

Application of web Metrics and Text Mining on the IRCTC Portal



Adwaith KT, Athira K, Krishnapriya S, Reuben Thomas Mathew

Abstract: *The Indian railways, a historical inheritance, the 4th largest railway network in the world by size, is an important force in our economy. The digitalization has enabled the customer to get the railway services on their finger tip. IRCTC, the subsidiary of the Indian railway offers variety of service like catering, tourism and the online ticket operation. As technology advancement is taking place, but the IRCTC has been widely criticized for many lacunas in meeting the customer needs and preference. So the study focuses on how Indian Railway needs to revamp the website to make it more contributing to customer expectations. An analysis of customer reviews, revealed that the customers experience many problems while using the IRCTC website to make their choice about the type of travel, coach preference, seat preferences, age group, payment gateways, time and date of travel, etc. The study attempts to find out the user-friendliness of IRCTC website from the point of view of the customers using four identified dimensions or variables. The analysis was done using various web metrics and a text-mining based on the customer reviews. It helped to know about the clicks rates pattern, visit rate, the various activities performed by the customer, time spent, type of device used, keywords used etc. The result shows that majority users have stated negatively towards the features and usability of the website. Based on the analysis of the study a brief summary of findings have been made and a meaningful conclusion have been obtained.*

Keywords : Customer preferences, Website, Web analytics, Website features.

I. INTRODUCTION

Indian Railways (IR) is India's national railway system functioned by the Ministry of Railways. It manages the 4th largest railway network in the world, in terms of size[3]. Indian Railways was started its operation in 1853 with its first passenger train running from BoriBandar to Thane, with 400 passengers is one of most successful ventures today. Since then, it has come a long way adding a numerous

features to the basic facility such as online ticketing and at present railway tourism. Railways holds the monopoly of the largest public transporter and the largest employer of the country by carrying more than 2.3 crore passengers daily and employing almost 13 lakh people, respectively.

IRCTC, the subsidiary of the Indian railway offers variety of service like catering, tourism and the online ticket operation of the latter, with around 5,50,000 to 6,00,000 bookings every day[2]. IRCTC has emerged as the biggest e-commerce portal in India and has to its credit a number of records in terms of number of tickets booked in a day[5]. Within a short span of its going online, it had become the largest and the fastest-growing ecommerce website in the Asia-Pacific region in 2013[5]. It began internet-based rail ticket booking through its website, as well as from the mobile phones. It also provides features like SMS facility to check PNR status and to know the Live Train Status as well. In addition to e-tickets, IRCTC also offers I-tickets that are basically like regular tickets except that they are booked online and delivered by post. Commuters on the suburban rail can also use Irctc to book season tickets through its website. It has also introduced a loyalty program called Shubh Yatra for frequent travelers. Under the Tatkal scheme, passengers who need contingent tickets can book in almost all Mail/Express trains through the Indian railways internet portal. These services have changed the face of IR over the period of time. It also has helped the user to access it from anywhere and at any time.

The evolution of the IRCTC had brought a huge transformation in the lives of the common people. The customer pain was correctly recognized by the IRCTC, by which they could cater the needs and wants of the people who travelled by trains. It eliminated the time that was spent on the ticket counters and standing in long ques by a customer. The service of IRCTC could be accessed by a customer from anywhere in the globe and in any device. But in recent times, we could identify that users had been facing problems while using the online portal from some more reviews which were found at downdetector.com[6]. Some of the other major problems that were identified by the users are logging in issues, booking problem, payment problems, refund problems, etc[1]. Looking on the last 6 months data provide at similarwebsite.com, it shows that the number of user visits have reduced which indicates there is change in the minds of the customers and they are moving towards alternative sources to book[4]. So this study makes an attempt to evaluate the current IRCTC website, design and the features which would help to understand and analyze the various problems encountered in the IRCTC Website from a common user perspective.

Manuscript received on March 15, 2020.

Revised Manuscript received on March 24, 2020.

Manuscript published on March 30, 2020.

