"Waiting Ticket Optimization using Reservation Chart Cluster for Indian Railway Reservation System"

Javed Akhtar Khan
Computer Science & Engineering
Indur Institute of Engineering & Technology
Telangana, India

Abstract: Data mining KDD and Cloud Concept is used for optimizing the waiting ticket in Indian railway passenger who are make reservation but not get the confirm reservation. Today Indian Railway has a very biggest network in all over world. Indian rail has 6853 stations in India, and up to 13 millions of passengers are being served by Indian railway in a day. Indian railway is improving day by day towards passenger neatness, safety, cleanness and tries to providing good facilities to maximum passenger during journey. This is very unfortunate to tell in the era of Digital India a major work is start with the help paper that is ticket checking and allotting of vacant seats on running status, And this is increases the corruption in Railways, Indian govt. and the rail authority has to do a very hard work to improve on this area but up to now no alternative are found. So This is the major problem include in this manuscript, This research article is to provide the alternative solution for Indian railway passenger who make reservation but not get the confirm ticket after chart prepare by system .

Key word: Cloud Computing, Data mining, Cluster algorithm.

I. INTRODUCTION

Indian Railway is biggest network in the world . So this biggest network total 6853 railway stations are located and approx 13 million of passenger are travel in a day . Table 1 is describe the symbol that use by Indian railway system for the PNR Status or ticket status of passenger some time these concern information is shown in reservation ticket make by passenger .

Symbol	Description
CAN / MOD	Cancelled or Modified Passenger
CNF / Confirmed	Confirmed (Coach/Berth number will be available after chart preparation)
RAC	Reservation Against Cancellation
WL #	Waiting List Number
RLWL	Remote Location Wait List
GNWL	General Wait List
PQWL	Pooled Quota Wait List
REGRET/WL	No More Booking Permitted
RELEASED	Ticket Not Cancelled but Alternative Accommodation Provided
R##	RAC Coach Number Berth Number
WEBCAN	Railway Counter Ticket Passenger cancelled through internet and Refund not collected
WEBCANRF	Railway Counter Ticket Passenger cancelled through internet and Refund collected

Table 1 Abbreviation used by Indian Railway[1]

The some statists of ticket reservation system and ticket cancelation system are as mention in the figure 1.
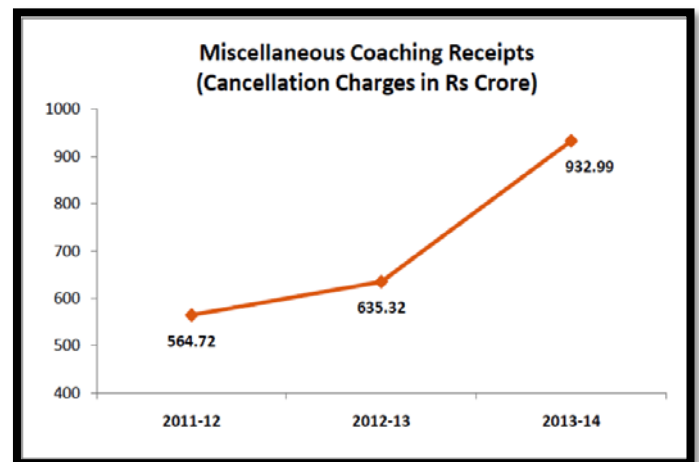

Figure 1 Ticket Cancellation Revue Statistic[2]

In 2013-14, a total of 840 crore passengers travelled via Indian Railways. While Maximum people is depend on the railways for their movement, many complaints have been received in the pervious about how the tatkal ticket booking system has been misused. In the latest bid to discourage misuse of reservation tickets and optimization of available alternative trains, railway make a new set of cancellation/refund and booking rules. In this introduction part of paper include the some static of ticket revenue of the Indian railway system yearly apart from that the mention information in this section maximum passenger are facing the problem after making the ticket but not get the confirm birth or reservation .A way of ensuring the reservation ticket is confirm or not via SMS , Train Chart and PNR status if ticket is not confirm before departure train form the origin station . Train chart is list of all passengers that have reserved seat on that train with all of their details like Name, Gender, Age and station of boarding and destination. It also consist of Coach number ,passenger birth number of passenger with confirmed ticket and RAC. This chart is usually pasted on notice board of your train platform or at

your coach. When you check PNR status, you may have found this word Chart not prepared. Passenger must be tens what does this means. As i told earlier if you have your seat in reserved category or confirmed then you do not need to worry about this chart preparation. But if your seat is in the waiting list or RAC then passenger have to wait till chart preparation, as only after chart preparation, so passenger will get to know whether seat is confirmed or not. If your seat got confirmed at the time train chart preparation, you will get your seat and coach number in the chart. Otherwise passenger will be in waiting list table . At this situation Railway Ticket Collector (Abbr. TC) play a important roll for solving this problem , TTE [1][2] (Travelling Ticket Examiner)/ TTI (Traveling Ticket Inspector) is the person who verify that all passenger in the train is traveling with valid ticket of appropriate denomination.]

