Big Data Analytics Based Recommender System for Value Added Services (VAS)

Inderpreet Singh[1], Karan Vijay Singh[1(✉)], and Sukhpal Singh[2]

[1] Engineer-Product Development, Mahindra Comviva, Gurgaon, India
inderpreet.singh993@gmail.com, karanvijay.singh93@gmail.com
[2] Computer Science and Engineering Department,
Thapar University, Patiala, Punjab, India
ssgill@thapar.edu

Abstract. The increasing number of services/offers in telecom domain offers more choices to the consumers. But on the other side, these large number of offers cannot be completely looked by the customer. Hence, some offers may pass unobserved even if they are useful for the particular kind of customers. To solve this issue, the usage of recommender systems in telecom sector is growing. So, there is need to notify the customer about the offers which are made on the basis of customer interests. The recommender system is based on demand or interest of consumer. In this paper we proposed a Big Data Analytics based Recommender System for Value Added Services (VAS) in case of telecom organizations so that they could gain profitability in the market by generating customer specific offers.

Keywords: Analytics · Recommendation · Big data · Telecom · Segmentation · Hybrid model

1 Introduction

In recent years, wireless data traffic has been increasing exponentially. This traffic is mainly driven by the usage of smart phones and increasing penetration of Internet services in the world. Nano and Pico cells have become a reality [1] due to increase in demand of smart phones and decrease in size of them. Thus resulting in a rapid increase not only in number and type of smart phones but also in the services needed to track and administer to provide high quality services to the customers. As most of data in the world is now generated by the smart phones and as telecom network is the carrier of this data, so telecom operators are experiencing the explosion of data making it difficult for them to decide about data needed for predicting customer behaviour.

Recommender systems are used to predict the liking that a customer will have for a particular item [2]. They are subclass of information filtering systems i.e. systems which remove unwanted data from the data stream. A recommendation produced by the recommender system is the output of analysis performed on data sets of customer's preferences. Recommendations in case of web services are difficult because of unavailability of customer's exact location and also service history of customer cannot be easily

K. Deep et al. (eds.), *Proceedings of Sixth International Conference on Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 547,
DOI 10.1007/978-981-10-3325-4_14

defined. But in the case of telecom, this problem can be pacified as smart phone becomes main access point for customer and its usage can be stored and analysed by the company to get ingenious insights about customer behaviour. The availability of customer data in case of telecom domain makes easier for performing different analysis which results in suggesting more worthy services to a target customer.

Using recommendation engines, giving customers the services of their choice and not loading them with the unnecessary services, companies can easily gain customer's confidence and improve customer experience [15, 16]. There are different ways that can be put to use in a recommendation system namely Collaborative filtering, Content Based filtering, and also hybrid approaches. The most commonly used, Collaborative filtering technique makes a prediction for the target customer based on gathering information of (older) customer's with similar interests and activities. Finding likeness between customers effectively is the pivotal part of Collaborative Filtering technique. The challenge occurs when there is only minuscule data available of older customer's or customer's with identical interests leading to cold starting problem [17, 18]. On the other hand, Content based technique uses the distinct characteristics of the product/item that the customer had earlier bought or had shown a liking, so as to give recommendations to the customer of the products with identical characteristics. The combination of the above two leads to hybrid approach that can be executed in several ways like making Content based and collaborative filtering predictions individually and then utilizing results of both the techniques to give final recommendation to the customer. Also this approach can also be used to eliminate cold start problem [19].

The main things playing roles behind the popularity of big data analytics are people skilled in analytics, efficient analytical tools, robust data infrastructure, fact-based decision-making strategy, association between IT approaches, durable committed sponsorship and clear business requirement [15, 16]. An efficient way of analysing and retrieve data in analytics makes it as popular decision oriented data management. In traditional data warehouses and databases, business value cannot produced from stored data. With the help of this new technology, data can create incredible value after storing it efficiently [17, 18]. A new products, techniques and technologies are emerged for data analysis which can be easily used for big data analytics, such as appliances, in-database analytics and in-memory analytics [19].

In this paper, we are giving two recommendation engine models. One, which generates the recommendations on the basis of customer segmentation and metadata details; and other which generates the recommendations on the basis of customer segmentation and service comparison. Former model is used for services like ringtone, games recommendation etc. While the latter is used for services like astrology, cricket, jokes etc.

