ISSN:0975-9646

A Research Study on Accountability of Dynamic IP Generation in a Campus Network (Wi-Fi)

Pawan Prakash Singh, Savita Shiwani, Ruchi Dave, Naveen Hemrajani
Department of Computer Science Engineering
Suresh Gyan Vihar University

Abstract--- **In this paper we have analyzed the algorithms, in which we can analyze the dynamic generation of IP addresses. This concept is based on the application level server logs. The algorithm will take the input (either DNS or IPs) from log on server to produce the identification and signification of dynamic IP addresses for the user in university campus use dynamic IP address. Dynamic IP address may be counted in a fraction of total number of static IP address which can change in a certain time of span. On the basis of this concept we can study and research that how dynamic IP address can use in our university (SGVU) campus WI-FI network. By these results we can analyze the importance of dynamic IP addresses for resolving different types of internet attacks and IP conflicts. We expect the benefits of dynamic IP generation is to enhance the security level of the university WI-FI network and to protect our network from different types of attacks (as virus, spam etc). This concept would produce automatically and identity and help understand IP address dynamics.**

This research study also analyzes the % age usage of dynamic IP generation for particular server log on and the %age of total numbers of IP which has accumulated by the span of particular server log on.

Keywords:DHCP, IP address, entropy, IP Dynamic mapping, mail server.

INTRODUCTION:

In the IP address assignment, e.g. whether IP addresses within an address block are dynamically or statically assigned. It provides some important information and clue in managing and securing the network. For example, there were a large significant amount of malicious activities have been encountered from dynamic IP addresses, as spamming, botnets , fishing and so on. It suspects to us with malicious activities (as email spam) and also allows us to associate multiple instances of such activities from the same dynamic address block [1]. In a campus or company network, dynamic IP address are assigned to the mobile devices (e.g. laptops) which can allocate (net connection) from one place to another place with the use of unprotected networks (e.g. wireless hotspots or access points within the range of that mobile device as located in labs, reception). In case of Suresh Gyan Vihar University, we have allocated two DHCP servers, one is at reception area named as SGVU and another is DLink. The DLink server has assigned by the pool of IP addresses from range 192.168.6.100-----------199. These IP addresses has assigned to the lease period of 2 days in DHCP server.

Hence, knowledge of such address block/pool can assist network operators/security analyst to suspicious activities on these blocks, detecting and preventing attacks from inside hosts. The behavior of network is also established with the knowledge of dynamic and static addresses with appropriate hosts for anomaly detection and behavior tracking. Whether IP addresses are dynamic or static is not be available on the network, even for those within one's own network. This is particularly true for large networks with decentralized management, where large block of addresses are allocated and delegated to sub organizations which control and managed, How these address are assigned and utilized, while it is possible to defer whether an IP address is dynamic or static by its DNS name, such an approach may not always be feasible not accurate for a variety of reason:

- Not all IP addresses have DNS names assigned or registered.
- DNS records are not always well updated.
-

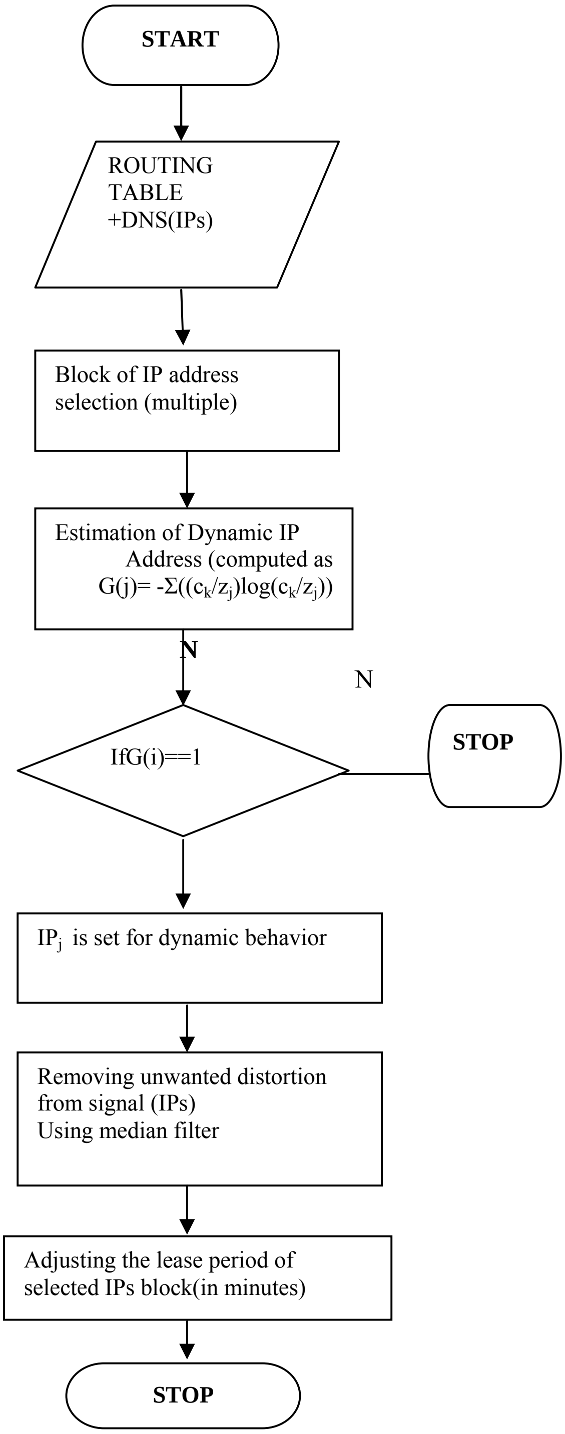
So ,the another methods for accurately classifying IP addresses , in particular for identifying dynamic IP addresses are required[4].

