

# Service Outages Prediction through Logs and Tickets Analysis

Sunita A Yadwad<sup>1</sup>  
Research Scholar  
Department of CS and SE  
Andhra University  
AP, India

Dr V. Valli Kumari<sup>2</sup>  
Professor  
Department of CS and SE  
Andhra University AP  
India

Dr S Venkata Lakshmi<sup>3</sup>  
Assistant Professor  
Department of CSE  
GITAM (Deemed to be University)  
AP, India

**Abstract**—Service outage or downtime is a growing challenge to the service providers and end users. The major cause for the unavailability firstly is failure of equipments and applications at various places and secondly failure for proactive diagnosis and rectification. The system activities that are logged and the response of customers and providers in the form of trouble tickets could be studied for minimizing network faults. The downtime can be reduced when the failures are predicted well in time and proactively corrected. Accurate prediction of faults helps in responding to downtime even before the customer tickets are raised or network trouble is encountered. Most of the research focuses on trouble shooting through forecasting the quantity of trouble tickets using the historical ones. If these tickets can be supported with the warning in the form of Syslogs and the technical support of network tickets the predictive models would be more efficient and accurate. Dynamic and truly adaptive machine learning algorithms are essentially required for processing the torrent of data and formulating predictions based on the trends and the patterns existing in it. The work refers to i) identifying number of trouble tickets that are related to the device a few days before the network component fails, ii) predicting fault will occur in broadband networks. Lasso and Ridge regression are used for the first and Bayesian structural time series analysis and prophet are used for the latter.

**Keywords**—Failure prediction; linear regression technique; network fault prediction; lasso; ridge regression; Bayesian structural time series; prophet

## I. INTRODUCTION

Internet is an indispensable service for mankind. Owing to its vivid and continuous usage for various purposes, life cannot be imagined in absence of this service. Detection and minimization of service unavailability or downtime is the need of the hour and a challenging task for the service providers. There are several reasons leading to the service outage.

One of the major reasons is the failure of equipments, service or application. Networks are complex due to rise in demand of novel and diversified applications. As these applications are provided by various vendors and providers it is difficult to detect network failures and diagnose their causes. The data for the detection can come from various sources like Social networking data, Syslogs, Customers tickets, Signal measurement. Customers internet usage data and the network trouble tickets. The data should have been harnessed using

several machine learning algorithms for fault prediction in networks.

The motivation for the present work comes from literature consisting of research papers that focus on fault prediction in networks using different techniques and datasets. The papers surveyed work on fault detection in two major ways:

- 1) Use quantity of customer trouble tickets with time series predictive models for prediction of the quantity of faults.
- 2) Using of System logs, Signal Measurements, data from social networking services and internet usage data that is generated by various network components for prediction of likelihood of the components being faulty in near future.

This paper proposes comparison of several machine learning techniques that help in assessing the priority and intensity of the trouble reports and choose the most optimal solution for investigation. The approach is used for detection and prediction of faults in an efficient and improved manner. To achieve the above goals for network services, the paper uses historical data obtained for the selected equipment several days before the failure which is made up of i) the customer trouble tickets used by the customers to inform about services affected, ii) Network trouble tickets which are technical information about break downs and services interruptions noted and iii) syslogs which are event logs, warning and alarms generated by the equipments.

The model proposed in the paper firstly effectively forecasts the count of customer trouble tickets for coming days using structural time series analysis overcoming all error limitation of ARIMA, GARCH, ARMA model and then uses range of warnings obtained from the historical data several days prior to the failure to formulate a pattern which can effectively predict the what happens to the equipment the following day. When a certain range of cumulative warnings are observed, that the equipment failure will happen the next day is predicted.

Due to confidentiality, the data presented in this paper are masked. The credibility of the model is supported by comprehensive tests.

Owing to confidentiality reasons data presented has been masked and model credentials are supported by several tests. The remaining portions of the paper are categorized in the

following sequence: Section 2 which immediately follows the introduction describes the related works referred and the motivation, methodology is explained in Section 3, Section 4 discusses both the regression and the forecasting results, and Section 5 is the concluding remarks.

