Leveraging Social Network Analysis to improve prediction of subscriber churn in Telecom

Paras Madan

Subscriber churn is rocking the boat for the Indian telecom operators

Churn is the loss of subscribers from their existing operator to competitor. The Indian telecom operators are suffering with an average monthly subscriber churn ranging from 5 percent to 8 percent in the prepaid segment.

If a subscriber spends INR 100 per month and a telecom operator has a subscriber base of over 5MM then at a 5 percent churn rate the operator stands to lose INR 25MM every month from then on. The operator not just loses these future revenues but also the commissions paid to retailers and distributors along with huge resources invested to acquire and cater to the subscribers. Clearly, the losses are immense and add up very quickly.

Subscribers may decide to cease the relationship with their current service provider for various reasons like:

- Poor network quality in voice calls, frequent call drops, scanty network coverage, weak data connectivity, higher prices than competitors
- Poor resolution of subscribers' grievances by the operator's call center

Operators can predict churn before it becomes a problem

Predicting subscriber churn is critical to an operator's profitability mainly due to the following reasons:

- While telecom penetration is increasing it is fast approaching the point of saturation
- Cost of retaining a subscriber is estimated to be one-fifth of acquiring a new one

It thus becomes imperative for the operators to switch their focus from subscriber acquisition to subscriber retention to maintain and enhance margins from their existing subscriber base.

Traditional models to predict churn have existed for long time. So, what's new?

People tend to influence each other when they are in a social relationship.

Traditional models enable the operators to understand who is going to leave in near future and why. This would help them take appropriate measures in advance to prevent them from leaving. But, with competition being at its peak right now and the industry likely to see a revolution in near future, the need is to add more dimensions to these models as well.

A new and significant factor that we observed impacting churn is 'a subscriber's calling community'. And the effect this community causes on subscriber churn is what is termed as 'Social Network Impact'. Whereas, the current traditional churn models treat a subscriber as an isolated entity, in real world a subscriber is a part of a large calling community. And, Social

Network Analysis captures the impact of relationships among socially connected subscribers in the predictive model, i.e. how these relationships influence a subscriber's usage behavior and loyalty.

Intuitively, a subscriber is more likely to churn if his or her connections also churn or if most of them are subscribers of a different network.

Figure 1 depicts a social network of 2 subscribers – A and B and their connections:

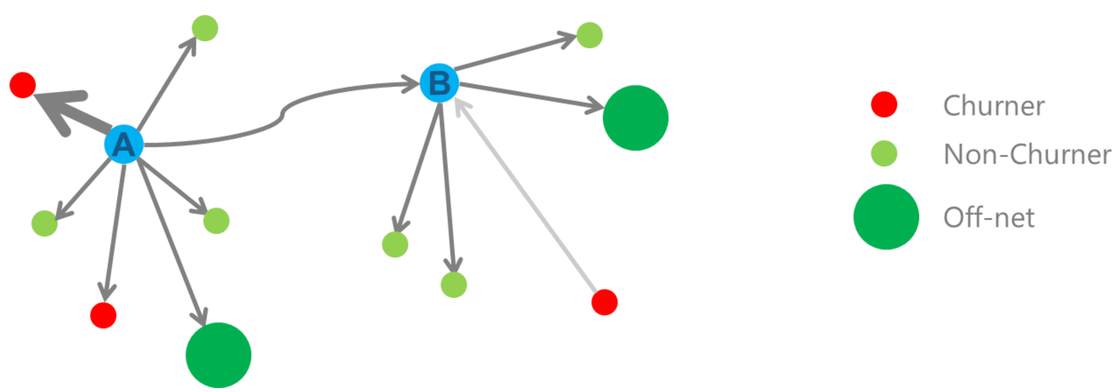

Figure 1 – A telecom social network

A is connected to more number of churners than B. This hypothetically indicates, A is more probable to churn relative to B.

A slightly different implementation of SNA

A typical way of implementing SNA is by way of construction of multi-relational social network and finding influencers based on various centrality measures such as degree (number of direct neighbors of a node), betweenness (number of times a node lies on the geodesic between two other nodes) and eigen vector (a node's influence increases when it is linked to more influential nodes). Influencers are people that are more strongly connected. These influencers then become persuadable subscribers for aggressive marketing campaigns.

However, at TransOrg, we have adopted a novel approach to implement SNA. Normally, a telecom network consists of thousands of disjointed clusters of strongly connected subscribers that communicate within their group only. With the aim of identifying probable churners, it makes sense to look at a subscriber in his micro community as a particular subscriber would not be talking to everyone else in the network. Two important things should be considered here –

- Have connections of a subscriber churned or not?
- Is a subscriber connected to more off-net than on-net subscribers?

Now refer to figure 1. We constructed a directed network in which nodes (red and green dots) represent subscribers and edges (the grey lines) represent their relation. The edges have been given weight according to call volume during a certain period, indicating how closely connected two subscribers are with each other.

Keeping in mind the processing time, one month or a quarter of call records is sufficient as social networks are not observed to change drastically overnight.

Illustratively, for a particular telecom circle that was considered for this experiment, the call network contained 2.03 million nodes and 6.40 million edges.

Following parameters were then defined for the subscribers of this circle to fine tune the churn prediction.

- Out degree for a subscriber i.e. number of distinct connections to which a subscriber makes an outgoing call considering a directed network
- In degree for a subscriber i.e. number of distinct connections from which a subscriber receives an incoming call considering a directed network
- Count of Out or in degree churners i.e. how many of the above connections have churned
- Total outgoing call volume from churned neighbors
- Total incoming call volume to churned neighbors
- Outgoing call volume proportions (w.r.t. total) from churners
- Incoming call volume proportions (w.r.t. total) to churners

A collaborated prediction algorithm observed improved impact

An ensemble learning algorithm was used to build the model. We have seen a substantial improvement in model's accuracy. The accuracy in this context is measured by the coverage of churners in top three probability-wise-ranked deciles.

Past coverage (traditional model)	New Coverage (including SNA)
72.9 %	80.5 %

Table 1 – Comparison of past and improved accuracy

We have also confirmed the robustness of this model by testing it on data sets from different time instants. A consistent improvement of ~ 7-8 % in accuracy is seen across data sets.

In a highly connected world today, SNA guarantees improved performance in predicting churn.

Author(s)

Paras Madan

Analytics Specialist at TransOrg Analytics

www.transorg.com | paras.madan@transorg.com

Contact us

TransOrg Analytics
www.transorg.com

US

Raajeev Aggarwal

President – North America

Mob: +1 703 568 0285

raajeev.aggarwal@transorg.com

Rest of the world

Shuchita Jain

Head – Client Development & Marketing

Tel: +91 124 4231894

Mob: +91 98112 60911

shuchita.jain@transorg.com

Anju Jain

Mob: +91 87655 42092

anju@transorg.com

Gaurav Srivastava (Mumbai)

Mob: +91 70217 96819

gaurav.srivastava@transorg.com

Prithvi Kathuria

Mob: +91 87006 06587

prithvi.kathuria@transorg.com