* Correspondence Author

Reuben Thomas Mathew*, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: reubenthomas97@gmail.com

Krishnapriya S, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: kppk25@gmail.com

Athira K, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: athira97krishna@gmail.com

Adwaith KT, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: adwaith.kt@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Application of web Metrics and Text Mining on the IRCTC Portal

The study also tries to access the number of visits, the average time spend by a user and the major search keywords used by the customers for logging on to the website. From study, we can suggest measures that would help to restructure or revamp the website in a more user friendly manner and provide better quality service to the users.

II. REVIEW OF LITERATURE

The literature has given variety of instruments to improve the service quality in numerous service sectors. The passenger transportation sector like Indian Railways, has pressure from passengers to adopt an attitude of customer-oriented service quality in the sector. After realizing this, service organizations like Indian Railways targeted on the passenger's perceptions of service quality as a result of it helps the policy manufacturers in developing service methods that cause passenger's satisfaction. Gundersen, (1996) has stated that an excellent service quality leads to higher profitability to the service providers with high customer's satisfaction. Most of the studies found a significant relationship between the quality of service delivered by the marketers and customer satisfaction. Today, SERVQUAL model is frequently used to measure the service quality and it had been used in almost all the countries and in all types of industries suggested by Parasuraman[10].

Barnes and Vidgen (2001) had developed an analysis for online booking using a WebQual scale with five crucial dimensions such as assurance, reliability, tangibles, responsiveness, and empathy[9]. Similarly Khare and Handa, (2011) indicated that the service is gaining acceptability in the India, there still exists considerable scope for improvement in customer experience with regard to the various service quality parameters for the online reservation system.

The online reservation system was an extension of the passenger reservation system (PRS) and served as an interface between the uses and the passenger reservation system. Kumar,(2006) revealed that the cost of maintenance and expansion of the railway infrastructure was monumental and also disclosed that the progressive implementation served as a acceleration towards higher responsiveness to the growing passenger demands. Sahney (2010) gave an opinion that E-ticketing is popular due to advanced technology, consumer behavior both psychographic and demographic and also of its user friendliness and computer proficiency. Sahney also said that technology know-how people and those people who want to avail the tourism connected services go for e-ticketing[11]. Wolfenbarger and Gilly, (2003) came up with a scale - eTailQ with four dimensions to evaluate e-service quality. Similarly, Bauer,(2006) proposed a transaction process-based scale called eTransQual, which includes emotional and intangible items to assess the e-service quality[12].

III. OBJECTIVE OF THE STUDY

The IRCTC could bring in a new wave of online customer service in the Indian Railway but the consequence of the poor customer service provide at IRCTC was evident from the reviews. Most of the customers reviews showed a negative sign towards the existing services provided by IRCTC. In the long run, many customers would look for substitutes sites which provided better customer services and higher

satisfaction. India, which is an emerging tourist destination attracts many foreign tourist and travellers. These downward trend may affect the future usability of the website. It is the right time to bridge the existing gap and restructure the web portal for provide user-friendly services. So this study aims to study the problems faced by the user while using the IRCTC web portal.

The objectives for the research paper are:

- To evaluate the IRCTC Website, design and its features using web analytics.
- To understand and analyze the problems faced in the IRCTC Website from the user perspective.
- To suggest measures to restructure/revamp the website in a more user friendly manner.

IV. RESEARCH METHODOLOGY

The current empirical study is mainly based on secondary data. The secondary data used for the study was collected through various online sources.

The sample size of reviews taken for the purpose of the study was 300 reviews. The population covers students, salaried, businessmen, professionals and others who are frequent users of IRCTC. The collected data has been analysed by using Voyant tools. The reviews were taken from the user who have already used the IRCTC website and its services.

To get a better understanding about the research, we adopted certain dimensions or variables to know about the service quality provide by the IRCTC.