## II. RELATED WORK AND MOTIVATION

Literature has in store abundant work focusing on network fault prediction using all possible machine learning techniques. The proposed techniques try to better them in terms of accuracy of prediction of faults. Most of the existing works use the following to deal with network downtime.

- They use total number of trouble tickets created for a time series network fault forecasting model.
- They use logs generated by network component or internet usage and measuring data.

This section discusses a few such works.

### A. Using Customer Trouble Tickets

Addressing customer services are highly desirable for both maintaining reputation of companies and for early prediction of outage [2]. Service outage experiences have been increasing among users of telecommunication industry [3] and root cause of the failure is inability to predict the failures proactively.

The measure of customer grievance tickets which are related to follies in network can be surmised with a model in time series. The quantity of tickets having equal intervals allows discrete representation of data. The trend can be exploited for future sequence forecasting. Customer on premise network equipment faults were focused by University of Telekom in monthly, weekly and hourly intervals using Autoregressive Integrated Moving Average (ARIMA), ARMA, GARCH, Kalman filtering and multivariate recurrent neural network model [1,2,3,4]. With sufficient amount of training data, the number of customer trouble tickets that would be generated was forecasted well ahead to let the service providers govern the allocation of their workforce. All methods were tested against the CMSE (Cumulative Mean Square Error) values for prediction of quantity of faults in broad band network. ARIMA proved to be a winner.

### B. Using System Logs and Network Usage Data

System logs (Syslog) are textfiles which provide an audit trail of events. Applications send information to the syslog process which stores the message in a text file in the order that they arrive [7]. Syslogs validate those tickets which are related to failure analysis and help in prior detection of network component failures. Logistic regression has been highly used for failure prediction. Logistic regression with rule based analysis has been proposed [5] to create a credible model and directly forecast future failure of network components at least four days prior to the actual occurrence. They utilized the historical data which includes the customer trouble ticket, network trouble ticket and Syslog warnings.

Rule based analysis model for the prediction of equipment failure the very next day was constructed with the cumulative sum total of the warnings and the gradients obtained from simple linear regression and best fit of line method. Challenges

are data is available on real time basis. It is huge and needs adequate storage and processing. Also the trouble ticket and Sysco follow different format for different manufacturers.

Study in paper [6] has used the classifiers like Random Forest classifier, C5.0 decision tree algorithm for bettering the forecast of network faults. The customer tickets are combined with network signal data and internet usage data to aptly describe the customer behavior and quality of the network components.

The sliding window for analysis was chosen for seven days, the obtained data in this span from customers was augmented with the internet usage data and signal measurement of network. The C5.0 decision tree and Random Forest (RF) classifiers were used for training the data. The results proved that the RF method had higher accuracy in prediction of network fault and also could estimate the importance of variable for prediction. C5.0 could present the expression leading to prediction outcome for better understanding of the potential causes of failure by network management.

In the purview of the above findings and motivation the main objectives of the paper are:

1) Obtain the warnings from the customer tickets and represent network fault with time series. The obtained trend can help forecast the future sequence in time series. So far the authors have proved ARIMA [1-4] is the best but prove Bayesian structural time series and prophet techniques give better result than ARIMA in warning prediction.

2) With Lasso, Ridge and Elasticnet regression technique and best-fit line methods predict the equipment failure atleast 9 days in advance. The pattern in which warnings occur is studied on the daily to set a trigger well in time to take necessary actions.

The above cited objectives are achieved through experimental results.

Further the work can be extended to improve the prediction of network fault using Hidden Markov model . Data training can be done by applying powerful approach of HMM and Bayesian network model in the decision function and create the fault detection method to find whether the aggregated log data is normal or failures have been detected. HMM is able to identify the important features and list decision rules that describe the relation among selected features from data aggregation of customer tickets, network tickets and System logs. The implementation is beyond the scope of the paper and is presented in subsequent publications. of th. The HMM is trained for normal non outage records.

## III. METHODOLOGY

Historical data is extracted from the grievances of customers as Customer Trouble tickets (CTT), the Network Trouble Tickets (NTT) form service providers and Syslog for a given enterprise.

CTT gives customers complaints about interruption of service whereas NTT gives technical details about breakdown of equipment.