In this paper, we have developed an algorithm as how many dynamic IP addresses are generated and the volatility of addresses within the pool of data collected within one-to-three month's mail server log[9].

Another application, which we use as case study in this paper, is spam filtering. By the spammer it can damage or destroy the original types of mail within the mail server [2]. Therefore, a mail server set up at dial-up or wireless connection is far more suspicious then one set-up with a statically configured IP address.

In other words, whether a mail server is mapped to a dynamic IP address or not is one of interesting feature for existing spam filtering system.

Algorithm for Dynamic IP generation:


START
ROUTING TABLE +DNS(IPs)
Block of IP address selection (multiple)
Estimation of Dynamic IP Address (computed as G(j)= -Σ((c_k/z_j)log(c_k/z_j))
N
N
IfG(i)==1
STOP
IP_j is set for dynamic behavior
Removing unwanted distortion from signal (IPs) Using median filter
Adjusting the lease period of selected IPs block(in minutes)
STOP

This algorithm is based on the selection of static IPs and to exhibit the dynamic behavior on these IP addresses. The fact is that no one IP address is dynamic in the real world ,whatever IPs are submitted to the any server are static. But by applying this algorithm on IPs, we can generate IPs in random fashion (dynamically) to the different users. The algorithm which has shown in the flow chart is divided on following steps:

STEP 1:-SELECTION OF IP ADDRESSES FROM A BLOCK:

The first step of algorithm is to search the IP address from the given block. Due to large number of server log on in the particular time, the IP addresses are accumulating more and more from the different host[4] .

Now it is not easy to count the number of users using the individual IP.

1> All the IP addresses in a block will not appear in the routing table (database or dataset)[2].

2> It is not certain that every dynamic IP address consistently remain dynamic, it may be static by using only and only with single users (as Dynamic IP assigned to a room router that rarely reboots)[3].

Hence this algorithm uses IPs in block of different types of IP address which can use by multiple users. For such, it select set of m continuous IPs from IP_1 to IP_m . These selected IPs block can contain in S(IP_1 ……IP_m) block which can be routing table[2].

This block has following properties:

a> IPs in a block must belong to the same classes (as class A, B, C and D) and mapped with same prefix entry in routing table. As in This observation all IPs are from class C categories. This is starting from 192.168.6.100 to 192.168.6.199.

b> Every block set a minimum size requirement having at least k IPs i.e. m>=k.

c> In the input trace, there must present starting IP address **(IP_1: 192.168.6.100)** and Ending IP address **(IP_m:192.168.6.199).** These blocks should not have any gap.

By using above properties we have defined our own IP block as given below:

The routing table had collected the data from router which has the facility of generating Dynamic IP addresses in random manner(DLINK) in the Suresh Gyan Vihar University Jaipur ,Rajasthan on 4th july 2011 and 15th july 2011.with IPs from 292.168.6.100 to 292.168.6.199(number of IPs are 99 in the block).

And IP_1 =192.168.6.100
IP_m=192.168.6.199

This observation follows all the properties of a dynamic IP block. It can validate that we can use this data for rest of the STEPS of This algorithms.

STEPS 2:- ESTIMATION OF DYNAMIC IP:

For concept : the estimation of IP dynamic is ----Dynamic IP address that had been assigned to multiple hosts and static IP addresses linked to a single host but shared by multiple users.

