11-2020

Base-package recommendation framework based on consumer behaviours in IPTV platform

Kuruparan SHANMUGALINGAM

Ruwinda RANGANAYANKE

Chanka GUNAWARDHAHA

Rajitha NAVARATHNA

Base-Package Recommendation Framework Based on Consumer Behaviours in IPTV Platform

Kuruparan Shanmugalingam
Singapore Management University
Singapore
kuruparan4@gmail.com

Ruwinda Ranganayake
University of Moratuwa
Colombo, SriLanka
ruwinda90@gmail.com

Chanaka Gunawardhana
University of Moratuwa
Colombo, SriLanka
jcpgunawardhana@gmail.com

Rajitha Navarathna
OCTAVE - John Keells Group
Colombo, SriLanka
rajitha.jkh@keells.com

***Abstract*—Internet Protocol TeleVision (IPTV) provides many services such as live television streaming, time-shifted media, and Video On Demand (VOD). However, many customers do not engage properly with their subscribed packages due to a lack of knowledge and poor guidance. Many customers fail to identify the proper IPTV service package based on their needs and to utilise their current package to the maximum. In this paper, we propose a base-package recommendation model with a novel customer scoring-meter based on customers behaviour. Initially, our paper describes an algorithm to measure customers engagement score, which illustrates a novel approach to track customer engagement with the IPTV service provider. Next, the content-based recommendation system, which uses vector representation of subscribers and base packages details is described. We show the significance of our approach using local IPTV service provider data set qualitatively. The proposed approach can significantly improve user retention, long term revenue and customer satisfaction.**

***Index Terms*—Feature engineering, Collaborative filtering, Content filtering, Machine learning, Clustering, Customer scoring, Customer Churn, Recommendation system**

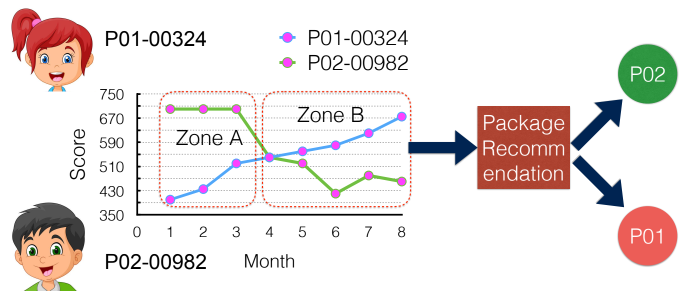

Fig. 1. **Overview:** Our framework consists with (a) A score-meter which provides score for each subscriber in every month based on their behaviours and (b) recommendation model, which can be use to recommend downgrade, upgrade or no action to their current package. Example: Two subjects in their base packages, package-01 (P01-00324) and package-02 (P02-00982) is shown here with their behaviours for few months. Initially, each user get a score based on their base package. Here, P01-00324 starts with 350 and P02-00982 starts with 700 score. Based on their monthly activities such as newly subscribed packages and VODs, payments and monthly watching behaviours, both subjects get a score. During the Zone A, subject P01-00324 is highly active compared to P02-00982 based on their scores. P01-00324 keep continuing the service in positive way and P02-00982's behaviour effect his scores to drop. Based on this behaviour, our recommendation model provides package recommendation for both subjects and recommend P01-00324 to upgrade to P02.

I. Introduction

Given enough number of consumer ratings for consumer services, one can build a recommendation system to recommend variety of services based on their previous engagement with the service. Often such recommendation systems such as Pandora[1],Amazon[2] and Netflix[3] are based on content-based and collaborative filtering methods. These systems widely use for enhance customer experience, reduce user churn and to increase profits [14].

With the development of modern day television, Internet Protocol TeleVision (IPTV) services can provide live television streaming, time-shifted media and Video On Demand (VOD) services. Typically, a registered user in a IPTV service can subscribe many TV channels. Due to large set of available TV channels, many customers don't have a clear inside about channels, content in each channel and even the IPTV service package, that they have already subscribed. Generally, each service package has a different intention with different number of channels and contents. However, majority of customers fail to identify whether they subscribe to a proper package based on their needs.

In this paper, we propose a package-level recommendation system based on user consumer behaviour. We measure consumer behaviour using a novel algorithm, which provides a user rating based on the base package details, customers subscribed package's consistency, past behaviour of the customer, payment history and current watching hours etc... The proposed approach can significantly improve user retention and provide positive experience.