Syslog contains event logs alarms and warnings generated by each equipment. As the network elements grow, there is huge volume of complex log data. Information in the logs needs to be extracted efficiently and precisely for trouble shooting and maintenance. Records are matched based on the equipments id. Equipments could fail due to several reasons. If the reasons are climatic like weather, power vandalism or theft it is filtered out.

By matching the tickets data from customer, the failure records of the equipments can be identified. The faults listed in customer trouble tickets are cross verified with the causes in Network ticket data which has detailed elucidation about the failure of the equipment. This is followed by analysis of the warnings generated by the equipment. Using Sequential Query Language (SQL) there is extraction of all important service failure related features causing the network equipment malfunction. The selected features from the historical data are listed below in Table I.

The block diagram in Fig. 1 represents the architectural model of the outage detection through fault forecasting and prediction.

The data collected from all the three sources is firstly stored using HDFS and later distributed across all the clusters. Each and every node processes the data using SQL in Hive. Due to presence of SQL there is no need to group or reduce the data. The native function of SQL classifies and processes the data all together to determine the total number of occurrences for the number of warning per equipment. The original raw data for very equipment in Syslog is more than million records as one month of data is used for the prediction model. The mappers key/value pairs are created with each Warning type and the respective counting values. The reducers are set to aggregate counting values giving the daily accumulative total for each Warning type and for each equipment. This in turn serves as the selected feature related to communication failure due to the failure of the equipment. These features selected are in turn fed to the regression model.

On the other hand the total numbers of customer tickets that are accumulated per day are fed to the BSTS a prophet time series model as fault data for further forecasting for coming times.

**A. Regression Model**

Large amount of data is collected by the equipment making the detection of failure time taking. Simple and effective methods are required for prediction with the real time and continual data. Linear regression model has been used for prediction of faults and the paper [5] has achieved accuracy of predicting faults days prior to the occurrence. The goal in this paper is to predict faults at least nine days prior to failure for proper preventive measures and continued service.

In linear regression theory before constructing the model a relationship is determined between the independent predictor variable and the dependent predicting variable. The predictor variable X is one on which prediction is based and the criterion variable Y is the one used to predict. Regression line which is a straight line is formed when Y gets plotted in terms of X as a function.

TABLE I. SELECTED FEATURES FROM HISTORICAL DATA

Customer Trouble Tickets	Network Trouble Tickets	Syslog
Date by Customer	Date by Network	Date of Syslog
Fault of Equipment E1	Causes of Fault of E1	Types of Warning for E1
Types of Equipments	Types of Equipments	Types of Equipments
Equip id	Equip id	Equip id

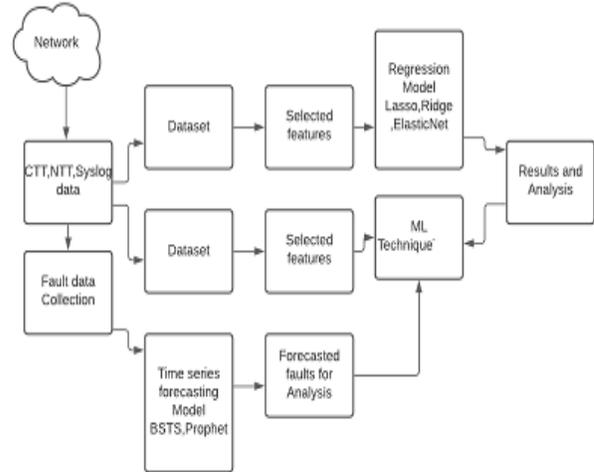


Fig. 1. Fault Forecasting and Prediction Model.

Regression formula can be written as.

$$Y = AX + C \tag{1}$$

Here A is the gradient Y and X are predicted outcome and predictor variable.

In [5], simple linear regression considers cumulative total warnings for previous days to predict what happens the next day. The variable Y is cumulative total of warnings and X represents the day the warnings will be recorded. The regression line is computed by taking values of Y against X. The value of square of correlation coefficient R<sup>2</sup> is computed which depicts the variance of one variable over the other. Its range varies from 0 to 1. As the value nears 1 the data has stronger relationship and determines the certainty of the predictions.

