

Churn Prediction in Telecom Industry using Social Network Analysis

Mr. Varun E
Research Scholar
Department of CS &E
AIT, Chikmagalur

Dr. Pushpa Ravikumar
Professor & Head
Department of CS &E
AIT, Chikmagalur

Abstract - Applied In the telecom portion, an enormous volume of data is being created each day in light of an immense client base. The telecom organizations consider that achieving new client is costlier than holding the current ones. This paper proposes an interpersonal organization examination based model used to distinguish the constructive key clients. The proposed approach assembles a telecom interpersonal organization; it investigates essential system includes, and distinguishes the network structures present in the system. The most significant piece of base mining is inner circle structures which share social structure, use structures and regular interests. The inner circle is a gathering of clients in the telecom interpersonal organization, they share comparative properties. The paper recognizes positive key clients, its club subgroup and that foundation is utilized to distinguish the churn in telecom client's system.

Keywords: Telecom social network, substructures, cliques.

I. INTRODUCTION

In the current world, a monstrous volume of data is being created by telecom associations at a really fast rate. There is an extent of telecom master fighting in the market to assemble their client share. The clients have various options as better and progressively moderate administrations. A complete goal of telecom associations is to support their clients and stay alive in an engaged business showcase. A client stir happens when an enormous degree of clients are not happy with the administration gave by telecom association.

Telecom associations consider arranging move when the amount of customers administration under a particular level which may achieve a huge loss of salary. Client stir distinguishing proof is basic in the telecom division as telecom managers need to hold their productive clients and improve their Client Relationship. The most significant factor is holding existing clients, since procuring new client is as yet costlier than holding existing [1]. The telecom organizations ought to have associations model that perceive and hold their clients by offering moderate and great types of assistance [2].

The informal community examination approach is utilized for agitate forecast in associations. Informal organization examination is the planning, estimating the connections and distinguishing the flood of correspondence between clients. The hubs present in the system are clients and the connection speaks to the connection between every clients. The informal organization investigation (SNA) approach gives scientific and visual methods to point by point examination. To get arrange and their member clients, the diverse centrality measures is utilized. The various estimates utilized for understanding the SNA properties are degree centrality, betweenness centrality, closeness centrality and Eigen vector centrality. The center centrality of informal community investigation is design present in organize connections, the impact on singular client, imperatives and practices.

The regular interests of basic investigation is in the "sub-structures" that is available in a system. A significant number of the methodologies for understanding the structure of a system focus on how thick associations are available and stretched out to create bigger factions or sub-groupings. The inner circles are a calculation created to recognize how bigger structures are exacerbated from littler structures. The particular definition to coteries is on the off chance that there are n clients; at that point there exist every single imaginable tie among themselves and nearby one another.

II. LITERATURE SURVEY

A. Churn prediction

Client beat is a noteworthy issue and one of the most huge concerns for enormous associations [3]. Due to the quick effect on the livelihoods of the associations, especially in the telecom field, associations are attempting to make expects to anticipate possible clients for maintenance [4]. Thusly, finding factors that impact clients to agitate is basic to take basic exercises to diminish the beat. The work on client agitate expectation focuses on

building a forecast model dependent on utilization conduct, use example and AI draws near. The a large portion of SNA model assembled so far focused on single social and homogeneous structure [5]. Another ordinarily utilized methodology for client agitate forecast is choice tree; there is an imperative in choice tree that it isn't reasonable for complex nonlinear associations between the qualities. The choice tree approach has preferences like it tends to be handily envisioned, can process clear cut and numerical information without earlier suspicions. A neural system based methodology is for forecast of agitate additionally can be created [6].

B. Substructure mining

The interpersonal organization contains numerous connections and hubs. The connections and hubs are consolidated with certain relations that can be demonstrated as a diagram. For the network mining these hubs and connections can give distinctive related networks. In interpersonal organization investigation (SNA), the foundation emulating is regularly viewed as a durable sub-network, for example, the factions, n-inner circles, n-groups, n-plexes, just as the semi clubs. The most existing methodology thinks about that there is just a single sort of connection in the system and the aftereffects of mining are autonomous of client inclinations. Anyway in actuality there are numerous heterogeneous social associations every relationship assume particular jobs. The relapse based methodology is utilized to discover the networks; however the methodology isn't appropriate for shrouded network structures. The occasion based methodologies are likewise proposed to identify the network; however they are not reasonable for telecom beat expectation network approach.