A. Contribution

We used one of the local IPTV service providers data in Sri Lanka for our research. They provide local and foreign TV channels to their subscribers to engage in 9 different base

First author is Kuruparan
[1]www.pandora.com
[2]www.amazon.com
[3]www.netflix.com

TABLE I
DISTRIBUTION OF IPTV SERVICE PROVIDER'S BASE PACKAGES

Base Package	Channel Distribution												Total Channels	Price (Rs)
	Local	Foreign	Sports	News	Movie	Infotainment	Travel	Kids	Music	Tamil	Religious	Learning		
P-01	30	16	11	10	9	7	4	8	4	6	8	8	121	1999
P-02	30	13	5	10	9	3	4	6	4	6	8	8	106	1799
P-03	30	12	5	10	4	3	3	6	4	6	8	8	99	1399
P-04	30	9	8	5	9	0	1	1	4	2	8	7	84	1399
P-05	30	8	3	10	2	4	1	4	4	2	8	9	84	1199
P-06	30	8	5	10	1	3	2	4	4	2	8	7	83	999
P-07	30	7	4	9	1	2	1	3	4	2	8	7	77	799
P-08	15	9	1	7	2	0	3	3	1	13	6	0	60	799
P-09	15	5	0	7	1	0	0	3	0	9	6	0	46	590

packages[4]. Available channels are categorized into 13 genres. They are Local variety, Foreign variety, News/Business, Movie, Sports, Infotainment, Travel and Living, Kids Learning and Educational, Music Tamil and Religious. Full breakdown is given in Table I with the package price in local currency.

When a new user register with the service, he/she can choose one of the nine base packages. Over the time, users can subscribe more channels to their active package or can switch to another base package. Switching packages happen very rarely. Users either keep engage with their current package or user just disconnect the service without knowing other options that they can enjoy based on their needs.

The objective of this paper is to develop a base package recommendation system based on consumer's past engagement with the service. This method will be financially beneficial for the service provider as they can target specific customers for different promotions/discounts based on user scores and also improve customer retention. Also, it helps customers to enjoy the full benefits based on their needs. An example is given in Figure 1.

II. RELATED WORK

A. Score Meter

Credit scoring models are popular among the existing scoring methods. A credit score reflects the likelihood that a consumer will repay his debts [3]. In other word it emphasize the churn or default behavior of a customer. Different scoring models are used based on the type of business and the domain. Some of the common credit scoring factors are listed below.

- **Payment history:** The most highly weighted factor of credit score.
- **Credit utilization ratio:** It compares the total amount of credit customer currently using with the available total amount credit of the customer.
- **Total debts:** This is the sum of all customer's debts.
- **Credit mix:** It looks at the different types of credit accounts customer is using such as a mortgage, an auto loan, a credit card, store credit etc...
- **Account age:** It indicates how old customer's credit accounts and the importance of aging.
- **Hard inquiries:** When someone runs a credit check on the customer, it known as a hard inquiry.
- **Public records:** tax liens, bankruptcies, or civil judgments are included.

Fair Isaac Corporation (FICO) and Vantage score are two popular credit scoring methods [10]. Weights of FICO score was calculated from payment history (35%), outstanding debts (30%), length of your credit history (15%), types of credit used (10%) and the amount of new credit (10%). Weights of the Vantage score are determined by recent credit amount (30%), payment history (28%), credit utilization (23%), account balances size (9%), depth of the consumer's credit (9%), amount of available credit (1%).

Due to generalization capability and associated memory characteristic of artificial neural networks (ANN), artificial intelligence (AI) has become a very popular alternative in credit scoring modelling [15]. Bellotti et al. describe a method based on Support Vector Machines (SVM).

The RFM indicator was one of the most popular tools for valuing customers based on their previous purchases. Components of RFM analysis are given below [5].

- **R - Recency:** How much time has passed since the last purchase.
- **F – Frequency:** How many times the customer made purchases
- **M – Monetary:** How much money did the customer leave

Customers are divided into n groups according to their value for these three components.

B. Recommender Systems

Recommendation system paradigms such as collaborative filtering and content-based filtering are widely proposed and employed in most of the recommendation engines [1], [8], [13]. Collaborative filtering was solely based on the past interactions between users and items [21]. These interactions were stored in a user-item sparse matrix [14] and then used to identify similar users and items for recommendations.

Collaborative filtering was further divided into two sub categories: (a) memory-based collaborative filtering and (b) item-based collaborative filtering [18]. The memory-based approach depends on the user-item sparse matrix, whereas the model-based approach create a latent model to understand the reasons behind each interaction in the user-item sparse matrix [22]. Memory-based approach was divided into two parts namely

[4]Note: Even though higher price packages contain more channels, they always don't include every channel available in lower packages.

(a) user-based approach and (b) item-based approach [22]. The user-based approach tried to predict a user's interest in an item by using ratings of similar users [6], [12]. Item-based approach used the same idea, but used similar items instead of users [9], [16]. In order to measure similarity between users or items, cosine similarity and Pearson correlation were used [16], [21].

Having less number of hyper parameters was considered as the main advantage of the memory-based approach over the model-based collaborative filtering. However, it failed to handle data-sparsity as effectively as the model-based collaborative filtering [22]. K-nearest neighbor method was a good example of memory-based collaborative filtering [20]. Matrix factorization techniques is considered as an example of a model-based approach [7], [14].

The main advantage in collaborative filtering was requirement of no any prior information or features about users and items as the model can able to mine that information from the user-item interaction matrix [21], [23]. With more user-item interactions recorded, the model is able to produce better recommendations. Here, the recommendation process was transformed to a binary classification problem [4] to find whether a certain user likes a product or not, or to a regression problem like predicting a user's rating to a selected product [19]. The main advantage of the content-based approach was the robustness to the cold-start problem. If the new users/items have the set of features required by the model, making a good recommendation is possible. But the main disadvantage of this method is the dependency on higher number of features [17].

