solitary VANET controlled area, although it can be executed in various VANET[11] coupled region utilizing numerous client-server Arduino and computer models in a cloud, pervasive cloud computing, portable cloud, or ubiquitous fog computing distributions.

CONCLUSION

The accuracy modeling can able to predict detection (e.g., accident) information with the maximum reliability in a short time and can send the functional approximated data in polynomials to their intended providers (e.g., local servers and MAS) across a single VANET coupled area.

The overall design of numerical modeling is complex in a manner because it attains accuracy of a system model through considering, device materials, throughput, threshold values, observations, knowledge-based functional programming rules, and probability estimations. The proposed numerical model works pretty well on low range polynomial data, usually an ignorable rate. However, in our simulations, we can see that the model can able to work also well in high range polynomial disperse data as well as in distributed VANET coupled area. The preliminary posturing results of this work can be a fundamental asset or a good adjustment for IoT based machine learning, cloud-based machine learning having IoT models, and ubiquitous fog IoT based machine learning through knowledge-based rules.

Nomenclatures	
Obs.	**Observable value**
n	Number of samples
n	Number of the process toward the end of time or iterations
p	Level of significance for sample testing
S	Systematic random error
P	Probability of particular function
Y	Defines the dependent variable
X	Defines the independent variable
b	Defines the slope of the straight line or tangent
a	Defines the intercept against the slope
$\bar{x}$	Defines the mean of a specific value, x
sim	Defines the simulated value of a particular variable
Greek symbols	
σ	Standard deviation
Σ	Sum of the total value in terms of function determination
Abbreviations	
PDF	Probability density function
RSR	Regulatory and statistical returns
RMSE	Root mean square error
LED	Light-emitting diode

ACKNOWLEDGMENT

The authors gave their sincere gratitude and wish toward Zamshed Iqbal Chowdhury[2] and Khandoker Abdul Rahad[3], Shamim Al Mamun[4], respectively, for supporting them throughout the ongoing research work of this article.

2 Lecturer, Institute of Information Technology, Jahangirnagar University, Savar - Dhaka; 1342.

3 Teaching Assistant at Department of Computer Science, University of Texas El Paso (UTEP).

4 PhD Research Assistant, Department of Information and Computer Sciences, Saitama University, Saitama - Japan

REFERENCES

1. Barba CT, Mateos MA, Soto PR, Mezher AM, Igartua MA. Smart City for VANETs Using Warning Messages, Traffic Statistics and Intelligent Traffic Lights. In: 4th Intelligent Vehicles Symposium, IEEE; 2012. p. 902-7.
2. Barrachina J, Garrido P, Fogue M, Martinez FJ, Cano JC, Calafate CT, *et al*. VEACON: A vehicular accident ontology designed to improve safety on the roads. J Netw Comput Appl 2012;35:1891-900.
3. Benslimane A. Optimized Dissemination of Alarm Messages in Vehicular *ad hoc* Networks (VANET). In: IEEE International Conference on High Speed Networks and Multimedia Communications. Berlin, Heidelberg: Springer; 2004. p. 655-66.
4. Bißmeyer N, Petit J, Bayarou KM. CoPRA: Conditional Pseudonym Resolution Algorithm in VANETs. In: Wireless On-demand Network Systems and Services (WONS), 10th Annual Conference, IEEE; 2013. p. 9-16.
5. Campolo C, Iera A, Molinaro A, Paratore SY, Ruggeri G. SMaRTCaR: An Integrated Smartphone-Based Platform to Support Traffic Management Applications. In: Vehicular Traffic Management for Smart Cities (VTM), 2012 1st International Workshop, IEEE; 2012. p. 1-6.
6. Charles S. Nrf24l01 Real Life Range Test, Charless Blog; 2013. Available from: https://www.hallard.me/nrf24101-real-life-range- test.
7. Available from: https://www.hallard.me/nrf24l01-real-life-range-test. [Last accessed on 2013 Jul 06].
8. Gilles ER, Bellaïche M, Pierre S, Quintero A. VANET security surveys. Comput Commun 2014;44:1-13.
9. Fogue M, Garrido P, Martinez FJ, Cano JC, Calafate CT, Manzoni P. Automatic accident detection: Assistance through communication technologies and vehicles. IEEE Veh Technol Mag 2012;7:90-100.
10. Goh KN, Jaafar J, Mustapha EE, Goh ETE. Automatic Accident Location Detection System. In: Information and Communication Technologies, 4th World Congress, IEEE; 2014. p. 63-69.
11. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: A guide for non-statisticians. Int J Endocrinol Metab 2012;10:486.
12. Hartenstein H, Laberteaux KP. A tutorial survey on vehicular *ad hoc* networks. IEEE Commun Mag 2008;46:164.
13. Houghton-Carr H. Hydrological Instrumentation: Preliminary Review of GSM/GPRS-based Telemetry Systems for Real-Time Data Acquisition. India Hydrology Project Phase II, Technical Assistance (Implementation Support) and Management Consultancy. Internal Project Report; 2009.
14. Domingo MC, Reyes A. A clean slate architecture design for VANETs. Wirel Pers Commun 2012;67:315-33.
15. Lakas A, Cheqfah M. Detection and Dissipation of Road Traffic Congestion Using Vehicular Communication. In: Microwave Symposium, Mediterrannean, IEEE; 2009. p. 1-6.
16. Rahman MT, Mahi MJ, Biswas M, Kaiser MS, Al Mamun S. Performance Evaluation of a Portable PABX System through Developing New Bandwidth Optimization Technique. In:

Electrical Engineering and Information Communication Technology, International Conference, IEEE; 2015. p. 1-5.
17. Mahi MJ, Rahad KA, Biswas M, Islam R, Chowdhury ZI. An Accident Detection System for a Single VANET at Low Cost Module. In: IEEE Students Technology Symposium (IEEE Tech Sym), International Conference, IEEE; 2016. p. 54-5.
18. Megalingam RK, Nair RN, Prakhya SM. Wireless Vehicular Accident Detection and Reporting System. In: Mechanical and Electrical Technology, 2nd International Conference, IEEE; 2010. p. 636-40.
19. Oche M, Noor RM, Al-Jawfi AS, Bimba AT, Nasir MK. An Automatic Speed Violation Detection Framework for VANETs. In: RFID-Technologies and Applications, IEEE International Conference, IEEE; 2013. p. 1-6.
20. Papadimitratos P, Buttyan L, Holczer T, Schoch E, Freudiger J, Raya M, *et al*. Secure vehicular communication systems: Design and architecture. EEE Commun Mag 2009;46:100-9.
21. Rahman MT, Mahi MJ. Proposal for SZRP Protocol with the Establishment of the Salted SHA-256 Bit HMAC PBKDF2 Advance Security System in a MANET. In: Electrical Engineering and Information and Communication Technology, International Conference, IEEE; 2014. pp. 1-5.
22. Mozumder DP, Mahi MJ, Whaiduzzaman M. Cloud computing security breaches and threats analysis. Int J Sci Eng Res 2017;8:117-29.
23. Ross SM. Introduction to Probability and Statistics for Engineers and Scientists. Cambridge: Academic Press; 1997.
24. Ross SM. Introduction to Probability Models. Cambridge: Academic Press; 2014.
25. Schoch E, Kargl F, Weber M. Communication patterns in VANETs. IEEE Commun Mag 2008;46:119-25.
26. Minarcin M. Manufacturing the next generation of connected and electrified vehicle. SAE Int J Mater Manuf 2016;9:368-77.
27. Kuang XH, Wang SM. A design of short range wireless communication module based on nRF24L01. In: Applied Mechanics and Materials. Vol. 455. Trans Tech Publications; 2014. p. 372-5.
28. Yang Y, Bagrodia R. Evaluation of VANET-based Advanced Intelligent Transportation Systems. In: Proceedings of the 6th ACM International Workshop on VehiculAr InterNETWORKING, ACM; 2009. p. 3-12.
29. Zeadally S, Hunt R, Chen YS, Irwin A, Hassan A. Vehicular *ad hoc* networks (VANETS): Status, results, and challenges. Telecommun Syst 2012;50:217-41.
30. Bowling SR, Khasawneh MT, Kaewkuekool S, Cho BR. A logistic approximation to the cumulative normal distribution. J Ind Eng Manag 2009;2:1.
31. Zogheib B, Hlynka M. Approximations of the Standard Normal Distribution. Canada: University of Windsor, Department of Mathematics and Statistics; 2009.
32. Agarwal Y, Jain K, Karabasoglu O. Smart vehicle monitoring and assistance using cloud computing in vehicular *ad hoc* networks. Int J Transp Sci Technol 2018;7:60-73.
33. Puar VH, Bhatt CM, Hoang DM, Le DN. Communication in Internet of Things. In: Information Systems Design and Intelligent Applications. Singapore: Springer; 2018. p. 272-81.
34. Gutzat F, Dormann CF. Decaying trees improve nesting opportunities for cavity-nesting birds in temperate and boreal forests: A meta-analysis and implications for retention forestry. Ecol Evol 2018;8:8616-26.
35. Altman DG, Bland JM. Standard deviations and standard errors. BMJ 2005;331:903.
36. Barde MP, Barde PJ. What to use to express the variability of data: Standard deviation or standard error of mean? Perspect Clin Res 2012;3:113.
37. Xiangsheng G, Min W, Qi L, Tao Z. Uncertainty analysis of torque-induced bending deformations in ball screw systems. Adv Mech Eng 2015;7:1687814015573625.
38. Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 2007;50:885-900.
39. Correa CD, Chan YH, Ma KL. A framework for uncertainty-aware visual analytics. In: Visual Analytics Science and Technology, VAST, IEEE Symposium, IEEE; 2009. p. 51-8.
40. Montgomery DC, Peck EA, Vining GG. Introduction to Linear Regression Analysis. Vol. 821. Hoboken: John Wiley and Sons; 2012.
41. Al Mamun S, Suzuki R, Lam A, Kobayashi Y, Kuno Y. Terrain Recognition for Smart Wheelchair. In: International Conference on Intelligent Computing. Cham, Switzerland: Springer; 2016. p. 461-70.
42. Whaiduzzaman M, Sookhak M, Gani A, Buyya R. A survey on vehicular cloud computing. J Netw Comput Appl 2014;40:325-44.
43. Qi H, Shiraz M, Gani A, Whaiduzzaman M, Khan S. Sierpinski triangle based data center architecture in cloud computing. J Supercomput 2014;69:887-907.
44. Li R, Lei KH, Khadiwala R, Chang KC. Tedas: A Twitter-Based Event Detection and Analysis System. In: Data Engineering IEEE 28th International Conference. IEEE; 2012. p. 1273-6.