From practical viewpoint, dynamic IPs are often assigned through random selection from a pool of IP addresses. The probability of occurrence of IP as dynamic can be expected roughly uniform.

For perfection choosing particular IP either dynamic IP address, we use entropy computation:-

This computation is performed on the basis of block-by-block IP selection. Let U is the set of all users and!U! the total number of user's trace. For selection of IP (multiple) S(IP_1……IP_m) with m IPs, we can construct a binary matrix:
X €$\{0,1\}^{!U!X\ m}$
Where K(i,j) to 1 only and only if user i logged on server from IP address IPj.
Now the set of all users U(j) who used particular IPj. The probability of user U(j) used other IP addresses in S(IP_1 ……IP_m). It can be calculated by the distribution law and used as entropy G(j)…

$$G(j)= -\Sigma^{m}_{k=1}((c_k/z_j)\log(c_k/z_j))$$

Where c_k is the kth column sum of Xj and z_j the sum of all the entries in Xj[2].

Real world example:
The setup I will describe is typical of a university (SGVU) office installation .Internet service is provided by a Telephone or cable company using DSL modem. The modem is connected to the DLINK router. The local machines, including the local port of router, are assigned address in a private domain such as 192.168.6.XXX. The external port is assigned a dynamic address by the internet provider's Dynamic Host Configuration Protocol (DHCP) server[1].
The dynamic IP address work fine for most typical application that might be run in an home/office environment, such as web surfing or checking email, but it presents a problem when using a server .If people want to connect to the server from somewhere on the public internet they need to know 18 addresses .
Normally they find these addresses using the internet 's Domain Name Server(DNS).But if no server 's address changes whenever the DHCP server change it then the regular DNS service can't keep with it[1].

What Dynamic DNS does is allow you to establish a URL in one of their domains, and they keep up with changing your IP address. When we enter that into browser DynDNS redirects the request to your current IP address [1].

In our network, we want to able to reach SGVU network from public Internet. There are two problems:
1>The web server is behind a local router, and has an "unreachable" IP address.
2>The IP address of the router is assigned by internet provider is DHCP server, and thus changes without warning.
There are three things that have to be done to solve the problem:-
A>First is to use the DYNDNS service to make the IP address of the router available to the users on the public internet.
B>Second is to program the router to forward requests for web pages to the SGVU server.
C>Third is to configure DHCP of SGVU as DLINK.
In our analysis: DLINK ROUTER FOR VLAN

Router IP address is 192.168.6.1
Subnet mask 255.255.255.0
DHCP IP address Range 192.168.6.100 to 192.168.6.199
DHCP lease time 1440(minutes) (24 hours)
In IP trace of 4th July 2011 is as: AT 2.30 to 3.20pm

Hardware address	Assigned IP	Hostname
00:a1:b0:80:11:8c	192.168.6.100	Chetan
00:1f:3c:e4:70:b6	192.168.6.169	Avdeshg
00:19:e0:84:ba:9c	192.168.6.170	Net
00:a1:b0:91:f8:60	192.168.6.171	Sneha
C4:17:fe:8c:f7:b7	192.168.6.172	Uphar-pc
4c:0f:6e:d7:84:b7	192.168.6.173	Me-VAIO
00:a1:b0:90:c8:ad	192.168.6.174	As-6d605fc30b99
00:19:5b:ce:d4:b8	192.168.6.175	admin
00:19:21:2b:22:f6	192.168.6.177	Def –dc8d3cd8823
C4:17:fe:c8:f8:f5	192.168.6.178	AJ-PC
00:a1:b0:91:f8:b7	192.168.6.180	Admin-d6ad75e78
00:1f:3c:d1:2b:a5	192.168.6.181	Sksingh
00:a1:b0:91:ef:79	192.168.6.182	It
00:00:00:00:00:00	192.168.6.183	
00:19:5b:ce:b2:0f	192.168.6.184	Admin
C4:17:fe:45:67:a3	192.168.6.185	Satyashrawa
70:f1:a1:89;e9:39	192.168.6.186	Pushpendra-PC
00:15:af:bf:4c:d0	192.168.6.187	xx-d4688654adb2
00:19:d2:1b:a0:3d	192.168.6.188	COMPAQ
2C:D2:e7:ae:29:95	192.168.6.189	
00:16:44:65:78:d6	192.168.6.190	User-PC
00:0c:e7:00:00:00	192.168.6.191	
00:a1:b0:90:ce:6e	192.168.6.193	Work
00:19:5b:ce:d4:bb	192.168.6.194	Dba48
C4:17:fe:d5:cb:17	192.168.6.195	Viren-PC
00:14:78:ee;05:cd	192.168.6.196	ADMIN
00:17:c4:45:b8:ad	192.168.6.198	divedi