There are several cases where the classical linear regression model doesn't handle data well and accuracy can be further improvised with dimension reduction or regularization. In the Ordinary Least Squares (OLS) approach, variance and bias can be reduced with an approach called Regularization which is beneficial for the improvement in the predictive performance. The process adds information to solve the issues of ill posing of problems and prevent the over-fitting. Commonly used regularization method includes adding a constraint to the loss function

$$\text{Regularized Loss} = \text{Loss Function} + \text{Constraint}$$

Most popularly used forms of constraints in regularization are the Ridge Regression, the Least Absolute Shrinkage and

Selection Operator (Lasso Regression) and the Elastic Net regression. The R package used for implementation of regularized linear models is *glmnet*. Elastic Net can be tuned with function called *caret*. Ridge regression minimizes the sum of square residuals and penalizes the size of the estimates of the parameter towards a shrink zero. Lasso is identical to Ridge conceptually. It adds penalty to non-zero coefficients like Ridge regression penalizes the sum of the squared coefficients (L2 penalty) and the Lasso penalizes the absolute values sum (L1 penalty) [8, 9]. A new regression technique called Elasticnet is combination of both the regression methods to even out what is best in them. The combined penalties of both the regressions when tuned by cross validation leads to minimization of the loss function.

Linear regression equation looks like:

$$Y = \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_n X_n \quad (2)$$

The regularized regression methods can be understood in terms of why and how they are applied to ordinary least squares. The OLS regression mainly tries to find a hyper plane which minimizes the sum of squared errors between a predicted response and observed values [10, 11, and 12].

The cost function for ridge regression is

$$\text{minimize } (SSE + \lambda \sum_{j=1}^n \beta_j^2) \quad (3)$$

Here we come across an extra term, which is known as the penalty term. The  $\lambda$  in the equation here is actually represented by the alpha parameter in Ridge function [8, 9, and 10]. The penalty term here can be controlled by varying the values of alpha. When alpha takes higher values, the penalty term becomes bigger and therefore the magnitude of coefficients is smaller. So, we can see that there is a slight improvement in a model because the value of the R-Square increases.

Lasso regression method selects only few features while reducing the coefficients of other features to zero. This feature selection property does not exist in ridge.

Lasso regression is similar to ridge except that we swap L norm for  $L_1$  norm. Instead of adding squares the absolute value is added for further improvement in the  $R^2$  value and increase in the fitness of the mode.

$$\text{minimize } (SSE + \lambda \sum_{j=1}^n |\beta_j|) \quad (4)$$

Elastic net uses both L1 and L2 penalty term, therefore its equation look like as follows:

$$\text{minimize } (SSE + \lambda_1 \sum_{j=1}^n |\beta_j| + \lambda_2 \sum_{j=1}^n \beta_j^2) \quad (5)$$

The best model is defined as the one that minimizes the prediction error. In this case Elasticnet overpowers the other regression methods making it best fit for the prediction of faults.

### B. Bayesian Structural Time Series

It is well known amongst data scientists that non-stationary series, shortage of data points and inability to distinguish trends of a time series will lead to inaccurate forecasts. Many forecast algorithms face this drawback. Structural Time Series models imitate and combine all the endemic features of regression models namely ARIMA, and the exponential smoothening

models. Bayesian Structural Time Series model is also popularly known as 'state space models' or the 'dynamic linear models' by different authors. It is a time series model that fits the overall structural change in time series dynamically. BSTS is the implementation of this model in R which is easy to use and is a function which requires a minimal mathematical understanding of the state space models [13, 14, and 15]. The benefits of Kalman filtering algorithm and Markov chain Monte Carlo (MCMC) are together used for fitting this model. Forecasts are then calculated from the predictive distribution of posterior. BSTS is more popular for its "Now casting" feature of predicting the values of time series in the present.

Prophet is a popular forecasting method which was developed at Facebook and is available as open source. It is a curve fitting approach, similar to how BSTS models the trend and the seasonality, except for the fact that it uses generalized additive models not the state-space representation for describing each component [16, 17]. For rigorous performance analysis of these methods calculating of forecast error related metrics like MAPE and RMSE is required. The findings of the study suggest that forecasting accuracy in the proposed models is better when measured against the frequently used ARIMA models.