III. OVERALL FRAMEWORK

Fig 2 elaborates the overall work flow and architecture of the project. Results of score meter and package recommendations are combined to make better decisions. Identifying the proper segments of customers, targeting, positioning are the key Segment Target Position(STP) concept of marketing [2]. Customer score meter is employed to cluster the customer and track their behaviours. Decision making process is conducted based on the customer scores.

There are important factors that affect the customer behaviours in IPTV domain. In this paper, we propose the following factors to calculate the customer's basic score.

- **Base package:** Indicates whether it is a high-value package or not.
- **Subscription packages:** How actively a person is subscribing and use packages. If a person is more actively subscribing and uses packages it means he is more engaged with the IPTV platform. Therefore he deserves a high value.
- **Current viewing behavior:** How actively he/she is watching TV. This also indicates how actively the customer engages with the service. When the customer is actively subscribing and use new packages his/her viewing behavior is less important. But when he/she is not actively doing things his/her viewing behavior is more important for IPTV service provider. Because if a customer is not actively watching the channels the probability, that the user is going to deactivate the service is very high.
- **Past months' behavior**: How valuable he/she was within the past months.

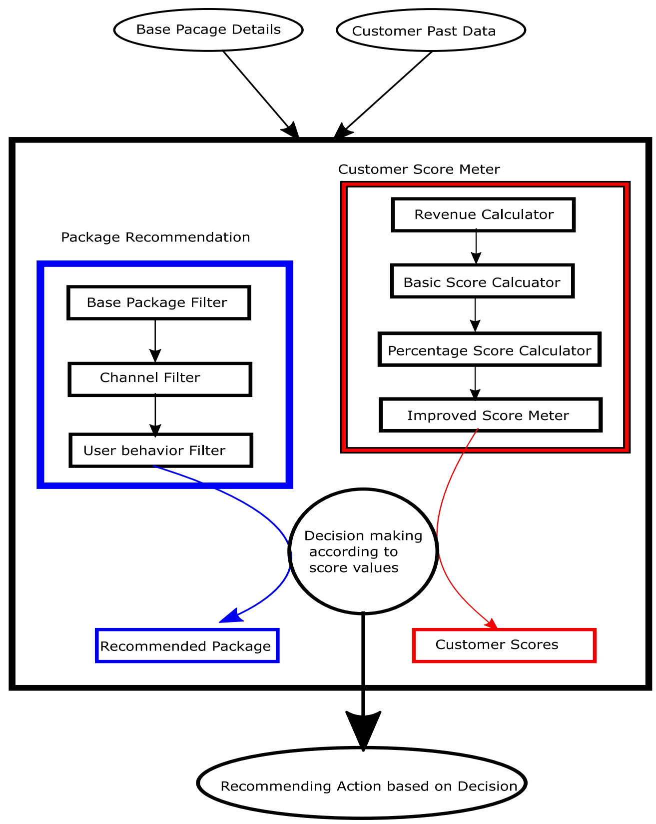

Fig. 2. Overall flow diagram with package recommendation and score-meter

A. Score Meter

1) Basic score: The basic score algorithm provides an individual score that indicates customers' individual behavior. The output is the current basic score. The algorithm that uses to calculate the basic score is shown given in Algorithm 1.

Here, m_rev_t denotes the revenue generated by the customer at month t. Similarly, parameters:
$score_t = the\ customer's\ score\ at\ month\ t$

Algorithm 1: Basic score algorithm

if $m\_rev_t != m\_rev_{t-1}$ **then**
  $score_t \leftarrow \frac{k*m\_rev_t+(k-1)*score_{t-1}..+1*score_{k-(t-1)}}{k*(k+1)/2}$;
if $\beta_t * score_{t-1} \geq \alpha * \sigma * m\_rev_t$ **then**
  $score_t \leftarrow \beta_t * score_{t-1}$;
else
  $score_t \leftarrow \alpha * \sigma * m\_rev_t$

$\beta_t = the\ score\ reduction\ factor\ at\ month\ t$
$\alpha = \frac{the\ upper\ bound\ of\ the\ minimum\ score}{m\_rev_t}$
$\gamma = \frac{the\ lower\ bound\ of\ the\ minimum\ score}{m\_rev_t}$
$\sigma = max(\frac{number\ of\ activedays\ within\ the\ month}{number\ of\ days\ of\ the\ month}, \frac{\sigma}{\alpha})$
$k = weighted\ average\ window\ size$

The parameter β_t depends on β_1 and β_2. Equations for calculating β_t and β_2 are represented as follows

$$\beta_t = \frac{2\beta_1\beta_2}{\beta_1 + \beta_2} \tag{1}$$

$$\beta_2 = max(\frac{activemonths}{n}, 0.5) \tag{2}$$

$$\beta_1 = \frac{\alpha^{\frac{1}{n}}}{2 - \alpha^{\frac{1}{n}}} \tag{3}$$

Here, the user define parameter n represents the number of months, that need to reach to the minimum score.

It is challenging to compare two customers using the basic score as there is no limitation to the score. Customers in higher base packages tend to get higher scores compared to customers in lower packages, even though customer is highly active. Therefore, we present a percentage score, which can be used to compare among customers.