II. DATA MINING AND CLUSTERING

Clustering is a mostly used technique in data mining, at present there exists many clustering algorithms, but so many clustering algorithms is not good for the large data sets. In the partitioning algorithm can be define as a set of similar objects of a databases into same groups. Clustering is a simple and very popular process to implement the partition of data. It is widely used in data mining, statistic, biology, and so on. The most distinct feature of data mining we work for the large amount of data and complex data sets. K-means [11] clustering method put forward by MacQueen is very easy and the most simply algorithm. However, it only works on numerical data set and can not cluster categorical attributes and mixed attributes K-modes and k-prototype [12] is the extending of k-means. They can cluster categorical attributes and mixed attributes respectively. But it has such shortcomings as instability, random city and so on. And when clustering big data sets it has bad efficiency. The data mining community has recently put a lot of efforts on developing fast algorithms for make a cluster of huge large data sets these algorithm are used like CLARANS [13], DBSCAN [14] and BIRCH [15] these algorithms have shown some significant performance improvements in clustering very big data sets. Again, these algorithms still target on numeric data and cannot be used to solve massive categorical data clustering problems.

III. RELATED WORK

This paper is the study about the passenger (Customer) satisfaction[3] Indian railways has its influence on service quality. For finding the Satisfaction of the Indian railway passenger author are part of this paper is make the study of how products & various services supplied by a vender to meet and satisfy the customer expectation yes or not . This study has tried to find out the level of satisfaction of Indian railway passenger .In this paper researcher are collect the total of 500 respondents have been taken by applying random sampling method and do the hypothetical analysis, and find the reveal that the passengers are fairly satisfied with the services provided by the rail system. During the study researcher are make a table and form for finding the level of customer satisfaction in researcher are make a form that include the some simple query related to customer and try to find the answer the some answer related to Indian

railway related problem , and then author make a sampling with using SERQUAL tool author find satisfaction level of passenger .The next article [4] is make concept of passenger reservation system using the online transaction processing .Author are make IMPRESS that means Integrated Multi train Passenger Reservation System , and proposed the general architecture of reservation system .so this article present the CONCERT is able to give the multiple facilities to the passenger like route oriented reservation system it means passenger are able to make ticket any ware from the India ,passenger are able to change the train profile , this is store the automatic database recovery .in the reference number [5] author are make the studying advantages and disadvantages of system structure of Bapat Chourah Indore MP, a alternating queuing model is introduce by the researcher in this model passenger are come the reservation counter and take the slip of token and wait till passenger number is come this system is optimize the queue in the ticket counter .UID based ticket reservation is introduce in reference [6] that able to fetch the information form the UID number of passenger with the help of this technique save the time . This paper proposes the Dynamic Seat Allocation (DSA) system considering the advantage of QR [7]code processing along with one of the standards of wireless communication, with a good QR device TTE check the passenger ticket and update the data base .Since the site with a bandwidth of 450 mbps cannot book more tickets than it is doing at present, researcher are suggestions which could improve the successful transaction rate and also help the in-emergency passengers to take reservation tickets easily, though at a higher cost. During the Tatkal[8] reservation time, 10:00 am to 12:00 pm, the customers face the problem of booking ticket due to heavy traffic. One of the step to help solve this problem to a great extent is that by providing a three-tier system of ticket booking inside IRCTC, i.e. to divide the booking timings into three slots:
1. 8:00 am – 10:00 am : in-Emergency Reservation
2. 10:00 am – 12:00pm : TATKAL Reservation
3. 12:00 pm onwards : Normal Reservation
Dividing the booking timing into these three parts will definitely decrease the load on the servers of IRCTC during the TATKAL timing and also help the Indian Railways in earning more from the emergency tickets booked in between 8-10 am. The work we propose is that IRCTC should divide the booking period into three intervals
In-Emergency (8-10 am):. TATKAL (10am – 12 noon):. Regular (12 noon onwards):
This is manual introduce by the intel system[9] that are describe the all the step for making the online reservation using the Indian railway web site www.irctc.co.in, in this manual technology are describe the Steps to create an IRCTC Account:
Opening the IRCTC site
Creating an Account
Activate your Account
Booking a ticket
Printing your ticket
Printing/ Cancelling a ticket
Sign out of your IRCTC account