III. PROPOSED METHODOLOGY

	A	B	C	D	E	F	G	H	I	J	K	
1	SUBSCR	churn_key	Churn_90	CUSTOMERBILLING	_avg_call_cnt	tot_call	(Latest_earliest	rday	GEOG			
2	8408	8408	0	205090	188659	12.9212963	462	2190	2100	26	293	
3	11504	11504	0	240422	177630	13.27314815	406	2190	2100	23	332	
4	11642	11642	0	1067510	177203	5.708333333	223	2129	2100	23	331	
5	12010	12010	0	326369	765976	7.75	320	2190	2101	24	332	
6	12450	12450	0	2727874	1791120	5.347222222	289	2127	2101	16	325	
7	12739	12739	0	1764199	1792857	6.962962963	290	2190	2100	22	267	
8	12868	12868	0	2481850	1791814	6.791666667	255	2190	2100	25	337	
9	15360	15360	0	1062037	57800	6.662037037	272	2190	2100	26	324	
10	16859	16859	0	377888	63277	12.125	482	2190	2100	26	334	
11	19388	19388	0	1346870	57047	6.49637037	236	2190	2100	26	340	
12	21917	21917	0	1222912	60156	8.75	402	2190	2100	26	333	
13	24165	24165	0	1508226	52783	6.777777778	252	2190	2100	24	311	
14	24318	24318	0	634390	52866	8.333333333	101	2190	2108	6	65	
15	25052	25052	0	1169884	52729	10.00926926	312	2190	2102	24	333	
16	25785	25785	0	5277444	6297555	2.861111111	55	2190	2103	12	257	
17	31967		1	1263138	53964		1	2	2128	2128	1	204
18	33077	33077	0	59950	62347	1.722222222	22	2129	2106	6	338	
19	36056	36056	0	956440	75575	12.23148148	463	2190	2100	26	338	
20	36928	36928	0	1271707	75809	3.333333333	89	2129	2103	14	84	
21	37281	37281	0	808649	68858	33.76851852	1349	2190	2100	26	324	
22	38977	38977	0	380897	74099	5.194444444	64	2127	2100	19	337	
23	39722	39722	0	931463	73886	10.39351852	359	2190	2100	26	327	
24	41321	41321	0	405684	70433	5.456185185	193	2190	2100	25	309	
25	43962	43962	0	14397	86773	6.287037037	244	2190	2100	24	280	

Figure 1: Sample telecom data set

This segment speaks to proposed agitate expectation approach. Figure 2 shows the model and depicts the means. The initial step is information pre-preparing, in information pre-handling covering approach is utilized to perform include choice [7]. The irregular timberland, boruta calculation, stepwise forward choice and stepwise in reverse choice calculations are utilized to separate significant highlights [8][9]. After element choice the comparability measures is determined, which yields nearness lattice. The subsequent stage is to assemble telecom informal organization and estimating diverse centrality esteems to distinguish the positive key client and negative key client. The last advance is to recognize the stir network. The informational collection of 24588 clients is appeared in the figure 1. The dataset is organized information which contains 76 distinct characteristics and 24588 records of clients.