## IV. RESULTS

The fault forecasting and prediction process as described in the Fig. 2 can be divided into two major parts:

- The forecasting of faults using time series model for dealing with outage.
- The Network fault prediction nine days prior to the occurrence for proactive dealing with outage.

### A. Hypothesis Testing on Lasso and Idge Regression for Failure Prediction

The dataset has approximately millionn records for months of data that has been logged about the event track records of equipment E1. The hypotheses are tested on the Equipment E1 which is known to have the maximum customer complaints. The complaints in the first month for the equipment failure as well as the consecutive two months are recorded. For consistency in the equipment failure's trend, this regression technique can be validated on other equipment of the same model .Testing is done on the equipment E1 for warning X. The choice of equipment depends on the highest customer tickets generated in the months listed. The trend developed on the equipment can be further used to test other equipments. In paper [5], where a simple regression method is used to predict equipment failure the value of  $R^2$  is 1 for four days prior to the equipment failure E1 and  $R^2$  was 0.6685 when checked about 9 days prior to failure of the equipment , which means smaller number of days is better in prediction of equipment's failure. To improvise the method of linear regression and to increase the accuracy of prediction the paper uses Ridge and Lasso regression methods. The value of  $R^2$  increases significantly. The value of coefficient  $R^2$  is 1 for four days and 0.8811, 0.8283, 0.8844 for 9 days with the Elasticnet, Lasso and Ridge regression techniques making it easier to fit for the model if the number of days is more. This helps us predict the fault nine days advance making it go easier with the service providers.

This has been tested on the equipment E1 with nine days and can be further tested on other equipments varying the number of days. The results of the regression technique is depicted in Table II.

In conclusion, the proposed method is valid for predicting failure in network equipments nine days before the actual occurrence. This means that the regression line better fits the actual data when the number of days is 9 and the methods are Lasso and Ridge regression. The cumulative warning range, gradients obtained and the other features of the dataset can be used to effectively predict the equipment failure and also estimate the variables that are important for prediction.

The value of  $R^2$  when predicting the upcoming warnings for equipment E1 with Linear, Ridge and Lasso regression are as follows. Lasso and Ridge outperform the linear regression method in terms of best fitting of the model for prediction.

The hypothesis tested can be represented and the results can be summarized as below. The value of  $R^2$  shows how well the regression line fits the actual data.

The paper achieved value of  $R^2$  as 1 by linear regression and proposed their model can predict the fault 4 days prior to actual occurrence. But our regression model tested on three other methods lasso, ridge and elasticnet show the value of  $R^2$  as 1 for 6 days and 0.8843 for 9 days. This proves that linear regression can be improvised by our model to achieve the fault occurrence warnings up to 6-9 days prior to the failure giving the service providers to take preventive actions.

**Advantages of proposed Regression Model**

The  $R^2$  values that are obtained by the model are closer to 1 for not only 4 days prior to failure but show the strongest linear relationship to 6 days and 9 days prior to failure. This indicates that the regression line is fitting closely in terms of data to the actual.

**B. Time Series Model for Fault Forecasting**

Generally speaking, faults in the broadband network can be identified in two basic ways, first way to detect using a variety of surveillance systems that monitor network operation, another method deals with customer reports. These two sets of data are more or less overlapped. Union set includes those faults that have been reported by customers through tickets and those that have been recognized by the network supervisory system through network tickets at the same time. The prediction results of the methods mentioned in the methodology are given below. Mean Absolute Percentage Error commonly known as MAPE is used as the criterion for results evaluation. The results are represented visually in diagrams describing the relationship existing between both the original and the predicted values.

A 24 hour distribution of fault occurrence having period with 5 min interval is shown in Fig. 3. The figure shows how the faults have been distributed over a period of a day. This distribution is subjected to the ARIMA, BSTS, and Prophet forecasting time series methods for fault forecasting. This forecasting is a lead to the service providers to buck up their resources to overcome failure and avoid further outages and customer enrage.