2) Percentage Score: We apply log transformation for all the individual customer scores and transformed values are normalised between zero and 1.

$$score_n = \frac{score_t - min\_score_t}{max\_score_t - min\_score_t} \tag{4}$$

Normalized scores are multiplied by 100 to obtain percentage scores.

$$score_p = score_n.100\% \tag{5}$$

Parameters α, γ, n and k are user define input arguments to the score meter.

B. Package Recommendation model

1) Vector Representation - Base Package: In our model, every base package is represented as a 13x1 vector. The base package vector v_{i} calculates as follows.

$$v_{\mathrm{i}} = \begin{bmatrix} c_{i,1} \\ c_{i,2} \\ \vdots \\ c_{i,j} \\ \vdots \\ c_{i,12} \\ p_i \end{bmatrix}_{13\times1} \tag{6}$$

Where,

$$c_{\mathrm{i,j}} = \frac{\textit{channels belonging to } j^{th} \textit{ category}}{\textit{total no. of channels in } i^{th} \textit{ package}} \tag{7}$$

And,

$$p_{\mathrm{i}} = \frac{\textit{price of } i^{th} \textit{ base package}}{\textit{sum of prices of all base packages}} \tag{8}$$

We consider the default genres defined by the service provider as the basis for vector v_{i}. Note, that only 12 categories are considered to get the base package vectors (See Table I). We combined the original music category with the local music category into one category as local music category only contains a single channel.

2) Vector - Subscriber: Every subscriber is represented using two 12x1 vectors. They are,

- User-time vector
- User-channel-count vector

User-time vector, $u_{\mathrm{k,t}}$ for the $\mathrm{k^{th}}$ subscriber obtains as follows.

$$u_{\mathrm{k,t}} = \begin{bmatrix} w_{k,1} \\ w_{k,2} \\ \vdots \\ w_{k,j} \\ \vdots \\ w_{k,12} \end{bmatrix}_{12\times1} \tag{9}$$

where,

$$w_{\mathrm{k,j}} = \frac{\textit{TV watching time for } j^{th} \textit{ category}}{\textit{Total TV watching time}} \tag{10}$$

The user-channel-count vector $u_{\mathrm{k,n}}$ for subscriber k obtains as follows[5]

$$u_{\mathrm{k,n}} = \begin{bmatrix} x_{k,1} \\ x_{k,2} \\ \vdots \\ x_{k,j} \\ \vdots \\ x_{k,12} \end{bmatrix}_{12\times1} \tag{11}$$

where,

$$x_{\mathrm{k,j}} = \frac{\textit{No. of channels watched in } j^{th} \textit{ category}}{\textit{Total no. channels watched}} \tag{12}$$

3) The Model: The base package recommender consists of 3 stages.

1) Base Package Filter
2) Channel Filter
3) User-behaviour Filter

In base package filter stage, we choose the 5 most similar base packages for the subscriber's current base package. We calculate the cosine similarity between all base packages and the user's current package.

[5]Note that to create these user vectors we considered 3-week data

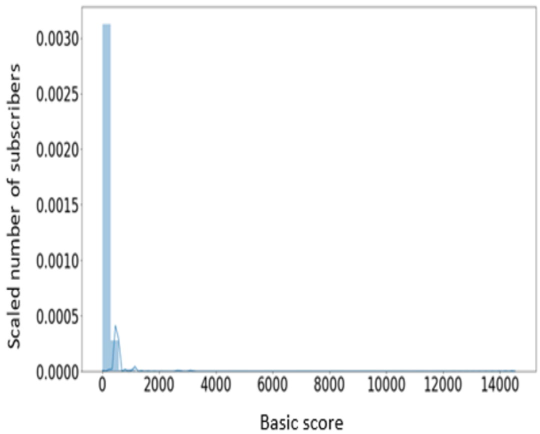

Fig. 3. Histogram of basic scores

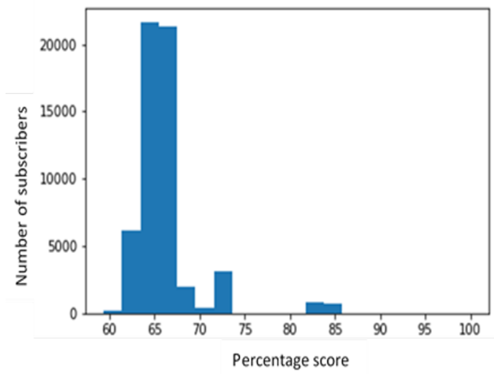

Fig. 4. Histogram of percentage scores

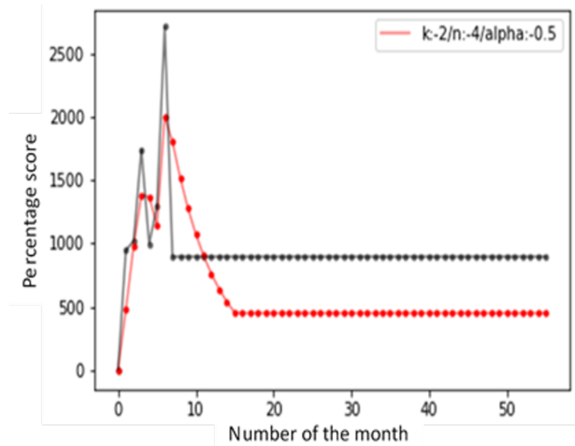

Fig. 5. Score variation for Subscriber A

In channel filter stage, we input the selected 5 base packages. Then, we use subscriber's most-watched 3 genres (popular genres). For each genre, we choose 3 most-watched channels (popular channels). Then we score each of the selected 5 base packages. This score is obtained for the package i as follows.

```
For genre in popular_genres:
    For channel in popular_channels:
        score += 4 - rank(channel);
```