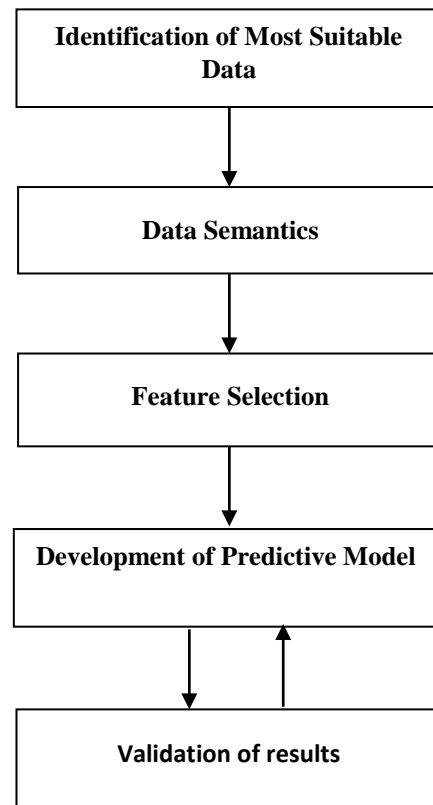


Figure 2: Churn prediction framework using social network analysis

IV. EXPERIMENTS & RESULTS

A. Construction of directed multi-relational heterogeneous network

Telecom informal community speaks to various connections, communications among clients. The system developed for investigation object is multi-social and heterogeneous that is clients of different sort of administrations like post-paid and paid ahead of time. The telecom arrange built utilizing numerous characteristic

from call detail record and the client data. The trait considered are customer_age, number_of_SMS in seven days, number_of_calls in seven days, duration_of_the_calls. Utilizing the euclidian separation comparability measure, the nearness grid is developed. The qualities present in the nearness framework speak to relationship separation between the clients in the system. The telecom interpersonal organization developed by utilizing all the properties is appeared in figure 3.

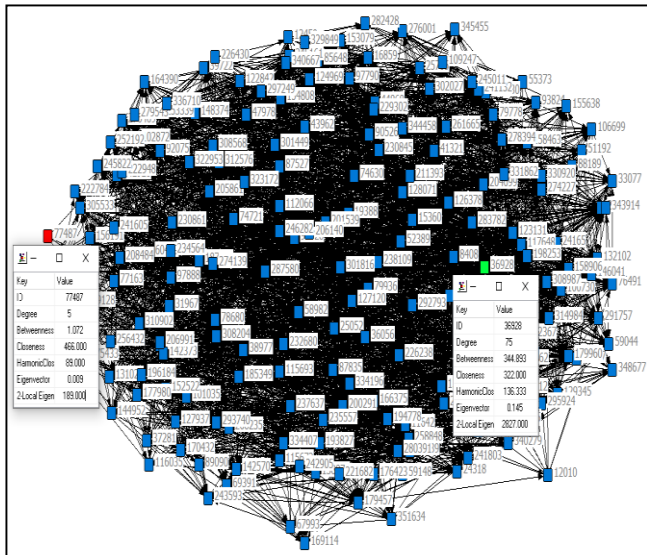


Figure 3: Customers telecom social network

B. Centrality measures

Centrality measure is the most noteworthy undertaking in a diagram. It is utilized to locate the most significant hubs in the client's system. Distinctive centrality estimates utilized are degree, betweenness, and closeness and Eigen centrality. The centrality measure can be utilized to recognize positive key players and negative key players in the system.

Given a telecom interpersonal organization with n clients, to such an extent that, adverse key client implies expelling that n-p set will bring about the system with least attachment. The key client positive methods it is maximally associated with all the hubs. The negative key players of size k=1 is 76899 having less centrality measures.

C. Identifying subgroups

The inner circle in a telecom interpersonal organization is a maximally complete foundation, the clubs with more noteworthy than size k is recognized as a subgroups. The figure 4 shows number of inner circles distinguished in telecom informal community, there are 364 clubs or sub bunches in the given telecom interpersonal organization. The biggest coterie in the system is made out of 6 clients, they share basic properties. The various clubs of size k=3 to k=6 were found.

