Identification of Valued Users to Generate More Telecom Diligence

N. KHAN*, A. SHAH, K. NUSRATULLAH

Department of Computer Science and Engineering ,Yanbu University College, KSA

Abstract: Stay connected" is a well-known advantage of telecommunication. Telecom diligences have introduced an easy and enormous way to people to keep in touch. From the last decades, a remarkable growth i noted in telecommunication industry. The number of telecomm users has enormously increased. With increasing demands, it is becoming challenge for CSPs (Communication Service Providers) to provide best services. In order to provide finest services a very important step is to monitor the activities of subscribers on the network. Call detail records (CDRs) can play a vital role in monitoring process. CDRs provide a diversity of information about the subscribers, Even though Call Detail Records were initially gathered and stored for billing purposes, but the huge amount of digital data is generated by calling, texting and internet usage etc., generates events, states and errors that provides useful insights.

These digital footprints can be analysed to produce worthy information like network optimization, user's patterns and behaviour identification, valued users identification, suspects and their connected associates and much more. This paper has identified an important use of call detail records to discover and analyze the networks generating more income for offering new services and promotions to those valued user's networks.

Keywords: Valued users, Call detail records, Telecom diligences.
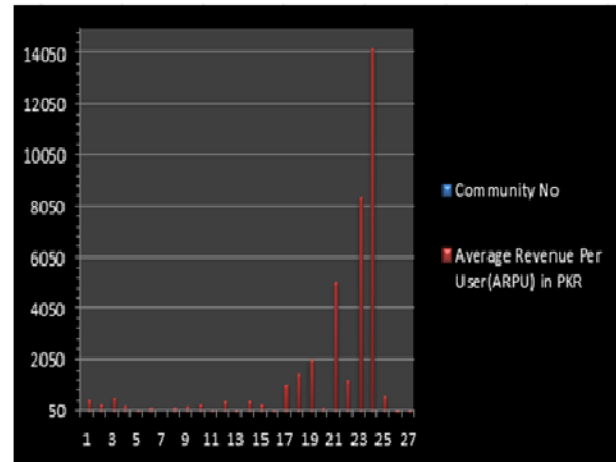
1. INTRODUCTION

The European Commission has published the results of a study on electronic communications (Digibyte, 2016). According to that people prefer to communicate over a mobile phone call (nine out of ten respondents). Another report by Julie Stoughton, in 1990s the mobile subscribers were around 12.4 million. In 2011,the number reached to four billion. Nielsen stated a massive augmentation in the usage of mobile data in a between 2010 and 2011, which is 89%. And global mobile data traffic is on track to increase 13-fold between 2012 and 2017.Hence it is sure that telecom diligence has grown tremendously within the past few years. The number of subscribers is increasing progressively. Increment in subscribers is definitely giving a good corporate to telecom industries but meanwhile it is a challenge as well to retain these subscribers or users. Telecom industries are trying to implement a number of methods to keep their users satisfied. The very first step towards this strategy is to monitor the activity of users in the mobile network. A CDR is a kind of record generated by different telecom kind equipment contains the information of transaction carried out on the telecommunication network. A call detail record is kind of metadata with fields that describe a each and every occurrence of a telecommunication transaction. A call detail records are more itemized information, and Contains features such as the phone numbers of both

parties caller and callee, call start time, duration of that call and call termination cause either normal or abnormal termination. Call metadata can be used to construct user's networks and analyze how they are related to each other (Kiss, *et al,* 2016). To have a network structure that contour on user's contacts, makes easy to analyze the CDRs. The deduced results from these analyses could be very Informative like who are the people whom communicating more? A community that uses to have communication at certain period of time, what could be the best time to offer packages, the role of subscriber in the network or the identification of valued users (Brandes,*et al,*2005).

Community detection is also a kind of analysis that could be valuable for telecom companies. Using networks analysis communities can be defined as how people are linked to each other, and bunching all these into alike groups. It is a statistical measure of connectedness, and furthermore it's not grounded on whether these people in alike groups be familiar with themselves as being part of the same community.

Many other community discovery algorithms either set a universal optimization goal for detection, such as optimizing for modularity, betweenness, or conductance and centerality (Viale 2015) This work is intended to propose a technique that can identify the valued communities that generate

Kulliyyah of Information & Communication Technology, International Islamic University, Malaysia.