The most-watched channel in a genre has a rank of one and the second most-watched channel has a rank of two. The third most-watched channel's rank is three. After obtaining scores for all five base packages, we select the top 3 packages with the highest scores.

For user-behaviour filter stage, we use the outputs of the channel filter stage. For every input base package (selected packages), we obtain another score as follows.

```
For package_i in selected packages:
    score_final_1_i=cosine_similarity(package_i,u_k,t)
    score_final_2_i=cosine_similarity(package_i,u_k,n)
    score_final_i=score_final_1_i + score_final_2_i
```

Note that the dimension of base package vectors is 13x1 and user vector is 12x1. For cosine similarity the last element in the corresponding base package vector is not considered. Next, we output the base package i with the highest $score\_final_i$ as the final recommendation to the selected subscriber.

IV. RESULTS AND DISCUSSION

We collected over 4 years of data and generated score values for each subscriber in every month using details such as current base package, subscription packages, current viewing behavior and previous month's behaviour.

The histogram of basic scores is shown in Figure 3. As shown in Figure 3, identify different user clusters such as valued customer, churn-able customers are challenging task. As discuss in Sec III, data transformation is applied to reduce the Skewness. As shown in Figure 4, 2 different separable histogram bins of customers are identified. Majority of the customers' percentage scores lie between 62 and 74. Subscribers who have scores more than 80 can be identify as revenue generating customers in this IPTV platform.

Two customer examples are shown in Fig 5 (*Subscriber A*) and Fig 6 (*Subscriber B*). Variation of revenue generation and generated monthly score is shown in black and red colour respectively. We set adjustable parameters K=2, α=0.5 and n=4. Thus, scores have reached to its minimum value after four months and minimum scores are equal to half of their revenues.

As shown in Fig 5, *Subscriber A* engaged positively during the first 10 months due to package upgrade and additional channel purchase. After 10 months, he showed less engagement with his current base package. During the higher engagement period he maintained higher scores. Less activities caused to decrease the customer score between 10^{th}–20^{th} month. After 20^{th} month, minimum score is given from our algorithm due to no actions and it remains for next several months. It indicates customer's churn-ability.

Fig 6 depicts the flow of the scores for *Subscriber B*. After having a positive engagement with the IPTV platform

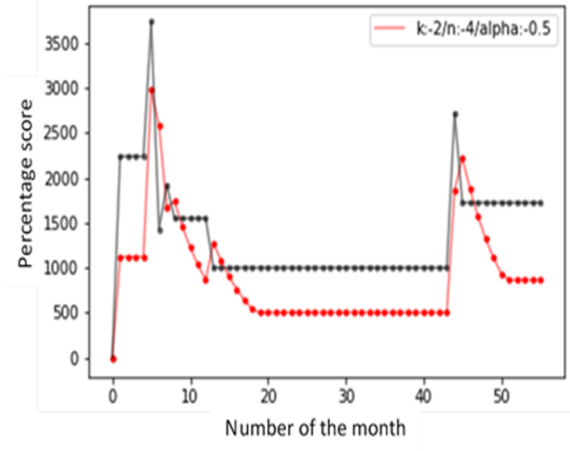

Fig. 6. Score variation for Subscriber B

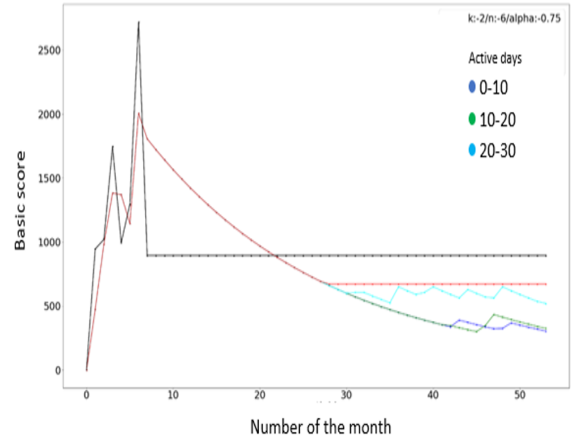

Fig. 8. Example plot for the improved algorithm

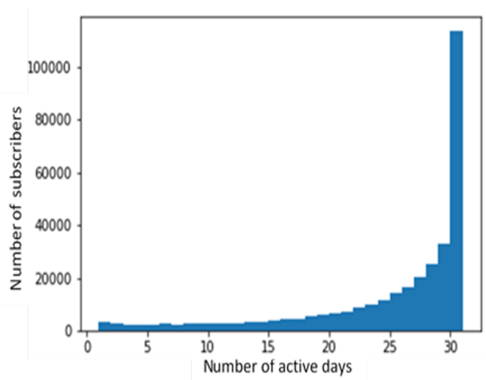

Fig. 7. Histogram of active days within a month

for several months, his scores started to decrease due to less activities and remained in score 500 till 40th month. Due to package update and new subscription, his scores started to increase after 40th month.