But in our routing table has the IPs = S(IP_1 , IP_2................IP_m)
=(192.168.6.100 to 192.168.6.199)
m =99Total number of user in one trace= !U!=27
Now using algorithm:
X€$\{0,1\}^{!U!\ !\ X\ \ m}$ = X€$\{0,1\}^{27\ X\ 99}$
X(I, j)== 1 if and only if i has log server; Now we can take compute the user who use only U(j) particular IPj. For such we form a sub-matrix Aj$^{!U(j)!X\ m}$
The entropy calculation for the data k=1, 2, 3, 4, 5, and 6
G(1)= -((1/99)log(1/99)
=-((1/99)(log1 –log99))
= - ((1/99)(0-log99))
= (log99)/99
=1.923/99
=0.193

AS:
G (2)= .193 +0.2932= 0.4862
G(3)=0.638, G(4)=0.754 ,G(5)=1.008, G(6)=1.5

Since the block size m may vary across different multi-user blocks, we define two normalized versions of usage entropy, called normalized usage –entropy G_B(j) and normalized

sample usage entropy $G_U(j)$. $G_B(j)$ is $\log_2 m$ time fraction of G(j) . And $G_U(j)$ is $\log_2(!c(j)!)$ time fraction of G(j).
C(j) : G(j) quantify the probability skew ness only across the set of IP address (denoted as c(j) , that were actually used by U(j).
In ideal case, where IP addresses are selected randomly from the entire block, we can expect the normalized usage entropy $G_B(j)$ most of IP address to be closed to 1.

But for the DYNAMIC IP Generation C(j)-> m and also $G_U(j) \rightarrow G_B(j)$.

STEP 3: **Reorganization of Dynamic IP address (block):-**

In previous steps we have calculated the entropy for IP selection, it might be include that those IPs with entropy closed to 1 are dynamic IP address.

Means G(j)==1 for dynamic behavior of IP address

As in previous we have selected these IPs with multiple selection criteria. So for identifying these IPs addresses, we can convert the multiple-block into different types of sub-block of IP addresses [2].

In our observation each multi-block IPs are segregated into multiple IP addresses, and entropy can be calculated for in the basis of particular IP address not a block. This is because our observation is working on small number of block size.
If the majority of IP addresses are tested for dynamic IP address and declared OK, then we use a pre-specified threshold h for solution IP traffic problems[9].

To achieved the fine grained and smooth signal. We use well known median filter method for suppressing isolated out of range noise. This method replaces every signal value with the median of neighbors. Specifically for the IPi smoothed signal value S'(i) is computed as :
S'(i) = median({s(i! – w/2!-------s(i!+w/2!))
Where w is parameter of median filter which determine the neighborhood size[2].
The main idea of the median filter is to run through the signal entry with the median of neighboring entries.

The pattern of neighbors is called "window" which slide entry, user the entire signal. For 1D signals, the most obvious windows is just first few preceding and following entry, where 2 D (or higher dimensional) signal such as images, more complex windows patterns are possible.

Note that if the window has an odd number of entries, then the median is simple to define. It just the middle value after all entries in the windows is sorted numerically .For an even number of entries there is more than one possible median.
For This algorithm, we have a predefined threshold he. The entropies lower then this threshold he value .We do not apply the median filter , due to small volume of IP block does not have enough boundary conditions.
At last applying the median filter, we can recognize the IP block easily.

STEP 4: **Adjustment of lease Period of IP block:**
This is the final step in classifying dynamic IP block for particular of lease time, means till when IP address could be changes. For this the algorithm considers two matrices here. Which are as follow:-

1> The number of distinct log on server user that has used this address in input data.
2> The average log on server inter-user duration e.g. interval b/w two different users, consecutive in time, using the same [2].

How sequentially change the I.P address is called leased lines. In our algorithm, this can be done by averaging the dynamic IP blocks with respect to given time (in days). As in our case the lease time is given to the DHCP server is 1440 minutes. It means 1440/60 = 24 hrs, means leased period are 2 days. And also we can analyze that, the I.P from 192.168.6.100 to 192.168.6.199 has 99 IPs But at the time only 27-30 IPs enable.
Why? Due to another IPs has used as previous day. So the destination of IPs START from 192.168.6.100 - 192.168.6.198.