APPENDIX

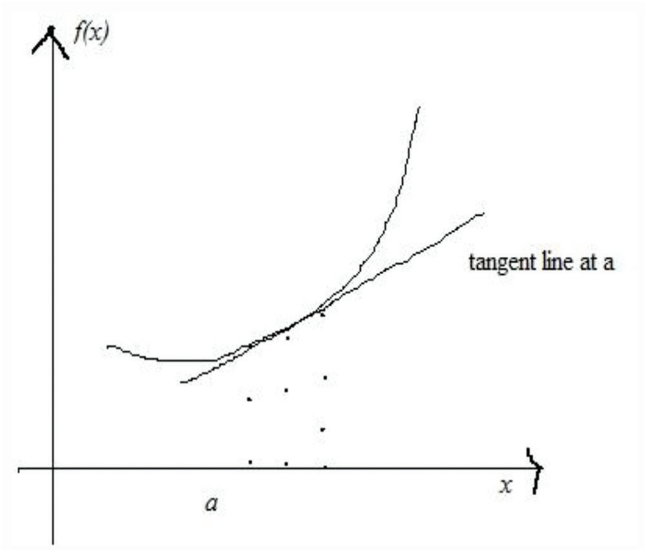

Figure illustration of the function of a straight line and its tangent in compares to linear regression.

Taylor approximation is speculation of linear estimation. It can be characterized basically in the below-stated way.

Definition 1

The n th order Taylor estimation P_n (x) of a function f (x) at a point c is the one of a kind degree n polynomial. Whereas, P_n (x) described in such a way that P_n (c) = f (c), P'_n (c) = f′(c),…, $P_n^{(n)}$ (c) = $f^{(n)}$ (c). In other words, it is the extraordinary degree n polynomial which coordinates the value and first n derivatives of the function f at x = c.

Observation 1

Say the polynomial P_n (x) is proportionate to commit the value of first n subordinates of f (x) at x = c. For expressing the definition a "computable" form is always a need. Supposing, the rest is on the observation that the function f (x) = xn has the property that f (n) (x) = n(n−1) 3, 2 1, which is usually denoted n!. So in particular f (n) (0) = n! But also noting that f (0) = 0 and that all other derivatives of, f at x = 0 are 0. Similarly, if we translate this function to be centered on the point x = c, we obtain the function f (x) = (x−c) n, which has the property that f (n) (c) = n! But f (c) = 0 and all other subordinates of f are 0 at x=c.

To construct the polynomials P_n (x) "each by each, "through viewing merits the following formula is obtained for the Taylor approximation which is centered at x = c.

$$P_n (x) = f (c) + f '(c) + f ''(c) (x–c)^2+f '''(c)^3$$

$$Pn (x) = f(c) + f'(c)(x-c) + \frac{f''c}{2}(x-c)^2 + \frac{f'''c^2}{3!}(x-c)^3 + .. + \frac{f(n)(c)}{n!}(x-c) -\infty < x < \infty$$

This is often written using Σ notation as follows

$$Pn (x) = f(c) + f'(c)(x\text{-}c) + \frac{f''(c)}{2}(x\text{-}c)^2 + \frac{f''(c)^2}{3!}(x\text{-}c)^3 + .. + .\frac{f(n)(c)}{n!}(x\text{-}c) - \infty < x < \infty$$

Definition 2

The Taylor linear approximation is a property to estimate a general function using a linear function. Given a differentiable scalar capacity f (x) of the genuine variable x, at that point the straight guess of the capacity at point a, as appeared in the figure beneath, is acquired by, f (x) ≈ f (a)+f′(a)(x−a); where f′(a) = df (x)/dx|x = a. The articulation on the right-hand side is recently the conditional equation for the tangent line to the graph f (x) at point a. The above description is true when x is nearer to a.