The identified variables are

- Website – Website variable includes the design, usability and features of the website .
- Payment – Payment includes the refund mechanism and the processing of payment for the ticket.
- Time – Time includes the loading time of the website.
- Grievance handling – It is about the customer services such as tatkal booking, the handling of the customer queries.

Based on the variables (Fig.1), we would measure the quality of the IRCTC web portal. As to get a further understanding of the study, we use 2 methods for the analysis of the data:

a. Web analytics

Web analytics is that the mensuration, collection, analysis and reportage of internet knowledge for functions of understanding and optimizing internet usage. It will be used as an instrument for business and research, and to assess and improve the effectiveness of a website[7]. Internet analytics provides info concerning variety of tourists to a web site and also the number of page views. It helps gauge traffic and recognition trends that is beneficial for research. The IRCTC official analysis link wasn't offered so analysis that was wiped out a secondary supply was wont to track the shoppers activities.

b. Text mining

Text mining is that the method of etymologizing high-quality info from text. High-quality info is usually derived through the making of patterns and trends through suggests that like applied mathematics pattern learning[8].

Text mining typically involves the method of structuring the input text etymologizing patterns at intervals the structured knowledge, and eventually analysis and interpretation of the output. For this study, the reviews were taken from mouthshut.com and that were text deep-mined with the assistance of voyant tools.

This could facilitate to spot the terms that occur in an exceedingly higher frequency. These words square measure the key terms that a customers uses to convey their feelings concerning the web site.

V. DATA ANALYSIS & INTERPRETATION

The data for research study is taken from secondary data source available in the internet. The graphs below are taken from an existing analysis which was performed on the Irctc website which was available online. The reviews were taken from mouthshut.com. The findings and interpretation of the data is done from the reviews of the users. There are positive as well as negative reviews.

The graph in Fig.2 shows the number of visits on the IRCTC website (April 2019 - September 2019). As it is evident from the graph, the user population has reduced the usage of the IRCTC and Services through desktop and mobile. Last 6 months' visits show a downward trend. The number of visits have come down from 44M in May 2019 to 32M in September. This clearly shows a fallback in the mindset of the users and the transaction and service charges have increased a lot which may be taken as the reasons for the decline.

Fig.3 shows the average visit rate duration on the website and the bounce rate of the users. The site has a low bounce which is excellent. And the average time which a user spends is approx. 6.68 minutes. The time which a user spends on the website can also be looked as the time taken by a customer to book the ticket and make the payment or browse through the website. It varies according to customers depending upon the server, type of system used and the mode of payment. It is also evident from the visits that there is a fall in the number of people using the website.

Fig.4 depicts that traffic attracts towards the website and by what source they get on to the website. The major share i.e. about 63.54% goes to direct entry to the website, then followed by search using various keywords which is about 32.72%. There is a less traffic that is coming from referrals, social, mail and display. The customer prefers to login directly to the website when in need. So the attraction of traffic from other sources is low as the service is the only for the traveller. Also there is an increase in the competitor websites which provide the similar service like railyatri, RailRestro, YatraChef.

The graph shown in Fig.5 is the top keywords which are used by the users. They usual search these keywords in order to access the website. So 63.95% of the users search for the word "IRCTC" for logging in to the site. Rest is divided among the irctc login followed by irctc then irctc account login.

The diagram in Fig.6 is a word cluster which was formed from the text mining of the reviews taken from the internet. The reviews were classified into positive and negative terms which convey the similar meaning using MS Excel. These are words which are commonly used by the users convey their feelings. The words larger in size shows the highest

frequency of occurrence in the reviews. Respondents used the words "easy booking", "refund problem", "booking problem", "load time", "tatkal booking issues" in their reviews about the service provided [refer table 1]. These words have a close visibility to the website and its features. These terms evolve around the customer mindset while going through or while looking for the IRCTC website.

The reviews of the users (Fig.7) shows more of a negative impact towards the website. The user reviews show less of positive side of the website. The server error, refund problems and tatkal booking were the most identified problems from the reviews. Around 33% of the collected reviews show a positive response. Almost 67% of the reviews have a negative impact.