IV. PROBLEM STATEMENT

In this article we are introduce the problem related to waiting ticket of Indian railway reservation system and problem face by the TTE for filtering the vacant seat with its allotment . let as see the problem definition in detail . If two person are traveling[10], one with more waiting then other and if that person come in contact with T.C and offering some bribe, than that person get there ticket conformed. This happen because there is no any cross checking on that. This is a major limitation of present system. This improves paper consumption and improves manual work also. So problem of passenger who do a reservation but not have a confirm ticket . Now exactly what happen A person find the TC during a traveling for getting a information about the availability of seat , he was say to TC there is any seat is vacant ? when did birth is free ? one major problem in the existing train chart is all the Ticket Type of particular train is printed in the one table . This all the information is get from TTE only at the journey time in this case Passenger are not able the get the any kind of information about the vacant seat information at the journey time .

V. PROPOSED METHODOLOGY

The concept of Cloud computing and Data mining cluster technique are used for making a cluster of Train chart that able to reduce the TTE complex task regarding the vacant seat allotment at the journey time also reduce the paper work this proposed solution is consume the less time with existing system . So in this we are proposing new architecture for the Indian railway services for passenger based waiting ticket reservation status checking model we are include some new application that gives the information about the vacant seat at the time of journey.

VI-Algorithm

Step 1- A passenger - approaches to a PRS (Passenger Reservation Counter) to book his ticket.
Step 2- A Passenger – Get the confirm Ticket .
Step 3- If Passenger get RCA or Waiting Ticket . (After the II Train Chart)
Step 4 –Cluster I Number of Confirm Ticket (Source to Distention Station)// Filter using the Data mining Cluster Technique .
Step 5- Cluster II Number of RAC Ticket (Source to Distention Station)
Step 6- Cluster III Number of Waiting Ticket (Source to Distention Station)
Step 7 – TTE Filter the vacant seat during the checking and immediately allot the vacant seat to RAC Passenger and waiting passenger without checking the whole chart .

VII. PROPOSED CATEGORICAL CLUSTER AND ATTRIBUTE

This section of paper is include the technical concept of Data mining Cluster technique a Categorical clustering data mining technique proposed in this article three attribute are use mention in the table of cluster I cluster II cluster III these data set cluster are maintain the entries of attribute

each station route , so using this technique at the last final cluster that is II cluster table include the all information .
A categorical data set if it is finite and not ordered it means that in the existing system of railway show the categorical data in train chart this train chart is include the all category of ticket like confirm ticket , RAC ticket , Waiting Ticket in a single table .However, such relationships may exist in real world databases.
Cluster I(Table Include only All Confirm Ticket) = Number of Ticket Book(always grater than Train capacity 650) Number of Ticket Confirm (Equal to Train Capacity500)
After First Train Chart = number of Ticket Cancel after first Train Chart (30 ticket Cancel) so RAC is convert in to the Confirm ticket.
Cluster II(Table Include all the RAC Ticket that are not confirm after the first chart) and waiting are optimize .
650 = booked against train capacity 500 RAC =50 WL =100 = 650
500 confirm (at the time of reservation because this actual train capacity)
Suppose 30 ticket cancel RAC become confirm (20 balance) //After the II Train chart
At the journey time After II train chart 70 passenger are not present in the allotted seat at the time of journey
20 RAC seat become confirm (70 -20 =50) and 50 waiting are clear after this 50 is waiting.
Cluster number III is only showing these data.
So this is very easy for TTE to allocate this seat to passenger after the vacancy creates.