2 Related Work

Data generated from smart phones is very complex due to which it poses difficulty in analysis. It has problems like heterogeneity and noisiness associated with it. In [3], authors have made a mobile recommender system that helps in providing profitable routes for taxis. It takes present location of taxi, time status and operational status i.e.

with or without passengers. While in [4], authors give example of a recommendation system which suggests appropriate information depending on the customer's situation and interests. Recommendation systems are now utilized in different areas like: In [5], authors discusses about news recommender system that pushes news articles to smart phone customers based on customer's contextual information as well as news content. Where customer's information needs are estimated by the use of Bayesian network technique. While in [14], author has proposed a real-time location-tagged contents recommender system based on smart phone social network. This system locates a customer with the use of GPS (Global Positioning System), and then applies filtering methods. In addition to the use of GPS, the correct location of the customer can also be found with the help of Wi-Fi positioning system (WPS) to get high precision [6]. In [7], authors have developed a smart phone recommender system for indoor shopping based on positioning approach by using received signal patterns of smart phones. It eliminates the disadvantages of existing positioning technologies. Apart from these, there are also recommender systems for movies [8], personalized blog content [9], experts [10], collaborators [11], financial services [12] and Twitter pages [13].

3 Various Types of Data Needed

To generate recommendations, recommendation engine needs certain sets of data. Depending on what type of recommendation needs to be generated, recommendation engine will use specific set. In this paper, we will be discussing about the recommendation engines which will be used for recommending Ringtones and Value added Services (VAS). Consumer application form details are the details that consumer fills while registering for a particular service or set of services provided by the telecom company. It contains details about the consumer's basic information. The set of CAF details which the company has is mentioned in below Table 1.

Table 1. Customer application form (CAF) details

CAF details		
S. no	Data	Purpose of data
1	Name	Basic Information
2	Age	Differentiate customer on the basis of Age
3	Sex	Differentiate customer on the basis of sex
4	Marital Status	Finding whether the customer is Single/Married
5	Address	Get the customer's home address
6	City	To know about the customer's home city

Company level details are the details of consumer transactions which the company stores in their databases (Table 2). Based on these details and metadata of ringtone, recommendation engine generates the ringtone recommendations for the customer. In case of other VAS (Value Added Services), recommendations generated are based on the segment to which consumer profile belongs rather than on the basis of metadata as metadata details are not associated for these kind of services. The set of company level details which the company maintains in their databases is given below in the table.

Table 2. Company level details

Company level details		
S. no	Data	Purpose of data
1	Talk time Value	Total Duration of calls used for generating recommendations on the basis of call plan
2	Data Usage	Details of data used for generating recommendations on basis of data packs
3	No. of SMS Sent/Receive	Details of messages sent/received for generating recommendations on basis of messages packs
4	Customer Location	To know about the customer's present location
5	Most Frequently Called Numbers	To give customer customized calling plans
6	Handset Type	Various types of services supported by customer's handset
7	Metadata of VAS Services	Generate recommendations on the basis of customer's VAS profile
8	Metadata of Customer Ringtone	Generate recommendations on the basis of customer's ringtone

Metadata details of ringtone as shown in Table 3, is used in case of recommending caller tones to the customers. In this case, metadata of customer's past and present caller tones is used to know to which segment the customer belongs, and on the basis of that segment, recommendation is generated for the customer.

Table 3. Metadata details

Metadata details		
S. no	Data	Purpose of data
1	Type	Type of Ringtone e.g. pop, jazz etc.
2	Film/Album	Film/Album to which Ringtone belongs
3	Composer	To generate recommendations on basis of composer
4	Singer	To give recommendations on basis of singer
5	Director	To generate recommendations on basis of music director
6	Language	Language of the ringtone

4 Model

Figure 1 Represents the architecture of the recommendation engine which works on the basis of customer segmentation and metadata comparison.