Based on the positive reviews (Fig.8), a classification of reviews was done using the identified variables. The reviews were sorted under each variable. 46% of the positive terms were related to the website, and 35% of the reviews were related to the payment. Rest of the reviews were about time(7%) and grievance handling(12%).

Similarly based on the negative reviews (Fig.9) were also classified depending on the variables. The negative terms constituted 37% about the website, 29% about grievance handling, 18% about payment and 16% about time.

VI. FINDINGS

IRCTC has created a large impact on the user and made it simple for them to travel and book tickets on-line. But we could pinpoint that there were many problem arising from the users point of view. The study of the reviews revealed that the procedure of booking ticket has problem and also the website is slow to load [table 1]. The server needs to have higher broadband which will enable to provide service at a higher pace which will remove difficulties like login problem server error. There are payment problems and as well the refund mechanism is not proper. The study shows that about 63.54% of the user population enter into the website directly and about 32.72% of the user population use various keywords to enter into the IRCTC portal[Fig.4]. Almost 67% of the reviews state a negative impact towards the website. It is also understood that majority of the users search for the word "IRCTC" in order to log into the site and also the average time consumed by the user within the IRCTC portal is around seven minutes[Fig.2]. However, the extent of visits throughout the last 6 months shows a downward trend which indicates a fall in the number of users. This also indicates a shift of the customer mindset towards the website. The overall appearance of the website needs a make over which make help to attract more user friendly features to the website.

VII. RECOMMENDATION, SUGGESTION AND CONCLUSION

A. RECOMMENDATIONS AND SUGGESTIONS

IRCTC website needs to adopt the many changes to make the website more user friendly.

First and foremost, payment gateways should be made easily accessible because many times payments get stuck at the final stage of the booking process and the refund policies must be monitored. There must a single window for the processing of the payments.

Then, the login time should be increased as it gets timed out fast and the customer needs to login again. Along with that it must improve the logging in procedure as entering the captcha may cause mismatch and lead to time loss. IRCTC must think of adding an auto save option that would help to save the data in case the user lands in an any error. The increased bank service charge has to be reduced or eliminated so that the amount of traffic into the website can be increased. The Website speed and performance has been pointed to as a matter of concern, so there needs to be a change in the design and connection to a high speed server must be brought in for better performance. IRCTC should provide fast retrieval of information to save more time. The booking facility availability for 24 hours to make it more convenient. The online ticket booking is a complex system. Thus, efforts should be made to reduce the number of steps in the booking process and the number of pop-up ads should be removed. The site maintenance should be made more faster as the customer can access the site at any point of time. Providing offers to the daily commuters may also help in to add more traffic and provide better service. And also more awareness about the usability of the website should be provided to the customers so that they can get a better experience of the website in future.

B. CONCLUSION

The research paper was designed to focus on the user experience while using on the IRCTC online portal. The study reveals that IRCTC online services provided by Indian Railways is just satisfactory and still needs to improve. The study helped to draw the following conclusions:

- Downfall in the number of online users was evident from the study.
- There were many concerns raised by the users while using the online portal such as the refund issue, booking problem etc.
- The Website content and features are good but some restructuring is appreciable by the user community.
- The Website speed and performance has been pointed as a matter of concern.
- Services like tatkal booking and refund mechanism were highlighted as a major concern from the users.

As it is used by a large population of this country, the website and its features needed to be restructured or upgraded using the latest technology so that it could cater according to the needs and preferences of the customers. This would enable the customers to experience a better quality service and also would help in developing the Indian Railways to a large extend.

ACKNOWLEDGMENT

Authors would like to thank Dr. Rejikumar G, Associate Professor, Department of Management, Kochi for his esteemed guidance, inspiration, expert observations and timely advice throughout the study.

