UDC 004.023

V.L. FLAKS

FUNNEL BASED REVENUE ATTRIBUTION OF THE ONLINE CAMPAIGNS

Abstract. *The article reviews the problems of static attribution models of online stores transactions and suggests a solution in the form of dynamic attribution model based on customer behavior while customer journey on a website.*
Keywords: *Ecommerce, online store, conversion, cost of customer acquisition, customer, merchant, revenue, attribution.*

Introduction

Ecommerce field is one of the most dynamic segments of Ukrainian and worldwide economics. In many ways this was made possible due to more efficient processes including those for customers acquisition. This efficiency is based on the availability of a large amounts of personalised customer data, especially before transaction. For instance, data on a specific advertising campaign, that a customer contacted before transaction or data on the products that interested a customer who hadn't placed an order.

In classic brick-and-mortar retail there is a saying that there exist three key factors of success: "location, location and location". It is a location of a physical store that determines the quantity and the quality of potential customers, who can be attracted by the trade offer of a store.

In internet there is no possibility to place a server on crossroads or another crowded place. Thus the task of customer attraction in ecommerce is solved by online marketing. As a result the major part of a gross margin is invested in marketing of online store. Like in classic brick-and-mortar retail, a big part of gross margin is invested in stores rent and maintenance. These expenses ensure that a potential customer gets familiar with a product.

That's why management of the marketing budget is a key challenge for online retail.

Glossary for article:

Attribution — a process of defining the impact of interaction, that took place during a customer journey.

Conversion — user action that took place on a website, usually a transaction.

Retailer — businesses that sell goods directly to individuals.

Conversion value — the value from a user action that was conducted on a website or in mobile application. For instance, the revenue from an order.

User session — a chain of actions of an exact user, that are combined by a traffic source and a period of time.

Traffic source — a source that a user had interacted with before a transition took place. For instance, "google / cpc".

1. Business task description

Managing of marketing budget is based on a segments estimation with grouping by properties of advertisement channels, products and users. To put such a grouping into practice, the value estimation should be implemented based on unsampled data, granulated to each transition.

There exist a few peculiarities of the customer behavior that complicate the business task:

1. Over half of transactions are made after 2+ paid transitions. Thus we need a model of revenue attribution from each transaction for several transitions.

2. A considerable part of paid visits does not result in transactions. That's why it's impossible to clearly estimate the impact of these expenses on the revenue from the interaction chains resulted in transactions.

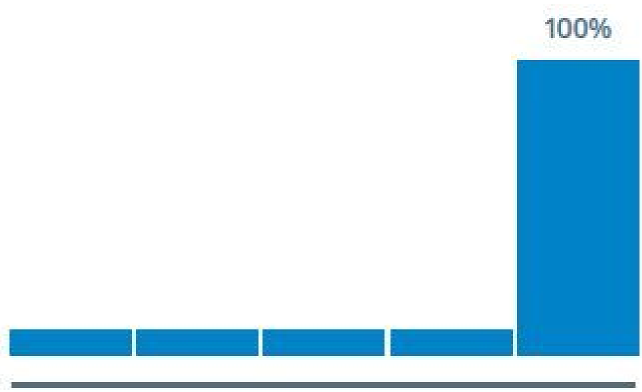
3. In many cases a customer uses several devices. It means that technically it's impossible to define the relations between user actions on different devices. Even if a certain chain resulted in a transaction. At the same time a possibility to connect interaction chains of authorised users does not solve this task for 100% visits.

4. There is a natural time lag between the moment of costs expenditure for a visit and the moment of getting revenue from a user, who conducted this visit. The time lag is caused by a period needed for making decision as for the purchase. This period can make from a couple of days to several weeks and depends on the product price. At the same time online store has to make decisions as for budget allocation as soon as possible, since the delay determines inefficiently expenses and revenue loss.

2. Limitations of existing models

The most popular models to estimate the efficiency of advertisement channel are static attribution models. In these models the value of visit is determined by its position in an interaction chain before transaction.

In terms of a Last Click (Pic. 1) Attribution Model 100% of conversion value is attributed to the last channel in a user interaction chain.



Picture 1 – Last Click Model

Since this model is the most common, let's consider its limitations in more detail.

