

# Consumer Behavior Analysis through Network captures using Big Data

Vishal Mehta  
Mphasis Limited  
Cybercity, Tower IV  
Magarpatta City, Hadapsar,  
Pune: 411013  
India

## ABSTRACT

For any business it is crucial to listen its target audience, understand market trends and must possess the ability to predict the future of the business. But understanding the mindset of the consumer has always been challenging for the businesses. Hence to understand the requirement of its customers and market trends, companies are investing to develop a feedback platform which can help them to listen and understand their specific customer needs. In this paper we are trying to propose a design of strategy and decision making platform which can help telecom companies to zoom into the consumer needs, in making futuristic strategies and recommending tailor-made offerings to the user.

## Keywords

Packet Data Unit (PDU).

## 1. INTRODUCTION

How many people are accessing the E-commerce website using laptops? How many people are accessing the website using mobiles? What are the geographical locations the users are trying to use the portal? How often they use the service? Are there any alternative services which they use in parallel? What is the average time the users are trying to request for the service in major? How we can retain the high value existing premium customer? How we can onboard the new user to the business? Can we cut down the cost of on-boarding per new user? These are some questions which every telecom company wants to answer but the question is how?

The answer to all the above questions is data. Data is the key. Data is the highest level of abstraction. To decipher the meaning hidden inside the data is science. It helps us answering the bigger questions by using the fundamental data dogma. In this paper we are trying to propose a design of decision and strategy making platform which uses network packets to analyze the behavior of the customer and help businesses to understand and serve their user base better by keeping the privacy of user.

## 2. FRAMEWORK

### A. Decision Platform Design

In the design we propose:

- **Endpoints:** Android phones as datasource of the system
- **Hadoop Clusters:** To store network packets
- **PcapParser:** Parses pcap files to network events
- **Network Event Parser:** To normalize network events to attributes.
- **iRecommender:** recommendation engine
- **Statistical Analyzer:** To detect the trend in the data
- **Consumer Behavior Learner:** learns the behavior of consumer and can predict the future
- **Consumer Cluster Analyzer:** module does analysis on set of consumers

### B. Implementation

Today mobile phones are not just communication devices used for communication but are the gold mine of user behavior. On the basis of data generation source the data can be broadly categorized into two categories:

- Machine generated data
- User generated data

Whenever we try to communicate with some other entity (human beings or software's) on some other part of geography these requests are translated to language (protocol) which network can understand, interpret and can easily transport the message to the other side of the world. Anything and everything we do on the mobile phones today gets wrapped up and decomposed to discreet bundles of bytes of information called packets (abstraction). These network packets not only contains the metadata of the device on which the information got bundled but also its surrounding information, user profile information and all the network transactions taking place with unique geographical and time stamps of the user.