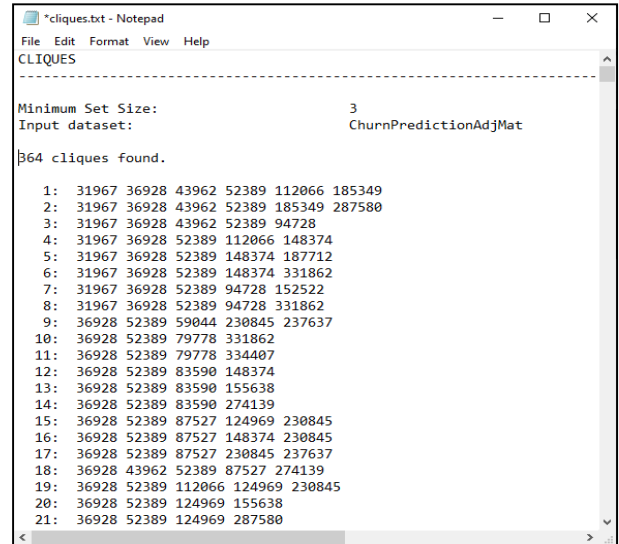


Figure 4: Clique sub-structures present in the telecom social network

There are 364 coterie of size k = 3 to 3 of size k = 6 are found in the telecom informal organization. The gathering of clients like 3196, 36928, 43962, 52389, 112066 and 185349 offer basic properties and all are near one another. Think about coterie with number of clients or hub n, to such an extent that no hub with more n-1 vertices found.

Table 1: Number of cliques found in the overall network

Clique size	Number of cliques found
3	364
4	228
5	39
6	3

Table 1 records all out number of factions found for the telecom informal community. There are 364 coterie of size k = 3 in the clients organize, 228 factions of size k = 4 is found. At the point when the quantity of client hubs expands, the size of longer factions will diminish.

V. CONCLUSION

This paper utilizes interpersonal organization investigation way to deal with discover bases. Clients are allotted with the centrality scores, which distinguishes positive key clients present in the system. In the wake of allotting scores the bases like factions are distinguished. In the event that the inner circle has positive key clients, at that point that faction is recognized as likely gathering and plausibility of stir suggestion is more in that bases. The outcomes contrasted and class mark adjusts 80% of churners present in that bases. The subgroups inner circles

distinguished follows exacting arrangement of gatherings, it tends to be reached out to increasingly loosen up n-cliques.

REFERENCES

- [1] Preeti K. Dalvi, Siddhi K. Khandge, Ashish Deomore, Aditya Bankar, V. A. Kanade, "Analysis of customer churn prediction in telecom industry using decision trees and logistic regression", in Proc. IEEE Symposium on Colossal Data Analysis and Networking (CDAN), pp: 1 – 4, 2016.
- [2] Hui Li, Deliang Yang, Lingling Yang, YaoLu, Xiaola Lin, "Supervised Massive Data Analysis for Telecommunication Customer Churn Prediction", in Proc. IEEE International Conferences on Big Data and Cloud Computing, Social Computing and Networking, Sustainable Computing and Communications, pp: 163 – 169, 2016.
- [3] Ammar A Ahmed, Dr. D. Maheswari linen, "A Review And Analysis Of Churn Prediction Methods For Customer Retention In Telecom Industries", in Proc. IEEE International Conference on Advanced Computing and Communication Systems (ICACCS - 2017), January, pp: 06 – 07, 2017, Coimbatore, India.
- [4] Sepideh Hassankhani Dolatabadi, Farshid Keynia "Designing of Customer and Employee Churn Prediction Model Based on Data Mining Method and Neural Predictor", The 2nd International Conference on Computer and Communication Systems, 2017.
- [5] Dr. Pushpa And Dr. Shobha "Social Network Classifier for Churn Prediction in Telecom Data" 2013 International Conference on Advanced Computing and Communication Systems (ICACCS - 2013), Dec. 19 – 21, 2013, Coimbatore, INDIA [978-1-4799-3506-2/13 ©2013 IEEE].
- [6] B.N. Krishna Sai ; T. Sasikala, "Predictive Analysis and Modeling of Customer Churn in Telecom using Machine Learning Technique", ICOEI – 2019.
- [7] Varun E, Pushpa Ravikumar, Chandana S, Spandana K M, "An Efficient Technique for Feature Selection to Predict Customer Churn in telecom industry", ICAIT – 2019.
- [8] Varun E and Pushpa Ravikumar, "Attribute Selection for Telecommunication Churn Prediction", International Journal of Engineering & Technology, 2018.
- [9] S. Jagannatha, M. Niranjanamurthy, and P. Dayananda, "Algorithm Approach: Modelling and Performance Analysis of Software System", Journal of Computational and Theoretical Nanoscience (American Scientific publishers), December 2018, Volume 15, Issue 15, PP. 3389–3397
- [10] Varun E and Pushpa Ravikumar, "Telecommunication Community Detection by Decomposing Network into n-Cliques", IEEE International Conference on Emerging Technological Trends (ICETT), 2016.