1. This model does not consider the impact of all the other interactions except for the last one.

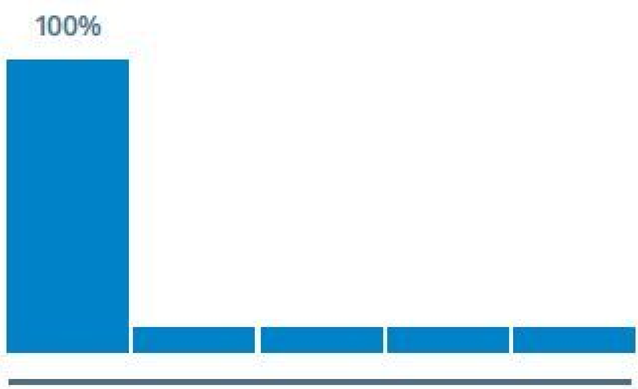
2. ROI is calculated as follows:

$$ROI = \frac{Revenue - Expenses}{Expenses} \cdot 100\% \quad , \qquad (1)$$

where the revenue and the expenses are taken for the same period. The problem is a part of revenue in the beginning of a reporting period is generated due to the expenses of a previous period. And at the same time the revenue at the end of a reporting period will be generated in the following reporting period.

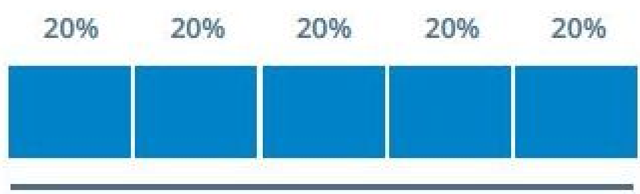
3. This approach considers the revenue on the 1st transaction, but ignores the revenue on the subsequent transactions of this user and the expenses for his/her retention.

In terms of a First Click (Pic. 2) Attribution Model 100% of conversion value is attributed to the 1st channel in a user interaction chain.



Picture 2 – First Click Model

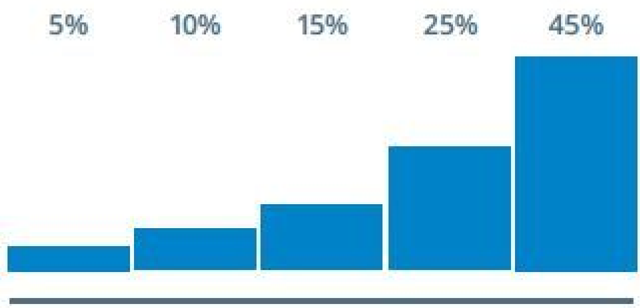
In terms of a Linear (Pic. 3) Attribution Model all the channels in a user interaction chain are attributed the same value.



Picture 3 – Linear Model

3. Attribution based on limitation period

The model is built on time decay (Pic. 4). The closer to the moment of conversion there is a point of interaction, the more valuable it is considered. In terms of this model the period of half life makes 7 days by default. It means that an interaction that took place 7 days before the conversion is twice less valuable than the one that took place the same day with conversion, and the one that took place 2 weeks before conversion is 4 times less valuable compared to the interaction of the same day. Exponential decay is analysed for the whole period of retrospective analysis, usually for 30 days.



Picture 4 – Attribution based on limitation period

The key limitation of all the static models is that they disregard real impact of all interactions to passing the customer journey.

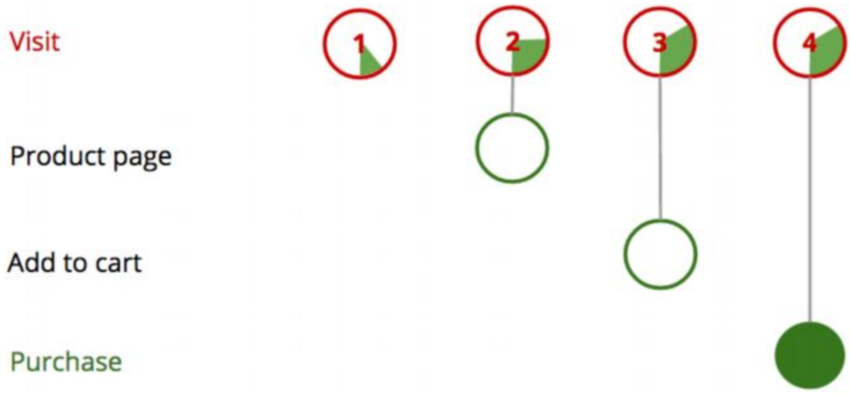
Let's consider an example. For instance there exist a customer funnel including 4 steps:

1. The first visit to a website
2. Visiting a product page
3. Adding to cart
4. Transaction

And there are 4 traffic sources:

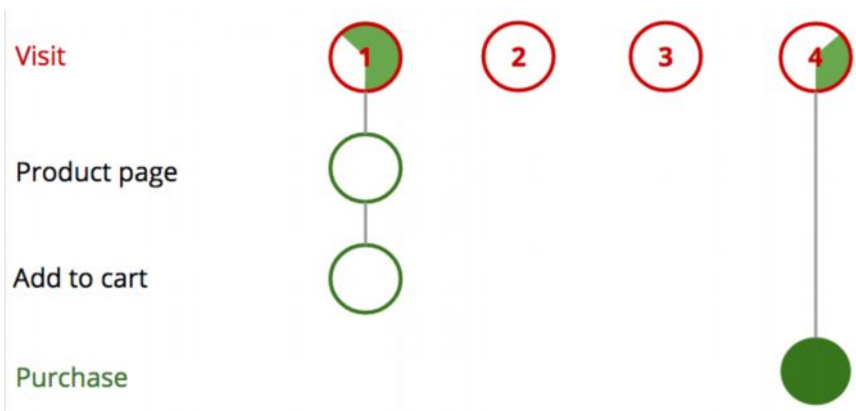
1. A — a channel through which a user came to a website for the first time
2. R — retargeting
3. E — email marketing
4. D — direct

With the same sequence of traffic sources and visits the different scenarios are possible (Pic. 5 – 7):



Picture 5 – Customers journey through the sales funnel

Picture 6 – Product page view, add to cart and transaction take place during the last visit



Picture 7 – Product page view, add to cart take place during the first visit

It is obvious that the impact of a traffic source in each scenario differs, though the sequence of interactions remains the same.

4. Requirements for the model

Thus we can formulate the requirements for the model of traffic sources estimation:

1. It should consider the impact of each transit in an interaction chain, not only the last one.
2. The connection between expenses and revenue should be considered for the users cohorts, not for static periods of time.
3. The value of a visit should not be fixed. It should be determined by the visit's impact to customer passing the sales funnel.
4. The expenses spent for visits that did not result in transactions should also be considered.

5. Solution

The input needed for calculation is the following:

1. User sessions
2. User actions and their priority within interaction chain
3. Values of conversions.

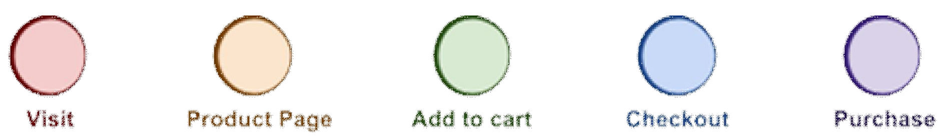
The output of the calculation is the following: conversion value attributed to each session.

Let's base the solution on the following thesis:

1. The business objective is to lead a customer through the sales funnel. That's why we should estimate not only transaction, but also micro conversions resulted in transaction.
2. The lower probability to pass the funnel step, the more difficult it is for a customer to make it. Thus the higher value of a session that resulted in such a transition.

Let's use the following method of calculation:

1. Defining user actions in a sales funnel and prioritize them (Pic. 8)



Picture 8 – Steps of the sales funnel

2. Calculating probability of each user action as a proportion of sessions that resulted in a transition, among those in which the previous transition took place.

Let's have a look at an example where users are distributed between the funnel steps in the following way (Tab. 1):

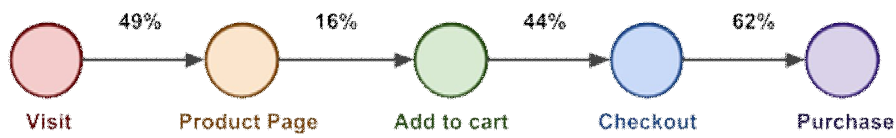
Table 1 – Users share for sales funnel steps

Action	Users
Visit	100.0%
Product page	49.0%
Add to cart	7.8%
Checkout	3.4%
Purchase	2.1%

In this case the probabilities of transits to the next step will be as follows (Tab. 2, Pic. 9):

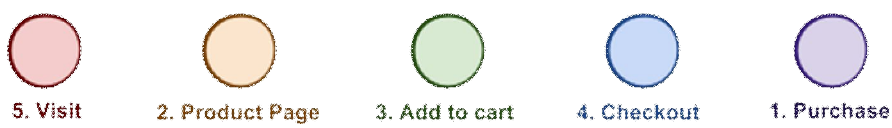
Table 2 – Probabilities for sales funnel steps

Action	Users	Probability
Visit	100.0%	100%
Product page	49.0%	49%
Add to cart	7.8%	16%
Checkout	3.4%	44%
Purchase	2.1%	62%