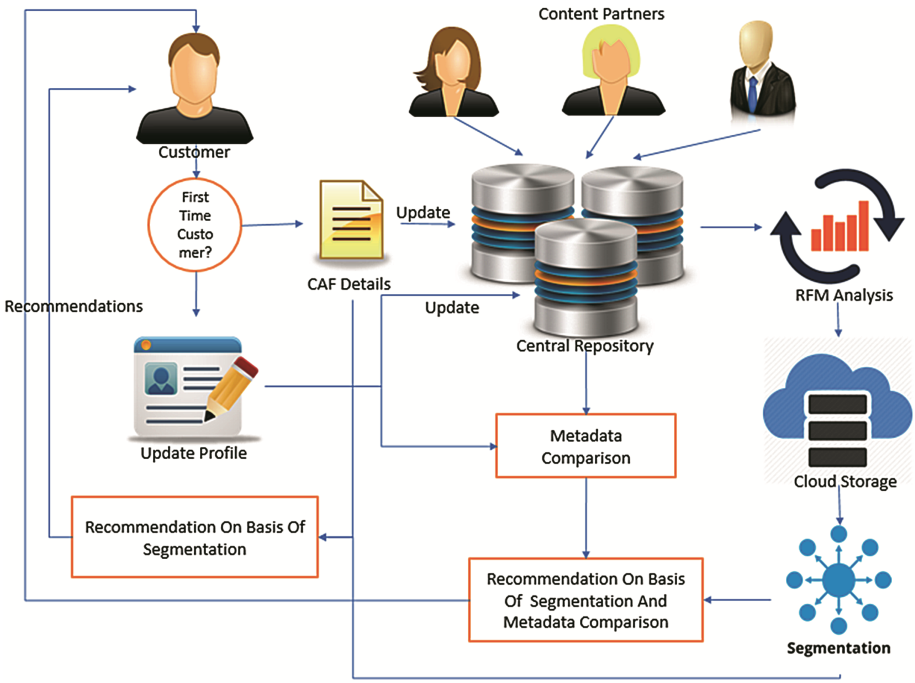

Fig. 1. Architecture of model based on customer segmentation and metadata comparison

4.1 Customer Profile

A new customer profile gets created when the customer registers himself/herself for the services of the telecom company. Customer profile contains basic details about the customer like name, age, gender etc. and these are the details which the customer gives himself by filling consumer application form (CAF). Then, it is checked whether the customer is first time customer or he/she wants to change the caller tone of smartphone. In case of former, recommendation engine generates the recommendations on the basis of customer segmentation i.e. the segment to which the customer's profile belongs. But in case of lateral, recommendations are generated on the basis of metadata and customer profile details.

4.2 Recommendation on Basis of Customer Segment

When the customer register himself for the first time, then there are no metadata details present for that customer, then the recommendations are generated on the basis of matching the customer profile parameters to the profiles of already existing customers in the repository. The profile which matches maximally is used to know the segment to which the customer profile belongs and on basis of that segment, the recommendation is generated.

4.3 Updating Profile and Metadata

If the already registered customer wants to change the caller tone of his/her smartphone, then firstly the metadata details of the customer's record is changed accordingly. So, customer's profile gets updated whenever customer wants to change the caller tone.

4.4 Metadata Comparison

In this step, recommendation engine will compare the metadata details of customer profile with the metadata of the other profiles i.e. we will be calculating the similarity index of the customer metadata details with other customer's metadata details and the profile corresponding to the maximum similarity index will be then looked up to see to which segment it belongs on basis of which recommendations will be made.

The similarity index between two customer's profiles is represented by a number between −1.0 and 1.0. The possibility of customer liking/selecting particular ringtone will be between −1.0 and 1.0. Similarly, in case of not liking/selecting the number will be between −1.0 and 1.0. For finding similarity index, we will have two sets corresponding to each customer. One corresponding to the customer liking/selecting the caller tune and other for not selecting/liking the ringtone. According to Jaccard's formula [20], the similarity index is calculated as follows

$$J(X, Y) = |X \cap Y| \div |X \cup Y|$$

The calculation involves the division of the total number of common elements in both sets by the total number of the elements in both sets (only counted once). The Jaccard index of two similar sets will always be 1, while for two sets with no common elements will always yield 0. Jaccard index for two profiles on the basis of liking of each parameter is,

$$J(X, Y) = |S1 \cap S2| \div |S1 \cup S2|$$

Now as two customers selecting same ringtone is similar, then two customers not selecting the same ringtone are also similar. So, by changing above equation we get,

$$J(X, Y) = (|S1 \cap S2| + |NS1 \cap NS2|) \div (|S1 \cup S2 \cup NS1 \cup NS2|)$$

I.e. instead of considering same selection we also have taken into the account the deselection. In denominator, we have taken the total number of selection/deselection that customer has made. Here, we have considered the customer selection or deselection in independent sort of way. But what if customer likes ringtone but other customer doesn't and vice-versa. To take this thing into account we again have to modify the equation as,

$$J(X, Y) = (|S1 \cap S2| + |NS1 \cap NS2| - |S1 \cap NS2| - |NS1 \cap S2|) \div (|S1 \cup S2 \cup NS1 \cup NS2|)$$

Now this equation will give 1.0 if two customer's profiles have same selection/liking for caller tones and −1.0 if two customers have deselection/disliking for caller tones.