Table I Cluster I

Categories → Station Name ↓	Confirm (sample data500)	RAC (sample data500)	Waiting (sample data500)
A	134 / 113	(134-113)21 / 12	134-(113+12)=9
B	78 / 53	(78-53)25/07	78-(53+7)=18
C	214 /192	(214-192)=22/5	214-(192+5)=17
D	96 /60	(96-60)=36/11	96-(60+11)=25
E	84 /52	(84-52)32/10	84-(52+10)=22
F	44 /30	(44-30)14/5	44-(30+5)=9
Total	Distributed 650 /500	50	100

Table II Cluster II (Final Cluster)Eg 30 ticket cancel

Categories → Station Name ↓	Confirm (sample data)	RAC (sample data)	Waiting (sample data)
A	134 / 113/ 2/115	(134-113)21 / 12 /10	134-(113+12)9/6
B	78 / 53/3/56	(78-53)25/07/5	78-(53+7)18/13
C	214 /192/5 /197	(214-192)22/5/0	214-(192+5)17/0
D	96 /60/36/96	(96-60)36/11/0	96-(60+11)25/14

E	84 /52/9 /62	(84-52)32/10/1	84-(52+10)22/12
F	44 /30/00	(44-30)14/5	44-(30+5)9
Total	Distributed 650 /500 /470	50/20	100/70

Table III Cluster III (Proposed Cluster during the journey)

Categories ➤ Station Name ↓	Confirm (sample data)	RAC (sample data)	Waiting (sample data)
A	101 +12/113	12	9
B	46+7/53	7	18
C	192	0	17
D	49+11/52	11	25
E	52	0	22
F	30	0	9
Total	470 /500	30	100

These Modified Cluster table able to help the TTE for identification the PNR , Vacant Seat Number , Total number of seat is in the buffer of Waiting how much RAC as well as Waiting ticket are clear at the journey time .

VIII. CONCLUSION

This paper is discuss problem face by the passenger who are not get the confirm ticket after the final train chart during the journey , in this article we are also identify the problem face by the TTE that is identifying the vacant seat at the time of journey. Proposed algorithm of this paper is make a more than one cluster that is confirm ticket cluster RAC cluster and Waiting Cluster is make by using the cluster technique of data mining , last cluster that is cluster number III   is only include passenger PNR and other information whose ticket are not confirm after the final train chart preparation . So this modification is very help full for the TTE as well as Passenger .

REFERENCE

[1] http://career.webindia123.com/career/options/ defence_government/railway-ticket-collector/intro.htm
[2] https://factly.in/new-railway-reservation-and-cancellation-rules
[3] G.Rajeshwari1 Dr.D.Elangovan et al   THE RAIL SYSTEM (A STUDY IN SALEM DIVISION OF SOUTHERN RAILWAY ZONE) RACST International Journal of Commerce, Business and Management (IJCBM), ISSN: 2319–2828 Vol. 3, No. 2, April 2014
[4] Seema Agarwal Computerized Passenger Reservation System for Indian Railways – Its Development and System Architecture Journal of Engineering, Computers & Applied Sciences (JEC&AS) ISSN No: 2319-5606 Volume 2, No.6, June 2013
[5] S. Vijay Prasad et al Alternate queuing system for tatkal railway reservation system www.pelagiaresearchlibrary.com Pelagia Research Library Advances in Applied Science Research, 2014, 5(6):120-125 .
[6] Asoke Nath Smart Computing Applications in Railway Systems - A case study in Indian Railways Passenger Reservation System International Journal of Advanced Trends in Computer Science and Engineering Available Online at http://warse.org/pdfs/2014/ijatcse01342014.pdf Volume 3, No.4, July - August 2014
[7] Man Mohan Swarup   et al A QR Code Based Processing For Dynamic and Transparent Seat Allocation in Indian Railway IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012 ISSN (Online): 1694-0814
[8] Saurabh Singh et al "A Most Efficient Price based Tatkal Scheme for Indian Railway"   International Journal of Engineering Technology & Management Research Journal Volume 1, Issue 1 ISSN: 2320-5288
[9] Introducing Online Railway Reservation Systemintel www.irctc.co.in
[10] https://www.quora.com/How-does-the-Waiting-List-status-work-in-Indian-Railways
[11] J.MacQueen. Some methods for classification and analysis of multivariate observations. Pro.5th Berkeley Symp.Math.Statist, Pro.,1:128-297,1967.
[12] Z. Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Mining and Knowledge Discovering, 2:283-304,1998
[13] R.Ng and J.Han. Efficient and effective clustering method for spatial data mining. In Pro.1994 Int.Conf .Very Large Data Bases,pages144-155,1994.
[14] M. Ester, H.-P.Kriegel, J.Sander, and X.Xu. A density-based algorithm for discovering clustering in large spatial database with noise. In Proc. 1996 Int. Conf. Knowledge Discovering and Data Mining.1996266 - 231.
[15] Zhang T., Ramakrishnan R. and Livny M. BIRCH: an efficient data clustering method for very large databases. In Proc. ACM-SIGKDD Int. Conf. Managament of Data, page103-114, 1996.