Picture 9 – Passing probabilities of the sales funnel steps

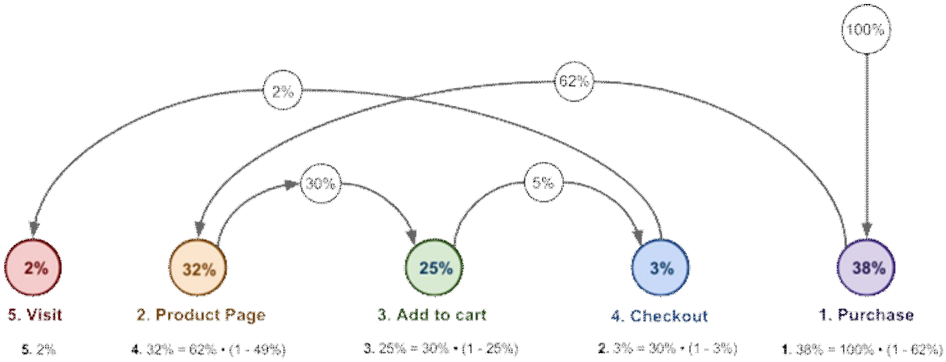
3. Assigning priorities to user actions when attributing values. The prioritization is based on business objectives (Pic. 10):



Picture 10 – Steps priorities of the sales funnel

4. Defining the value for each user action in accordance with its probability and assigned priority.

Attributing 100% of value starts from the step with the 1st priority. For the step with the highest priority the value is calculated as 1 minus probability of transition to this step. For all the other steps the value is calculated in descending order of priorities as 1 minus probability of transition to this step multiplied by 1 minus value attributed for the steps with higher priority (Pic. 11).



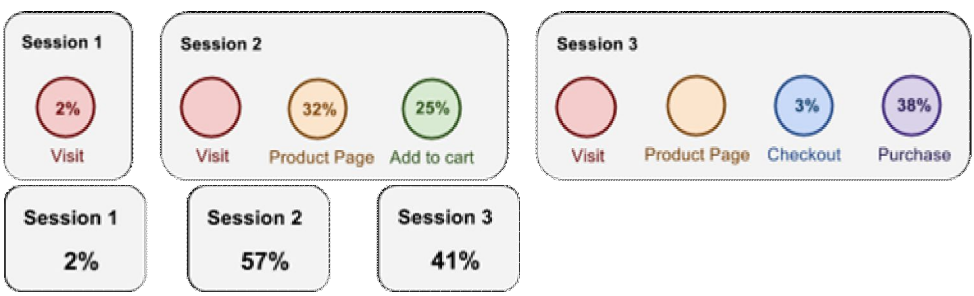
Picture 11 – Value distribtion by the sales funnel steps

The result of sales funnel steps' estimation (Tab. 3):

Table 3 – Value distribtion by the sales funnel steps

Step	Users	Probability	Priority	Value
Visit	100.0%	100%	5	2%
Product page	49.0%	49%	2	32%
Add to cart	7.8%	16%	3	25%
Checkout	3.4%	44%	4	3%
Purchase	2.1%	62%	1	38%

5. The value of session is defined as the sum of values for every user action that took place within this session for the 1st time (Pic. 12).



Picture 12 – Sessions value evaluation based on the steps value

The value of user action is directly proportional to the attributed value and directly proportional to the difference between 1 and probability of this action.

The attributed action is calculated as the difference between the conversion value and all the values attributed to actions with higher priority.

If the sampling is statistically valid, the probability of transition from one step to another has to be segmented between:

– new and returned users;

– depending on the type of a landing page;
– type of traffic source.
Let's have a look at a formula of the session value in a general way:

$$S = (1 - P_c) \sum_{a=k}^{n} (V_a \cdot (1 - P_{v(a)})), \tag{2}$$

where

$$V_a = (1 - \sum_{p=1}^{f(a)-1} V^p) \cdot (1 - P_a) \tag{3}$$

,

$$V^p = V_{f^{-1}(p)}, \tag{4}$$

S – conversion value attributed to the exact session;
k – the minimum step order that has been achieved in this session;
n – the maximum step order that has been achieved in this session;
$P_{v(a)}$ – probability of the value on step-*a;*
P_c – probability of conversion cancellation;
V_a – value of a step-*a*;
P_a – probability of a transit to the exact step of a sales funnel;
$f(a)$ – function the returns priprity of the step by its order a.

Conclusion

The suggested approach allows to attribute the conversion value in accordance with the true impact of a traffic source into the profit from a micro conversion. Unlike static models it allows to adaptively distribute the conversion value. To implement the suggested solution there can be classified as a limitation a necessity to collect enough statistic data.

REFERENCES

1. G. Cristal, Ad Serving Technology: Understand the marketing revelation that commercialized the Internet. – Paperback, 2014
2. M. Jeffery et al., Data-Driven Marketing: The 15 Metrics Everyone in Marketing Should Know Know. – Hardcover, 2010.
3. Mary E. Tyler, Jerri Ledford, Google Analytics. – Wiley, 2006 г.
4. Forrester eCommerce Forecast: http://www.forrester.com/US+eCommerce+Forecast+2013+To+2018/fulltext/-/E-RES115513

Стаття надійшла до редакції 30.07.16.