4.5 Recommendation on Basis of Customer Segmentation and Meta-data Comparison

RFM analysis of customer transaction information plus the metadata of profile that matched with the customer's profile gives us the segment to which the customer profile belongs. RFM is method for analysis of market which is used to examine which customers are best ones by examining how recent the customer has made purchase (Recency), how often he/she purchases (Frequency) and how much the customer spends on the purchase (Monetary). It is based on the fact that "80% of business comes from 20% of the customers". Customers are then given ratings on the basis of these three input parameters by the telecom organisation. The RFM score and metadata details matched profile gives the segment to which the customer profile belongs. On the basis of which the recommendations are generated for the customers i.e. Hybrid recommendation systems [21].

4.6 Content Provider

Content providers belonging to different partners are responsible for uploading the content to the telecom company's central repository on which RFM analysis is performed and output is stored on the central cloud storage, so that results can be accessible anytime and from anywhere. Based on these results, customer segmentation is done.

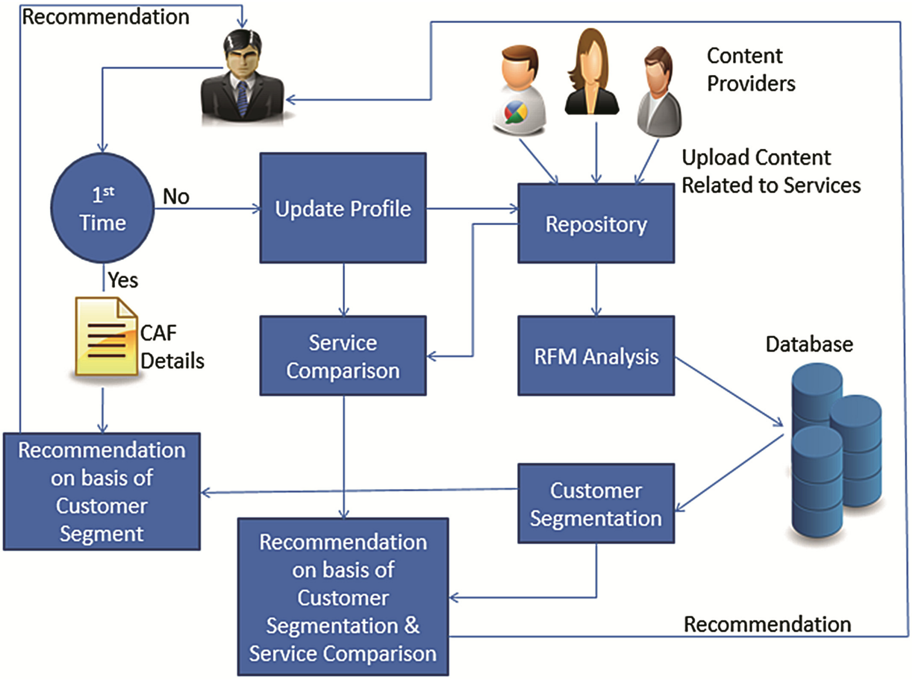

Fig. 2. Architecture of model based on customer segmentation and service comparison

The similar type of model can be used in case of Games i.e. based on customer profile, recommendations can be given whether one likes action, racing, puzzles, sports games etc. While new customer are given recommendations based on similar CAF details of other customers who have already subscribed to games, old customers get recommendations based on metadata of games that they have subscribed along with the segment of customer profile to which they belong. Figure 2 represents the architecture of the recommendation engine which works on the basis of customer segmentation and Service comparison.

In case of other VAS related services, if a new customer wants to subscribe to any service like Astrology, Cricket, Jokes etc., he/she is recommended based on customer profile segmentation of other customers with similar CAF details as we have no idea about the interests of the new customer. Also there is no meta-data related to these type of services. On the other hand, old customers are given recommendations based on their updated profile and RFM analysis of transaction information of the customer which gives the segment to which the customer profile belongs.

5 Conclusion

This paper discussed the approach that how recommendation engine could collect the data, analyse it and generate recommendations on basis of it. In it we have explained two types of recommendation engine. One, which uses the metadata details and customer segmentation to generate recommendations and other which uses services comparison along with customer segmentation. Recommendation not only helps telecom companies to target their customers effectively, but also help the customers in showing the content which is relevant to them.