Community No	High Revenue Generating Customers	Medium Revenue Generating Customers	Low Revenue Generating Customers	Total Number of Customers in each community	Average Revenue Per User(ARPU) in PKR
1	51	78	113	242	500
2	2	1	4	7	347
3	22	38	59	119	580
4	3	2	4	9	300
5	47	71	95	213	152
6	4	7	17	28	230
7	11	17	19	47	50
8	90	138	201	429	230
9	32	79	92	203	245
10	29	59	96	184	350
11	58	75	157	290	150
12	36	40	56	132	440
13	9	11	17	37	66
14	129	211	343	683	480
15	2	4	2	8	324
16	57	71	151	279	112
17	26	50	65	141	1115
18	30	47	105	182	1508
19	60	52	125	237	2100
20	36	53	63	152	230
21	2	13	18	33	5100
22	50	56	145	251	1300
23	0	1	3	4	8425
24	1	3	13	17	14258
25	3	10	12	25	645
26	3	1	3	7	52
27	4	8	14	26	110

2. PROPOSED METHODOLOGY

The goal of this analysis is to find out the network that composed on valued users. A valued user's network on the basis of CDR is the community of people

. The dataset consist of approximately 45000 CDRs collected from a major telecom operator in Pakistan In order to keep personal discretion, original phone numbers were anonymized by the communication service provider before handing over to us for analysis, and were identified with a security ID (hash code) and many fields were omitted as well. For the sake of simplicity, it is being assumed that all CDRs are on net calls. After cleaning of data the CDRs are transformed in to graphs. Several algorithms are available that can identify the common clusters from the graph structure. In this paper we have used luvain method of community detection which best suits to detect the networks ranging from small to large on the basis of modularity. The density of edges inside communities to edges outside communities is calculated by using modularity that uses a scale value between -1 and 1. In the Louvain Method of community detection, initially the small communities are identified by optimizing modularity locally on all nodes, then each small community is grouped into one node and the first step is reoccurs (Vincent,2008).

$$Q = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_ik_j}{2m}\right]\delta(c_i,c_j)$$

The clustering or identifying subnets within the dataset that is included on whole 45000 CDR, network has also been identified. Each cluster can be identified by its unique community number. These income

generating communities contains only the connected users which make frequent calls to each other. Features which are selected to generate communities are:

- Number of links (calls) between nodes (caller and callee).

- Balance consumed by each user.

(Fig 1): shows the detected communities with their respective income. The valued community has been selected on the basis of maximum profit. The threshold to compare ARPU with is being generated at run time on the basis of current ARPUs values.

Table 1:Top 5 Income generating Communities

Cummunity No	No of customers	ARPU
18	182	1508
19	237	2100
21	33	5100
23	4	8425
24	17	14258

Following is the algorithm to identify the valued user's networks.

```
{Valued Community identification

If ARPUS > 1500

Community ="valued"} //grading of communities
as per maximum income

Else community = "not valued"// Profit is less
than threshold.
```

3.   CONCLUSION

While concluding the results it has been revealed that 80% in telecom sector is generated by 20% users, to whom we named valued customer. All these valued users are somehow connected with some set of people and form the community. In this research paper we have detected these valued networks consists of majority of valued users rendering maximum profit for the communication services providers. By using this mechanism CSPs can find valued users and their networks by calculating ARPU and can introduce specialized offers to these groups or communities. Furthermore we can discover the networks of criminals.. For future research it would like to be recommended that the calls of identified valued networks should be monitored with well-planned techniques in order to treat them differently from other communities. And key players will be identified in each network.

REFERENCES:

Brandes U. and T. Erlebach (2005). Network Analysis: methodological foundations, Springer Verlag,.

Community detection in graphs, (2015) Santo Fortunato,Complex Networks and Systems Lagrange Laboratory, ISI Foundation, Viale S. Severo 65, 10133, Torino,I-ITALY.

Digibyte (2016). New data shows that mobile internet is used more but phone call remains most popular communication.

Kiss, C., A. AndreasScholz, M Bichler, (2006). Soc Networks Identification And Analysis Using Call Detail Records,. Evaluating Centrality Measures in Large Call Graphs. IEEE 0-7695-2511-3/06.

Vincent D. B., J. L. Guillaume, R. Lambiotte, and E. Lefebvre, J. S. Mech, (2008) Fast unfolding of communities in large networks. P10008.