Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7, July 2021: 2544 – 2551

Channel Attribution Using Markov Chain model

N. Sandeep[1], S. Vasavi[2], G. Shruthi[3]

[1,2]Department of Computer Science & Engineering, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India,
[3]Anblicks Cloud Data Engineering Ltd
[1]sandeepnidumukkala@gmail.com, [2]vasavi.movva@gmail.com, 3shruthi.gangam@anblicks.com

ABSTRACT

For every company it is crucial to know how the marketing channels performing towards the sales of a company. So, there is a need for a model to predict which channel is best performing and which is least performing. The multi-channel attribution helps to find which channels generate the more conversions and most value to the sales, they are very important to predict overall marketing ROI and helps to make key decisions based on analysis of the data. Multi-channel attribution is must for every business marketer to identify the growth. But most of them not able to understand how it works and cannot use it as it are difficult for them. The recent study shows that 59% of marketers didn’t have knowledge of channel attribution and admitted that a lack of understanding is the key reason why they haven't implemented an attribution model. Multi-channel attribution helps to track and find the marketing touch points which make the user to convert which leads to conversion/sales. Channel Attribution follows predefined rules that assign the attribution value to every touch point which contributes to the sales of company. There are many benefits like that it tracks each touch point across the entire conversion path. But there are different models which give us different results and also, they give credit to only one channel which leads to overconfidence. Markov Model is a probabilistic model which will assign value based on channel contribution and calculate for all channels so that marketer will know better influencer.

Keywords: ROI, Channel attribution, touch point, Conversion path, Markov model

1 Introduction

Generally, businesses that are running actively do marketing campaigns which helps them to identify the channels that make user to drive into actual conversion. It is very essential for the marketers to find the ROI which is a key performing index (KPI). Gradually the no of channels for marketing are growing day by day to increase more customers and to engage them to more products which lead to good sales. It is more crucial to know the impact of marketing channels on company’s sales. A recent study identified that 92% of users browsing a retailer's website for the first time didn’t buy the product and they searched different sources before made a purchase. There are different standard attribution models like first click, last click, linear model, position model and time decay model etc., which are used for predicting the channel efforts to conversion. Different companies use different methods according to their own need. But these models may not give same result for same data.

Last-click Attribution model

This attribution model is the default for Google Analytics. It delivers all the credit to the last track source that resulted in a sales conversion. It's a simple, rough-cut way to see what actions result directly in sales.

First-click Attribution model

The model is analogous to the last click model, and assigns all credit to the first channel a customer interacted with before ultimately converting. In the marketing attribution example above, all credit would go to the email. Both of the models run the risk of oversimplifying and inaccurately depicting true motivation.

Linear Attribution model

In this approach, the attribution value is equally assigned to the channels present in the user journey which leads to conversion. This model is better suitable to find the trend of multi touch channel behaviour. But it cannot distinguish between the different channels and this drawback of this model.

Markov Chain Attribution [6,7,8]

Markov Chain model describes the next state by using the current state only. It uses the previous/historical data and based on that data it evaluates the next state and predict the next state. It considers only the current state and predicts the next states. Markov model give credit to the channel/ touch point which contributed more towards sale.

Figure 1presents the channel attribution problem in which it has `Start' means the user just started the journey and visits the channels like `Facebook' , `Google' etc and after that the user may convert or not depends upon his need and utility if the product. If the user converted the `Conversion' is used to represent and if the user didn't convert then `Null' is used to represent it.

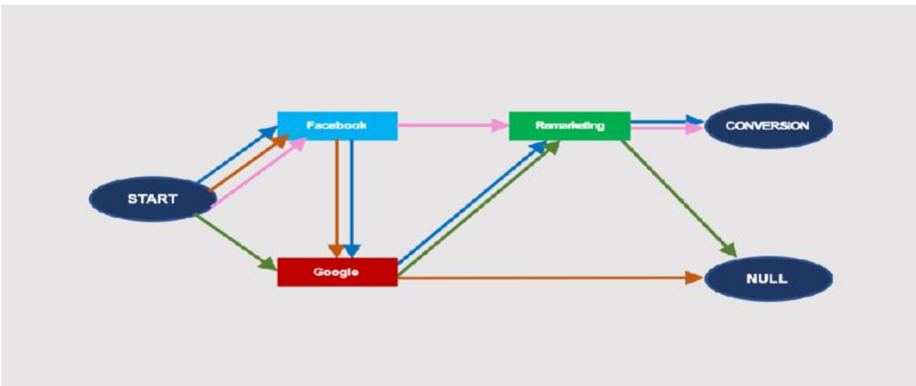

Figure 1. Markov Chain Model [5]

2 Background

There are so many standard attribution channels which are very easy and can implement them to find the Return on Invest of our company advertising channels. This leads to the overconfidence to the marketers in which the results are biased to only one channel and will not give credit to the remaining channels in the user journey .This leads to wrong estimation of future business. In order to overcome this oversight, the authors used a more advanced and good optimized model called Markov model which helps to eradicate the problems of all heuristic models.