Fig 7 shows the histogram of customers' active days within a month. Based on the histogram, an improved version of the algorithm is proposed to further accurate qualitative predictions. Results of scores generated by the improved algorithm for *Subscriber A* is shown in Fig 8. Note, that we only use watching time behaviour, when a customer reached to the minimum score level. Here, the scores were changed according to the number of watching days within the month. It demonstrates that the customer had more engagement with the service by watching more time even though he was not very active with other services. Fig. 8 illustrates all three different cases using watching time (For illustration purpose, we syntactically generated the watching time for each active days category in the plot). Based on this score, we can identify the probability to churn his account in the near future.

Recommendations obtained for some selected subscribers are discussed in a qualitative manner. We used three weeks of data buckets and provides a recommendation. We used next three week of data for validation our recommendations. There is no 100% accurate method to validate this recommender system [11]. The most reasonable method is to select a subset of users and let them to interact with the recommendation system and analyze their behaviour and feedback. Considering this limitations, we used current 3-week data and obtain a recommendation from our model. Then we generated subscribers actual behaviour using next three weeks data and compare it against with the recommended package from our method. Few examples are shown in Figure 9,10, 11, 12, 13 and 14,. As shown in Figure 9, our recommendation model recommend *Subscriber ID-34767*'s to upgrade to other package for next three week. His actual behaviour for next three weeks was same as the recommended package (See Figure 10). Some examples for downgrade and neutral recommendation is given in Figure 13, 14, 11 and 12 for *Subscriber ID 507497* and *Subscriber ID 510533*.

Experiments are conducted using a high end machine with one RTX2080Ti grpahics card. The current model consumed approximately 15 minutes to output a base package recommendation for one subscriber. The majority of the runtime was spent on loading the entire data set to the memory and for pre-processing.

V. CONCLUSION

In this paper, we describe a base package recommendation system based on novel customer score meter, which use consumer watching pattern, package interaction and package price. Different set of valued and premium customers are identified from the data mining approach. Based on customer scores, recommendation are pushed to customers to keep them engaged continuously with the service. We used a content-based approach to represent each subscriber and available base package as set of vectors. Vector similarity measurements and other predefined set of rules are used to come up with base package recommendations. Our propose recommendation platform based on customer score meter is proven in a qualitative approach to be the most efficient way of recommending base packages in IPTV platform.

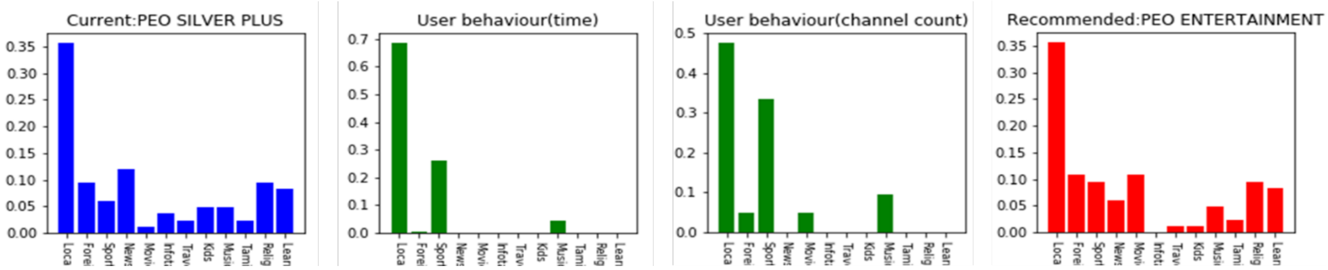

Fig. 9. First 3 week analysis for userID 34767

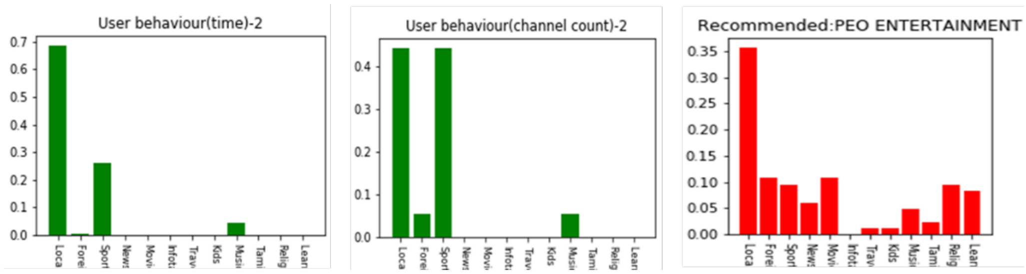

Fig. 10. Actual behaviour analysis for userID 34767

REFERENCES

[1] A Merve Acilar and Ahmet Arslan, 'A collaborative filtering method based on artificial immune network', *Expert Systems with Applications*, **36**(4), 8324–8332, (2009).