The appearance of various faults is a statistical process with respect to time which is represented with time series, sampling of the faults data happens in every 5 min interval and the interval sequence has a cumulative feature which facilitates the series of discrete time parameter. Members of series can be forecasted. Every forecasting method is distinctive in its own way. The characteristics of every method cannot be considered a hundred percent for sure. Adequate methods are required for increase in the accuracy of the prediction.

The aim of the work is examining the prediction of faults through three different models ARIMA, BSTS and Prophet. The model is designed on the basis of monitoring of the network broadband faults analyzed in the order in which they appear. The model is described in Fig. 3.

TABLE II. RESULTS –  $R^2$  COMPARISON

Computed value of $R^2$ for Equipment E1				
S.no	Regression Method	4 days	6 days	9 days
1	Linear	1	1	0.6685011
2	Elasticnet	1	1	0.8811015
3	Lasso	1	1	0.8283883
4	Ridge	1	1	0.8843964

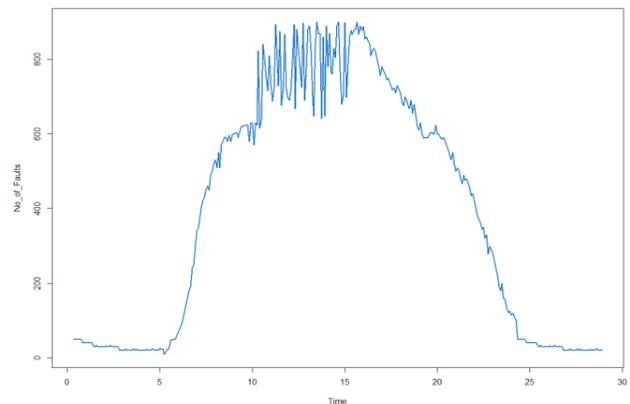


Fig. 2. Daily Fault Distribution.

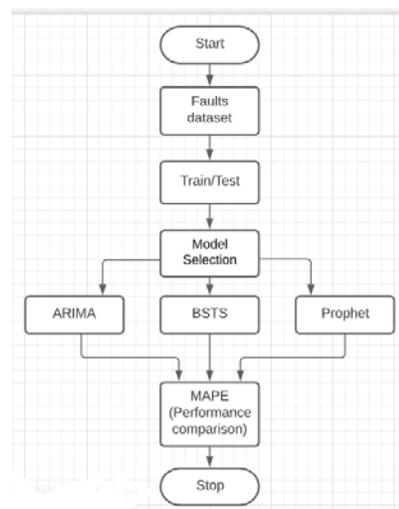


Fig. 3. Forecast Model.

Accuracy of the above mentioned models are assessed by comparing the predicted results obtained against the actual data. The results obtained in fault prediction for a randomly selected day are shown in Fig. 4 to Fig. 6 by ARIMA, BSTS and Prophet Method.

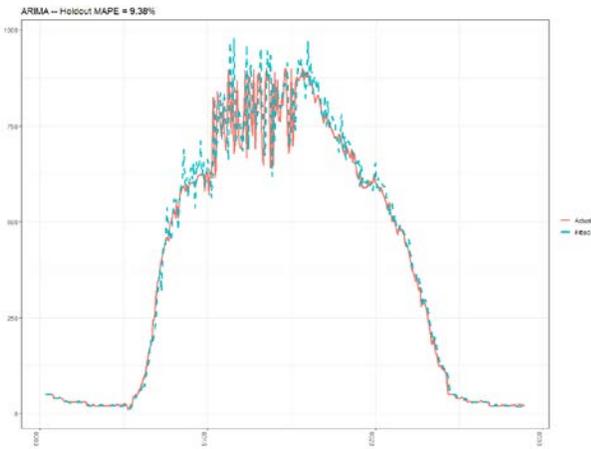


Fig. 4. Results of Forecasting using ARIMA.