In the work [1], it analyses the selected company's data by following Markov chain method. It has the customer journeys data and is analysed. In this the authors understand the Markov model which decreases the credit for the last touch models but it is more favoured by first touch and linear models. The Markov order estimator GDL is the best suitable model for buyer journey analysis and has a great use to find the conversion path. In this 40% revenue is achieved by the journeys with less than five interactions.

In the work [2], it helps us to jump from traditional Google heuristic models to probabilistic models like Markov chain. In this have to calculate the transitional probabilities which assign value or give credit to every channel/ touch point .Generally when buying a product customer check with different sources and may contact different channels through his journey before he buys anything. Google analytical models like last click , first click models attribute the channel values based on some condition like whole credit goes to first channel in case of first click model.

In the work[3], It is an analytical model which describes the problem of assigning credit value to the customer buying products from different channels of a digital market campaign. The attribution value is to give incentives to the advertiser which encourages the advertiser to more advertise the product and leads to increase in sales for a company. It has a two-stage marketing funnel, where the publishers should advertise and helps to conversion. The result helps us to understand the multi touch attribution thus giving higher credit to the portion of the funnel where the baseline conversion rate is higher. Next, helps us to understand the social welfare maximizing contracts may have a higher conversion rate than optimal multi-touch contracts.

In this paper [4], it follows deep learning methodology and assess the effect on every channel how much they impact is. In this its present Deep Neural Net with Attention multi-touch attribution model (DNAMTA) model in a fashion of supervised learning finding if a series of events leads to conversion, and helps us to understand the interaction of different channels .It uses study of population such as age ,sex, race and behaviour to decrease the deviations of channel effects and also includes user related context data to understand the impact of every media channel on sales. In this method they used a large real time marketing dataset to describe that the methodology used is very good performing and best method compared to already existing methods.

3. Methodology

For the implementation of this project, we are proposing an approach Markov Chain model which is nothing but a probabilistic model predict the next state value based on the current state only. The proposed system predicts the probability of conversion of channel to find the best and least influencer. Figure 3 presents the proposed system of channel attribution model which tells us the brief description of the project and the sequence of steps involved in it.

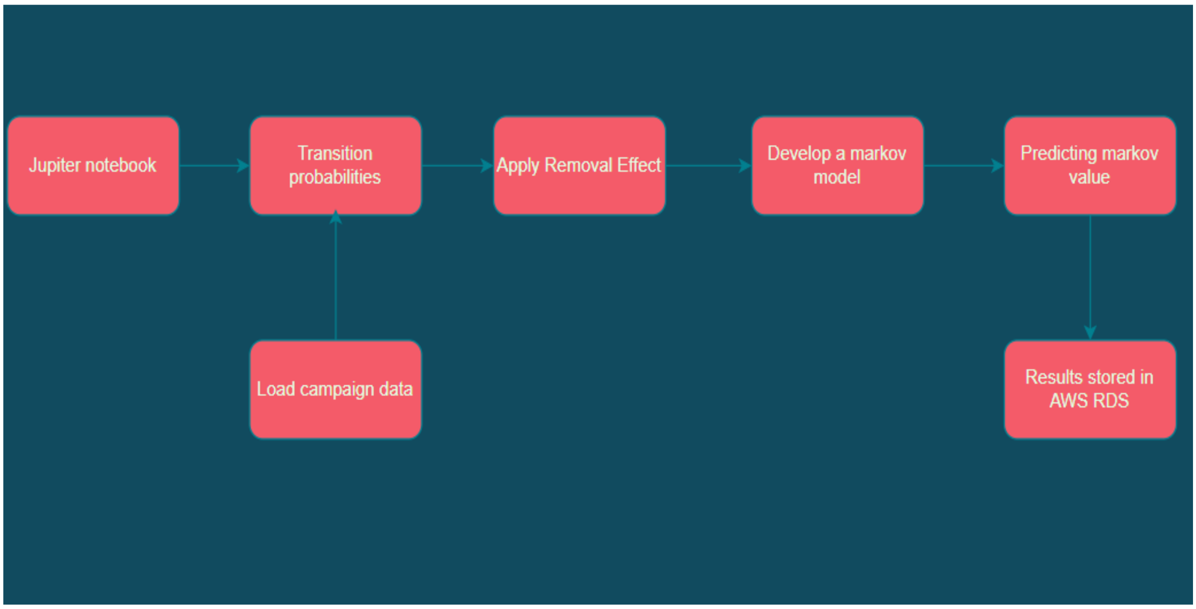

Figure 3: Proposed System Architecture

For the implementation of this project, we are proposing an approach Markov Chain model in which have to calculate the probabilities of every transition state and applying the removal effect and finding the Markov value by using the Markov principle.