6 Future Scope

The presented recommendation engines can be used to help the telecom companies in recommending appropriate services to the customer. Hence, gaining competitive edge in the market. Some open issues that need special attention in future research are,

- Approach to integrate above recommendation engine with the systems already used without increasing workload of them.
- To have more input parameters on basis of which recommendation could be given.
- To give real time recommendations to the customer on the basis of location, time etc.

References

1. Ericsson, L.M.: More than 50 billion connected devices. White Paper (2011)
2. Ricci, F., Rokach, L., Shapira, B.: Introduction to Recommender Systems Handbook. Springer, US (2011)
3. Ge, Y., Xiong, H., Tuzhilin, A., Xiao, K., Gruteser, M., Pazzani, M.J.: An energy-efficient mobile recommender system (PDF). In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 899–908. ACM, New York (2010). Accessed 17 Nov 2011

4. Bouneffouf, D.: Following the customer's interests in mobile context-aware recommender systems: the hybrid-e-greedy algorithm. In: Proceedings of the 2012 26th International Conference on Advanced Information Networking and Applications Workshops (PDF). LNCS, pp. 657–662. IEEE Computer Society (2012). ISBN: 978-0-7695-4652-0 [dead link]
5. Yeung, K.F., Yang, Y.: A proactive personalized mobile news recommendation system. In: 2010 Developments in E-systems Engineering (DESE), pp. 207–212. IEEE (2010)
6. Danalet, A., Farooq, B., Bierlaire, M.: A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures. Transp. Res. Part C Emerg. Technol. **44**, 146–170 (2014). doi:10.1016/j.trc.2014.03.015
7. Fang, B., Liao, S., Xu, K., Cheng, H., Zhu, C., Chen, H.: A novel mobile recommender system for indoor shopping. Expert Syst. Appl. **39**(15), 11992–12000 (2012)
8. Colombo-Mendoza, L.O., Valencia-García, R., Rodríguez-González, A., Alor-Hernández, G., Samper-Zapater, J.J.: RecomMetz: a context-aware knowledge-based mobile recommender system for movie showtimes. Expert Syst. Appl. **42**(3), 1202–1222 (2015)
9. Chiu, P.-H., Kao, G.Y.-M., Lo, C.-C.: Personalized blog content recommender system for mobile phone customers. Int. J. Hum. Comput. Stud. **68**(8), 496–507 (2010)
10. Buettner, R.: A framework for recommender systems in online social network recruiting: an interdisciplinary call to arms. In: 47th Annual Hawaii International Conference on System Sciences, Big Island, Hawaii, pp. 1415–1424. IEEE (2014). doi:10.13140/RG.2.1.2127.3048
11. Chen, H., Gou, L., Zhang, X., Giles, C.: Collabseer: a search engine for collaboration discovery. In: ACM/IEEE Joint Conference on Digital Libraries (JCDL) (2011)
12. Felfernig, A., Isak, K., Szabo, K., Zachar, P.: The VITA financial services sales support environment. In: AAAI/IAAI 2007, Vancouver, Canada, pp. 1692–1699 (2007)
13. Goel, A., Gupta, P., Sirois, J., Wang, D., Sharma, A., Gurumurthy, S.: The who-to-follow system at Twitter: strategy, algorithms, and revenue impact. Interfaces **45**(1), 98–107 (2015)
14. Kwon, H.-J., Hong, K.-S.: Personalized real-time location-tagged contents recommender system based on mobile social networks. In: IEEE International Conference on Consumer Electronics (ICCE), pp. 558–559. Las Vegas (2012)
15. Singh, S., Chana, I.: EARTH: energy-aware autonomic resource scheduling in cloud computing. J. Intell. Fuzzy Syst. **30**(3), 1581–1600 (2016)
16. Singh, S., Chana, I.: Resource provisioning and scheduling in clouds: QoS perspective. J. Supercomput. **72**(3), 926–960 (2016)
17. Singh, S., Chana, I.: QoS-aware autonomic resource management in cloud computing: a systematic review. ACM Comput. Surv. (CSUR) **48**(3), 42 (2016)
18. Singh, S., Chana, I.: QRSF: QoS-aware resource scheduling framework in cloud computing. J. Supercomput. **71**(1), 241–292 (2015). Springer
19. Singh, S., Chana, I., Singh, M., Buyya, R.: SOCCER: self-optimization of energy-efficient cloud resources. Cluster Comput. **19**, 1–14 (2016). doi:10.1007/s10586-016-0623-4. Springer
20. Jaccard, P.: Etude comparative de la distribution florale dans une portion des Alpes et du Jura. Impr. Corbaz (1901)
21. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005). doi:10.1109/TKDE.2005.99