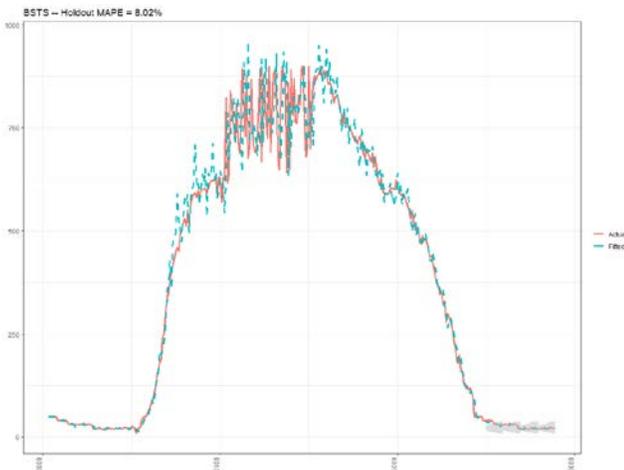


Fig. 5. Results of Forecasting using BSTS.

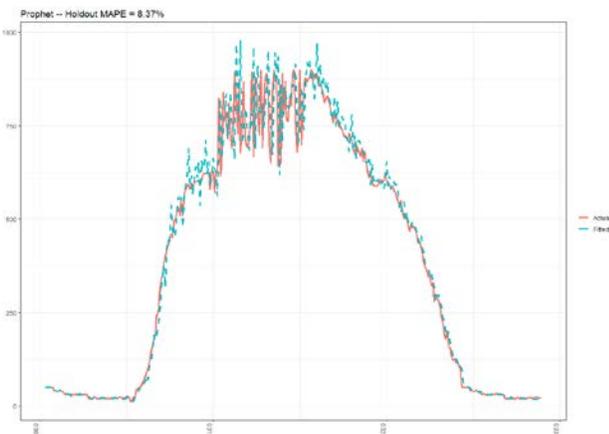


Fig. 6. Results of Forecasting using Prophet.

Focus of the paper is trying to find the best candidates for time series forecast of occurrence of faults in the broadband networks.

Comparison is made for the checking accuracy which is the difference between predicted and actual data for each of the three models. The accuracy of prediction criterion is expressed by using MAPE. The MAPE errors obtained are shown in Table III.

TABLE III. RESULTS – MAPE COMPARISON

Performance Metrics	MAPE %
ARIMA performance	9.38
BSTS performance	8.02
Prophet performance	8.37

Comparing these three models over prediction, we can conclude that better results were achieved by BSTS and prophet method over ARIMA STL. The tools show respective efficiency and achieve better results.

#### Advantages of Forecasting Model

The motivation for the work is that not much forecasting activities are undertaken to predict the rate of faults in a network for prevention of outages. There were some previous studies to predict fault using Kalman filters, GARCH, HMM, ARMA and ARIMA methods. Getting a formula, algorithm to with a good accuracy to predict the amount of faults is certainly a challenge. The advantage of the proposed model for forecasting is usage of Bayesian structural time series and Prophet Method which surpass the ARIMA method which has been the best contender in the surveyed papers.

#### V. CONCLUSION

The forecasting of faults for the broadband network accurately is the need of the hour for internet service providers. It helps them to properly strategise the future operational expenses and plan the strategies for increasing their business efficiency. The forecasted data is a means to make decisions concerned to the network maintenance, investments in terms of new equipments and proper work to resource allocation. Proactive actions can be directed in the areas which are identified to be the potential generators of network service faults. Increase quality of service to the customers is the major driving force behind the research.

Firstly the paper studies data about the equipment through customer grievances, through network tickets and the warnings of Syslog. Failures related to Equipment can be predicted to ascertain less downtime and more customer satisfaction. Algorithms namely Lasso and Ridge to process the huge amount of data through a regression technique to predict the equipment failure are proposed and implemented. They optimize linear regression and improve the model for early prediction with better accuracy. The work can be further expanded by adding additional warning types and by establishing significant relationship among the different types of equipment failure.

Secondly this paper compared three short-term prediction models to derive the number of faults that would occur in telecommunication networks namely BSTS, Prophet and ARIMA which previously has been proved to be the best method. The aim was finding the best contender for the analysis and forecasting of faults occurring. The conclusion drawn was that the BSTS and Prophet models showed higher prospects in forecasting network faults in telecommunications networks.

In future the achieved accuracy can be further improvised by expansion of the model with addition of different data sources, accumulated both from the network and other external information. These can be further assimilated with probabilistic chains of the dynamic Bayesian network and Hidden Markov Model. This can be the topic for future extension of the research.