The approach contains the following modules:

1. Data Pre-processing - Data cleaning and Feature Engineering.
2. Transition state probabilities
3. Removal Effect
4. Apply Markov model
5. Results stored in AWS RDS

3.1. Data Pre-processing - Data cleaning and Feature Engineering.

• Data cleaning - In this module we will identify all the null or N/A values associated with each attribute and will remove their corresponding rows as we have sufficient number of values for training and testing the model.

• Dropping the unnecessary features - In this process we will remove the attributes like `interaction', which are not useful for project.

• Removing Duplicates - In this process the rows containing same values have to be removed.

3.2. Transition matrix

• Based on the path we found add the `start' at the beginning.
• Add `Null' or `Conversion' to the end if it converted add `Con-version'.
• Find the different transition states present in the data
• Find the number of times each transition occurs
• After calculate the probabilities for every transition state
• Represent them in a matrix
• Matrix should have every value and put `0' to diagonal elements at first
• After fill the values accordingly and display it using heatmap.
• Heat map helps us to find the correlation between the different channels.
• So that we can predict users are migrating from which channel to Channel more.
• Later these transition probabilities are used in removal effect Calculation
The figure 4 tells us the probabilities of different transition states.

	Online Display	Instagram	Online Video	Facebook	Start	Paid Search	Conversion	Null
Online Display	0.000000	0.029363	0.017299	0.053986	0.0	0.092386	0.050324	0.756643
Instagram	0.023531	0.000000	0.024118	0.218777	0.0	0.045809	0.057979	0.629786
Online Video	0.018920	0.031908	0.000000	0.059323	0.0	0.047947	0.078146	0.763757
Facebook	0.024352	0.173072	0.025577	0.000000	0.0	0.050655	0.053220	0.673123
Start	0.142644	0.119188	0.142361	0.278408	0.0	0.317399	0.000000	0.000000
Paid Search	0.048053	0.034196	0.029221	0.064442	0.0	0.000000	0.053309	0.770779
Conversion	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	1.000000	0.000000
Null	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	1.000000

Figure 4: Transition matrix

3.3. Removal Effect

• Remove the channel
• calculate the total conversions present in the remaining data
• Find the probability of every channel.
• Probability is calculated by removing the channel and counting the total Conversions present. Figure 5 presents the removal effect value for every channel present in our dataset.

```
{'Online Display': 0.15435482356041264,
 'Instagram': 0.21731366149038456,
 'Online Video': 0.20691411655642178,
 'Facebook': 0.3547597674182721,
 'Paid Search': 0.3311037560086154}
```

Figure 5: Removal Effect Proportions

3.4 Apply Markov chain model

Markov Chain model describes the next state by using the current state only. It uses the previous/historical data and based on that data it evaluates the next state and predict the next state. It considers only the current state and predicts the next states.

1. To apply this, we have to calculate the transition from one channel to another channel and finding the probability of each channel. And also, we have to calculate the removal effect of every channel.
2. Markov chain calculates values based on the past information.
3. It is a probabilistic model.
4. It assigns the channel attribution value based on the contribution of each channel in the path.

Markov model = removal effect / Total removal effect * Total Conversion rate

3.5 AWS RDS

1. Open AWS account and search for RDS.
2. Create a database and specify the requirements.
3. Now launch the database.
4. Open MySQL and connect with RDS.
5. import the result and save it.

3.6 Dataset collection

The dataset we are going to use for this project is Campaign Attribution dataset. It contains the information about the cost of various customers buying orders

• There are 587626 entries with 6 parameters from which we are going to select the required parameters. Figure 6 presents the dataset in which it has 6 attributes. The attributes are:

1.Cookie Id
2.Time
3.Interaction
4.Conversion
5.Conversion value
6.Channel

	cookie	time	interaction	conversion	conversion_value	channel
0	00000FkCnDfDDf0iC97iC703B	2018-07-03T13:02:11Z	impression	0	0.0	Instagram
1	00000FkCnDfDDf0iC97iC703B	2018-07-17T19:15:07Z	impression	0	0.0	Online Display
2	00000FkCnDfDDf0iC97iC703B	2018-07-24T15:51:46Z	impression	0	0.0	Online Display
3	00000FkCnDfDDf0iC97iC703B	2018-07-29T07:44:51Z	impression	0	0.0	Online Display
4	0000nACkD9nFkBBDECD3ki00E	2018-07-03T09:44:57Z	impression	0	0.0	Paid Search

Figure 6 Dataset

4 Results and Discussion

The results of the proposed describes about every channel contribution to the conversions. Figure7 presents the Markov value of each channel.