[2] Santi Mutiara Asih and Sinta Kemala Asih, 'Marketing strategy implementation in developing sharia tourism in indonesia', *International Proceedings of Economics Development and Research*, **84**, 133, (2015).

[3] Robert B Avery, Kenneth P Brevoort, and Glenn B Canner, 'Credit scoring and its effects on the availability and affordability of credit', *Journal of Consumer Affairs*, **43**(3), 516–537, (2009).

[4] Chumki Basu, Haym Hirsh, William Cohen, et al., 'Recommendation as classification: Using social and content-based information in recommendation', in *Aaai/iaai*, pp. 714–720, (1998).

[5] Derya Birant, 'Data mining using rfm analysis', in *Knowledge-oriented applications in data mining*, IntechOpen, (2011).

[6] John S Breese, David Heckerman, and Carl Kadie, 'Empirical analysis of predictive algorithms for collaborative filtering', in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, pp. 43–52. Morgan Kaufmann Publishers Inc., (1998).

[7] John Canny, 'Collaborative filtering with privacy via factor analysis', in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 238–245. ACM, (2002).

[8] Long-Sheng Chen, Fei-Hao Hsu, Mu-Chen Chen, and Yuan-Chia Hsu, 'Developing recommender systems with the consideration of product profitability for sellers', *Information Sciences*, **178**(4), 1032–1048, (2008).

[9] Mukund Deshpande and George Karypis, 'Item-based top-n recommendation algorithms', *ACM Transactions on Information Systems (TOIS)*, **22**(1), 143–177, (2004).

[10] Laurie Goodman, 'In need of an update: Credit scoring in the mortgage market', *Retrieved from the Urban Institute website: http://www. urban. org/sites/default/files/publication/92301/in-need-of-an-update-credit-scoring-in-the-mortgage-market_1. pdf*, (2017).

[11] Asela Gunawardana and Guy Shani, 'A survey of accuracy evaluation metrics of recommendation tasks', *Journal of Machine Learning Research*, **10**(Dec), 2935–2962, (2009).

[12] Jonathan L Herlocker, Joseph A Konstan, Al Borchers, and John Riedl, 'An algorithmic framework for performing collaborative filtering', in *ACM SIGIR Forum*, volume 51, pp. 227–234. ACM New York, NY, USA, (2017).

[13] Mehrdad Jalali, Norwati Mustapha, Md Nasir Sulaiman, and Ali Mamat, 'Webpum: A web-based recommendation system to predict user future movements', *Expert Systems with Applications*, **37**(9), 6201–6212, (2010).

[14] Yehuda Koren, Robert Bell, and Chris Volinsky, 'Matrix factorization techniques for recommender systems', *Computer*, (8), 30–37, (2009).

[15] Tian-Shyug Lee, Chih-Chou Chiu, Chi-Jie Lu, and I-Fei Chen, 'Credit scoring using the hybrid neural discriminant technique', *Expert Systems with applications*, **23**(3), 245–254, (2002).

[16] Greg Linden, Brent Smith, and Jeremy York, 'Amazon. com recommendations: Item-to-item collaborative filtering', *IEEE Internet computing*, (1), 76–80, (2003).

[17] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro, 'Content-based recommender systems: State of the art and trends', in *Recommender systems handbook*, 73–105, Springer, (2011).

[18] Hao Ma, Irwin King, and Michael R Lyu, 'Effective missing data prediction for collaborative filtering', in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 39–46. ACM, (2007).

[19] Seung-Taek Park and Wei Chu, 'Pairwise preference regression for cold-start recommendation', in *Proceedings of the third ACM conference on Recommender systems*, pp. 21–28. ACM, (2009).

[20] Arkadiusz Paterek, 'Improving regularized singular value decomposition for collaborative filtering', in *Proceedings of KDD cup and workshop*, volume 2007, pp. 5–8, (2007).

[21] Badrul Munir Sarwar, George Karypis, Joseph A Konstan, John Riedl, et al., 'Item-based collaborative filtering recommendation algorithms.', *Www*, **1**, 285–295, (2001).

[22] Jun Wang, Arjen P De Vries, and Marcel JT Reinders, 'Unifying user-based and item-based collaborative filtering approaches by similarity fusion', in *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 501–508. ACM, (2006).

[23] Gui-Rong Xue, Chenxi Lin, Qiang Yang, WenSi Xi, Hua-Jun Zeng, Yong Yu, and Zheng Chen, 'Scalable collaborative filtering using cluster-based smoothing', in *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 114–121. ACM, (2005).