REFERENCES

1. Irctc.co.in,ReviewsandRatings.(n.d.).Retrievedfrom <https://www.mouthshut.com/websites/Irctc-co-in-reviews-925052665> W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.
2. Indian Railway Catering and Tourism Corporation. (2005, June 20) Retrievedfrom https://en.wikipedia.org/wiki/Indian_Railway_Catering_and_Tourism_Corporation.
3. IndianRailways.(2004,May 15). Retrieved from https://en.wikipedia.org/wiki/Indian_Railways
4. How big is IRCTC? (2016,February 16). Retrieved from <https://www.quora.com/How-big-is-IRCTC>.
5. IRCTC.co.in.(2019,September).Retrievedfrom <https://www.similarweb.com/website/irctc.co.in>
6. [6] IRCTC down? Current problems and outages. (n.d.). Retrieved from <https://downdetector.in/status/irctc/>
7. Web analytics. (2005, August 17). Retrieved from https://en.wikipedia.org/wiki/Web_analytics
8. Text mining. (2003, September 14). Retrieved from https://en.wikipedia.org/wiki/Text_mining
9. Barnes, S. J., Vidgen, R., (2001), An evaluation of cyber-bookshops, The WebQual method. International journal of electronic commerce,
10. Parasuraman, A., V. Zeithaml and L. Berry, 1988. SERVQUAL: a multiple item scale for measuring consumer perceptions of service quality. Journal of Retailing,
11. Sahney, S., Banwet, D.K., and Karunes, S., (2004), A SERVQUAL and QFD approach to total quality education, A student perspective, International journal of productivity and performance management,
12. Bauer, H. H., Falk, T., Hammerschmidt, M., (2006), eTransQual, A transaction process-based approach for capturing service quality in online shopping. Journal of business research
13. Bikramjit, Singh, Hundal., and Vikas, Kumar., (2015), Assessing the service quality of Northern railway by using SERVQUAL model. Pacific business review international, 1
14. Parasuraman, A., Zeithaml, V.A. and Berry, L.L., (1985), A conceptual model of service quality and its implications for future research, Journal of marketing.
15. Irctc.co.in,September2019,Retrievedfrom <https://www.similarweb.com/website/irctc.co.in>

APPENDIX

Table 1. Keyword Frequency

1	Poor service	35
2	Loadtime	30
3	Easybooking	28
4	RefundProblem	17
5	Loginproblem	16
6	Tatkalproblem	15
7	Bookingproblem	14
8	Paymentproblem	14
9	Worstwebsite	13
10	Slowsrver	9