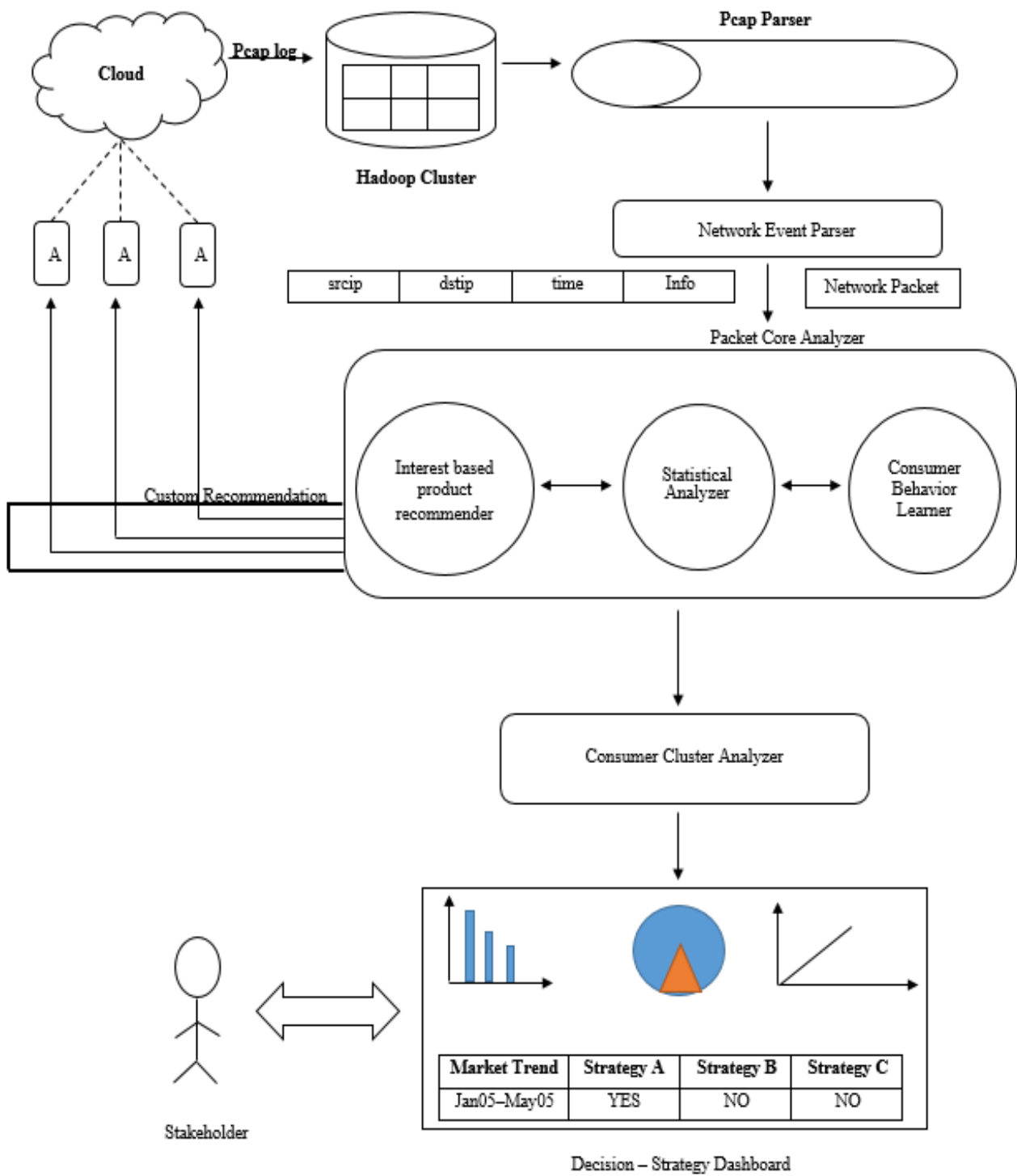


Fig-1

We deploy a packet sniffer on the Android mobile phone which capture every network packet arrive or leaves the network interface. All the sniffed network packets are sent to Hadoop cluster in the form of pcap files. Hadoop collects and stores packet captures from multiple android devices. Once they are collected inside Hadoop clusters they are forwarded to PacketPig Pcap Parser which will parse the pcap files to output one network packet at a time these packets are then forwarded to further get parsed by Network Event Parser to

break down the packet to attributes: srcip, dstip etc. as shown in Fig-1.

As we can see in Fig-2 we have a network capture which shows the sequence number of the packet, date, time, source ip, destination ip, protocol and the type of the request. When this network packet goes into packet core analyzer as shown in

Fig-1 the network packet is analyzed to find the consumer behavior. The fields in the network packet can help us answer the following questions and understand finding the pattern in the network packets.

- Till what time span the user was on the site?
- What all products he searched for?
- What's the daily average time when he surf's the site?
- What is the list of items he searched for and what was the time breakdown for all the products?
- What are the sites surfed in parallel?
- Which are the other sites user is switching to?
- Did the user read the review before buying the product?
- Which review the user stayed maximum time for?
- What is the frequency of the user coming back to the same product page?

If we closely observe on the sequence of the packets we need to observe on the time field. For all the packets we can compare the time fields and see if the value inside the time field is same for all the packets in the sequence we can

precisely tell, till what times the user stayed on the product page. The statistical analyzer module will help in finding the frequency of all the websites the user surfed that day.

The last field inside the packet is the type request made by the user. We can extract the field and split it on the basis of "/". All the tokens received after splitting the information field we get is all the directories of the website on the server user navigated. The iRecommender engine works on basis of correlation values calculated by product and time values. If the product filed shows same for the series of packets and time values keep on changing of all those packets then we can say that user is interested for that product and we can push the tailor made recommendation the user which has high probability that the user will buy that product. The Consumer behavior module learns all the patterns from the Interest based product recommender module and Statistical Analyzer and creates a model on the basis of which it can predict the trend in the future.

On the basis of analysis done by Packet Core Analyzer the information is directed to Decision – Strategy Dashboard where it's visualized and the stakeholders can plan their future strategies and make effective decision on the basis of current information and visible trends. The tailor made plans automatically feedback to the Android mobile phones on the basis of the network transactions by the Packet Core Analyzer.

Date	Time	Scrip	Dstip	Request	
30829	2015-05-10	10:54:58.20	143.255.100.22	16.114.29.150	GET H http://www.amazon.in/s/ref=nb_sb_noss_2/279-5268633- T 2360832?url=search-alias%3Daps&field- T keywords=macbookpro HTTP/1.1 P

Fig-2

## 2. CONCLUSION

The above proposed system will help the telecom businesses to push automated tailor made offers to the user's device. Will help in increasing the user base and will be acting as the eye opener for the stake holders by understanding the consumer behavior and market in the more effective way on the basis of which they can make their strategies on the fly at the time of business need.

## 3. REFERENCES

- [1] Consumer behavior on ecoproduct: why consumers do not buy an ecoproduct IEEE
- [2] Research on Consumer Behavior Influenced by ICT Using Structural Equation Model IEEE
- [3] Classification of Consumer Behavior Using Functional Link Artificial Neural Network, F. IEEE
- [4] A qualitative study in consumer behavior of Skype internet phone IEEE.