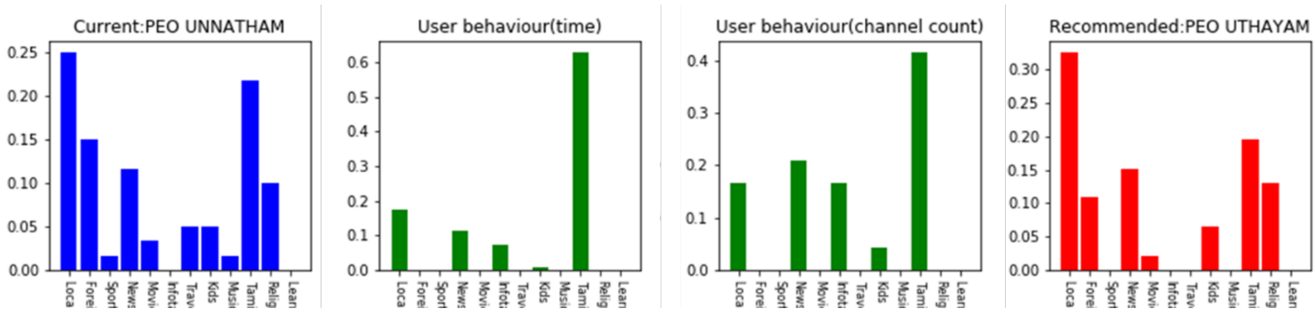

Fig. 11. First 3 week analysis for userID 507497

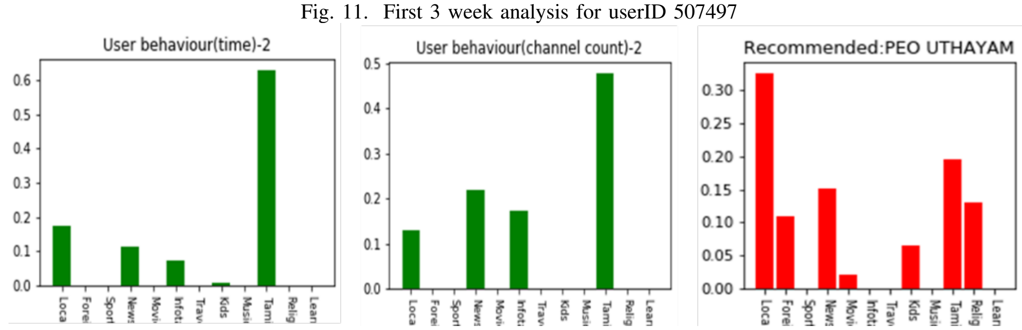

Fig. 12. Actual behaviour analysis for userID 507497

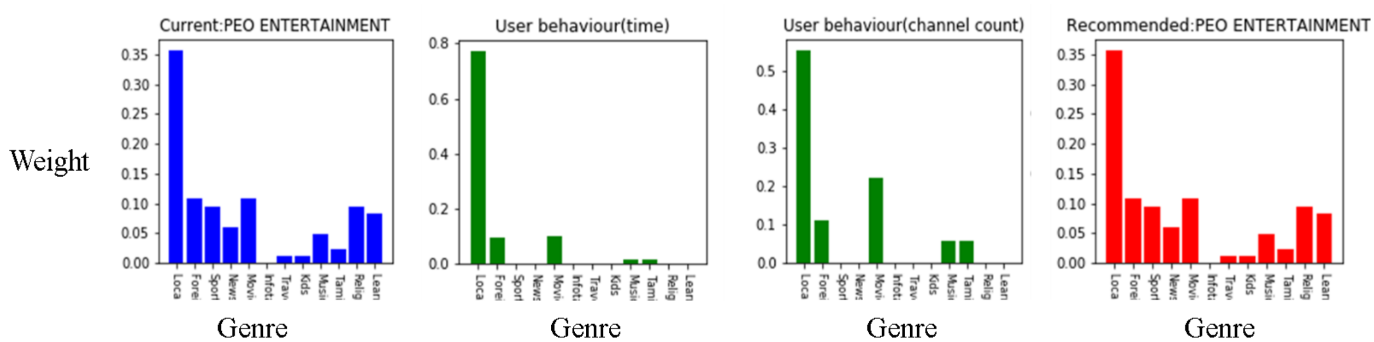

Fig. 13. First 3 week analysis for userID 510553

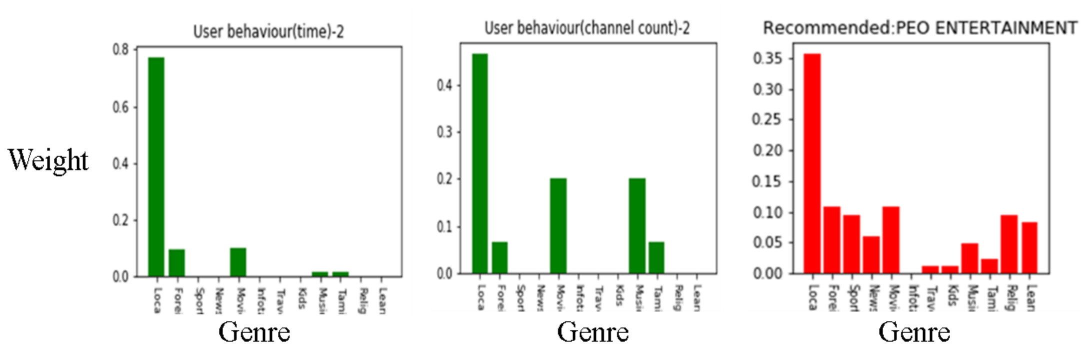

Fig. 14. Actual behaviour analysis for userID 510553