#### REFERENCES

- [1] Sonny Yuhensky, Hafiddudin, Rendy Mi, et al. "Forecasting the formulation model for amount of fault of the cpe segment on broadband network PT.Telkom using ARIMA method ". In Control, Electronics, Renewable Energy and Communications (ICCEREC) , 2016 International Conference on, pages 185–191. IEEE, 2016.
- [2] Zeljko Deljac, and Gordan Krcelic, Mirko Randic , "A multivariate approach to predicting quantity of failures in broadband networks based on a recurrent neural network" . Journal of network and systems management, 24(1):189–221, 2016.
- [3] Deljac, Zeljko, Boris Spahija and Marijan Kunstic, "A comparison of traditional forecasting methods for short-term and long-term prediction of faults in the broadband networks." MIPRO, 2011 Proceedings of the 34th International Convention. IEEE, 2011.
- [4] Zeljko Deljac and M. Kunstic. "A comparison of methods for fault prediction in the broadband networks." Software, Telecommunications and Computer Networks (SoftCOM), 2010 International Conference on. IEEE, 2010.
- [5] Lam Hai Shuan, Guo Xiaoning, Tan Yi Fei, Soo Wooi King and Lee Zhe Mein." Network equipment failure prediction with big data analytics". International Journal of Advances in Soft Computing & Its Applications, 8(3), 2016. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [6] Ji Sheng Tan, Amy Hui-Lan Lim ,Chin Kuan Ho, "Predicting network faults using random forest and C5.0," International. J. Eng. Technol., vol. 7, no. 2, pp. 93–96, 2018.
- [7] T. Kimura, K. Takeshita, M. Yokota, K. Nishimatsu, , T. Toyono and T. Mori, "Network Failure Detection and Diagnosis by Analyzing Syslog and SNS Data: Applying Big Data Analysis to Network Operations," NTT Technical Review, Nov 2013, Vol. 11, No. 11.
- [8] Rodríguez, Oldemar. (2013). A Generalization of Ridge, Lasso and Elastic Net Regression to Interval Data. 10.13140/2.1.3753.0883.
- [9] Friedman, J., Simon, N., Hastie, T, Tibshirani, R. , (2015). <http://cran.r-project.org/web/packages/glmnet/index.html> , glmnet: Lasso and elastic-net regularized generalized linear models.
- [10] Pereira, José & Basto, Mario & Ferreira-da-Silva, Amélia. (2016). The Logistic Lasso and Ridge Regression in Predicting Corporate Failure. *Procedia Economics and Finance*. 39. 634-641. 10.1016/S2212-5671(16)30310-0.
- [11] Tibshirani, Robert. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, 1996, pp. 267–288. *JSTOR*, [www.jstor.org/stable/2346178](http://www.jstor.org/stable/2346178). Accessed 18 Jan. 2021.
- [12] Friedman,J, Hastie.T, Simon .N ,Tibshirani R.,(2015). glmnet,Lasso and elastic-net regularized generalized linear models. R package version 1.9-8, <http://cran.r-project.org/web/packages/glmnet/index.html>.
- [13] KIM LARSEN , "Sorry ARIMA, but I'm going Bayesian" ,April 21,2016,San Francisco ,CA.
- [14] G. Papacharalampous, H. Tyralis, D. Koutsoyiannis, "Predictability of monthly temperature and precipitation using automatic time series forecasting methods", *Acta Geophysica* 66 (4) (2018) 807–831. .
- [15] Abhinaya Ananthakrishnan "Forecasting ? Think Bayesian" .
- [16] July 8<sup>th</sup> ,2018 , End to end problem solving ,Data nerd,Facebook ,UT ,Austin.
- [17] Almarashi, Abdullah & Khan, Khushnoor. (2020). Bayesian Structural Time Series. *Nanoscience and Nanotechnology Letters*. 12. 54-61. 10.1166/nml.2020.3083.
- [18] N. Bounceur, I. Hoteit and O. Knio, "A Bayesian Structural Time Series Approach for Predicting Red Sea Temperatures," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 1996-2009, 2020, doi: 10.1109/JSTARS.2020.2989218.