	Online Display	Instagram	Online Video	Facebook	Paid Search
0	2153.246927	3031.521549	2886.44809	4948.892178	4618.891257

Figure 7 Markov Values

Based on the above result we can decide that the major contributor and least contributor and will allocate the budget more to major contributor so the sales of the company will increase and it also saves time.

From the result we can conclude

The first best influencer is `Facebook'.

The second-best influencer is `Paid Search'.

The first worst influencer is `Online Display'.

The second worst influencer is `Online Video'.

From the manager will allocate more budget to the Facebook and paid search as they are most influencing their products. And manager may remove online display and online video from the marketing campaign so that he can save some money and use it on purposes which will impact his business more. Here the Figure 8 describes the proportion of each channel on the whole contribution. Here we used pie chart to visualize the channel contribution.

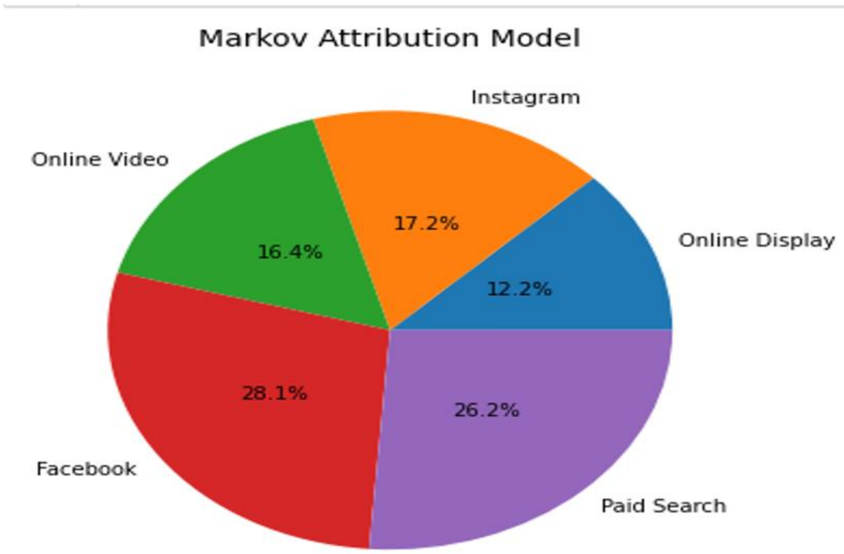

Figure 8: Pie Chart of Different channels

5. Conclusion and Future Work

In this project a methodology for channel attribution is proposed. Firstly, we loaded the Channel attribution dataset into Jupiter notebook. Transition probabilities are calculated for every transition state and calculated the removal effect for every channel and apply the Markov principle to find the contribution values. These values are used to find the best influencer and least influencer. Based on these marketers allocate their budget properly. This model assigns value based on the channel performance, how frequently users are visiting and number of conversions done. Whereas others models assign values to one channel and it leads to overconfidence. This is the best approach and marketers are adopting it very fatly due to accurate results. Future Work is that this data contains only 5 channels but where as in real there are many channels to explore. So it can be done with a greater number of channels.

References

1. Lukáš Kakalejč & Jozef Bucko & Paulo A. A. Resende & Martina Ferencova, 2018. Multichannel Marketing Attribution Using Markov Chains, Journal of Applied Management and Investments, Department of Business Administration and Corporate Security, International Humanitarian University, vol. 7(1), pages 49-60, 2018
2. Sergery Bryl, Marketing Multi-Channel Attribution model with R(part 1: Markov chains concept, https://www.analyzecore.com/2016/08/03/attribution-model-r-part-1/#:~:text=This%20is%20a%20Multi%2DChannel,to%20touchpoints%20in%20conversion%20paths. Last accessed on April 10th 2021.
3. Abhishek, Vibhanshu and Despotakis, Stylianos and Ravi, R., Multi-Channel Attribution: The Blind Spot of Online Advertising (April 27, 2017). Available at SSRN: https://ssrn.com/abstract=2959778 or http://dx.doi.org/10.2139/ssrn.2959778
4. Ning Li, Sai Kumar Arava, Chen Dong, Zhenyu Yan, Abhishek Pani, Deep Neural Net with Attention for Multi-channel Multi-touch Attribution, arXiv:1809.02230, 2018
5. Divyansh Saxena, Markov Chains in R, LondonR, 2017
6. Kaelin Tessier, Marketing Channel Attribution With Markov Models In R, Bounteous, 2016
7. Serhii Puzyrov, Markov Chain Attribution Simple Explanation of Removal Effect, https://serhiipuzyrov.com/2019/07/markov-chain-attribution-simple-explanation-of-removal-effect/ Last accessed on March 13th 2021
8. Ben Denis Shaffer, Introduction to Markov Chain Multi-Touch Attribution, Towards Data science, 2020