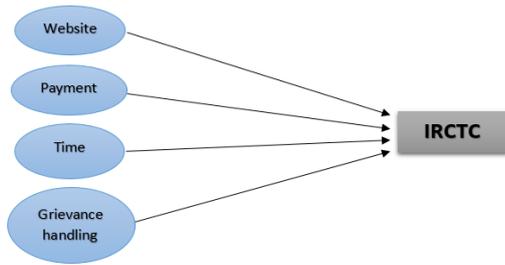


Fig.1. Variables



Fig.6. Word cluster made from the reviews

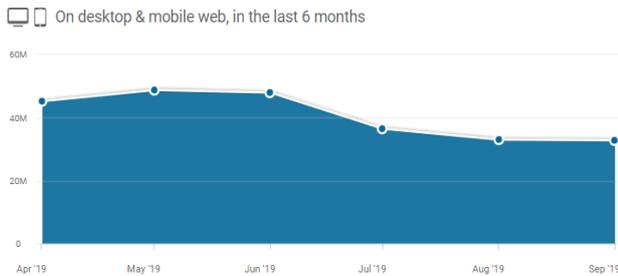


Fig.2. Total number of visits in the last 6 months
Source: Similar web



Fig.3. Customer engagement in the website
Source: Similar web

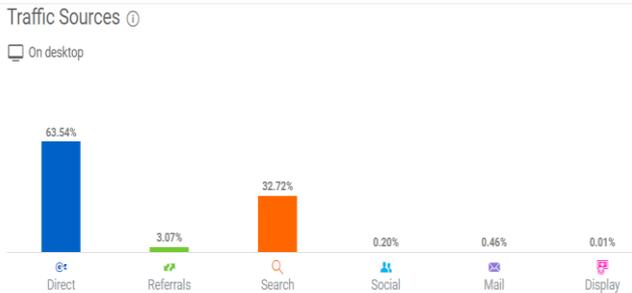


Fig.4. Traffic source on the website
Source: Similar web



Fig.5. Keyword search used for login
Source: Similar web

Reviews of users

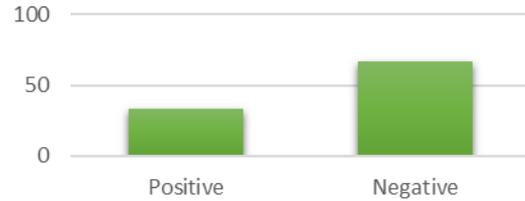


Fig.7. Reviews of users divided into positive and negative

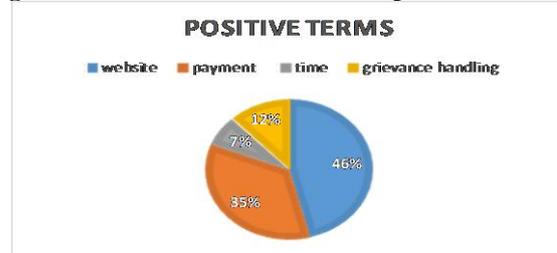


Fig.8. Classification of positive reviews in terms of variables

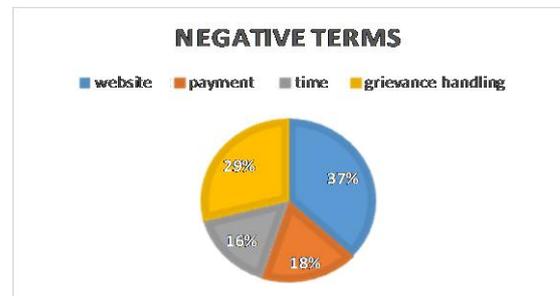


Fig.9. Classification of Negative reviews in terms of variables

AUTHORS PROFILE



Reuben Thomas Mathew is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. He wishes to make his area of expertise in Marketing and Finance. After securing his Bachelor's degree in BBA, he turned his attention towards world of Finance and Marketing to showcase his business and leadership skills. His leadership traits was seeded and developed during academic carrier by being responsible for coordinating various academic activities. He has actively participated in many well-known business skill development programs across the state during his academic carrier. His personal aim is to enter the Business world and pursue his passion of becoming an entrepreneur



Application of web Metrics and Text Mining on the IRCTC Portal



Krishnapriya S is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. She wishes to make her area of expertise in Marketing and Analytics. After securing her Bachelor's degree in BSC Physics, she turned her attention towards world of Marketing and Analytics to showcase her business and leadership skills. Her leadership traits was seemed and developed during academic carrier by being responsible for coordinating various academic activities. She has actively participated in many well-known business skill development programs across the state during her academic carrier. . Her personal aim is to enter the teaching world and pursue her passion of becoming a lecturer.



Athira K is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. She wishes to make her area of expertise in Marketing and Analytics. After securing her Bachelor's degree in BSC Mathematics, she turned her attention towards world of Marketing and Analytics to showcase her business and leadership skills. Her leadership traits was seemed and developed during academic carrier by being responsible for coordinating various academic activities. She has actively participated in many well-known business skill development programs across the state during her academic carrier. Her personal aim is to enter the teaching world and pursue her passion of becoming a lecturer.



Adwaith KT is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. He wishes to make his area of expertise in marketing and operations. After securing his Bachelor's degree in mechanical engineering, he turned his attention towards world of marketing and operations to showcase his business and leadership skills. His leadership traits was seemed and developed during academic carrier by being responsible for coordinating various cultural and academic activities. He has actively participated in many well-known business skill development programs across the state during his academic carrier. His personal aim is to enter the Cineworld and pursue his passion of becoming a Filmmaker.