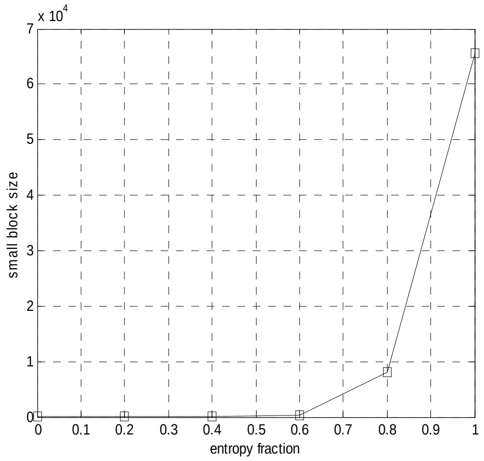
- Also the ending IP 1932.168.6.199 and starting IP 192.168.6.100 also appears in our observation list.
- Lease time also yields by the volatility estimation, also can easily filter the first class of the just positive by removing IPs with large number of concurrent accesses[4].
- More specifically our algorithm discards consecutive IPs address that are each associated with a large number of users(>= 1000) and also exhibit very short inter user duration(<= 5 minutes)
- It can be accomplished by RDNS (reverse domain name server) lookup table [3].

VALIDATION OF ALGORITHM:

	IPs type	Number of IPs	Number of block
UDMAP IPs	Static	102,941,051	958,822
THIS ALGORITHM IPs	Static	199-255	128

TABLE.1. Dataset of UDMAP and THIS algorithm

In above mention table, the UDMAP has taken the large number of IPs whenever I have observe the same thing with small number of IPs[6]. The standard graph shown in the figure 2.



Graph 1: shown with large number of IP block and with standard entropy computation.

This is a standard graph to our observation with fixed entropy computation. We can analyze by the graph that for large number of IP blocks graph is certain or steady for small number of IP blocks with fixed commutative fraction of entropy.

In table 2, data of UDMAP with fixed entropy computation.

S/N	Cumulative fraction of IP block with fixed entropy(X)	Large block size (Y)
1	0	2
2	0.2	8
3	0.4	32
4	0.6	256
5	0.8	8096
6	1	65536

Table 2: IP block selection with fixed commutative fraction of IP block

Relatively, if we take small number of IP block with fixed entropy and computation entropy calculated by this algorithm which is shown in table 3.

S/N	**Fixed entropy (X1)**	**Observed entropy(X2)**	**Small block sige Y**
1	0	0.192	2
2	0.2	0.485	4
3	0.4	0.638	16
4	0.6	0.754	32
5	0.8	1.008	64
6	1	1.6	128

Table 3: IP block selection with fixed and observed entropy computation.

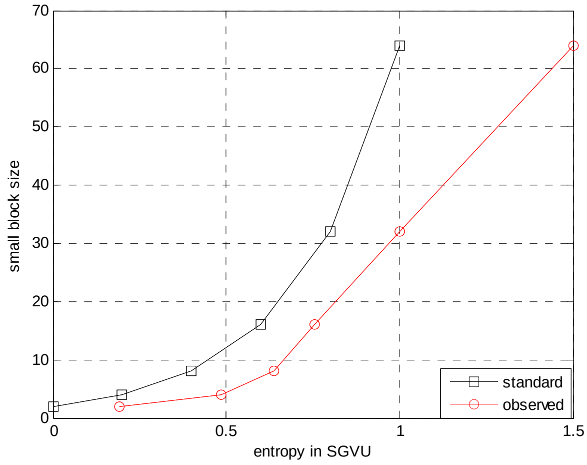
Relatively the graph computed by UDMAP is used large number of IP blocks which shown in graph 1. Now for observation I have compacted the data of large IP blocks in small number of IP block as in table 3 shows the extreme limit of 128.

The graph between fixed entropy computation and small number of IP block is mentioned by BLACK color, whenever graph between observed entropy computation and small number of IP block is mentioned by RED color.

The BLACK color graph is standard for the small number of IP block, whenever RED color of graph is following to the BLACK graph at small number of IP block, when the IP block exceed it scattered, due to the number of dynamic IP generated more frequently.

When G (j)==1 then IP will be dynamic.

So in our case the DYNAMIC IP generated more frequently.



Graph 2: shown with small number of IPs with standard and our observed entropy computation.

RESULT:

In our research, we have study about the algorithm of dynamic IP generation and analysis of algorithm for limited data set. As example university campus .We have compared the analysis with graphical representation in between standard of Microsoft data analysis and our data set analysis.

In our study graph 1 represent the IP distribution with large IP block size as in graph shown it start from 0 to 60000 data block. Whenever our research studies, we have taken data from D-link router which can produce 256 IP addresses at a time which we examine in the algorithm.

For such examinations we have made miner changes in the algorithm which we have study in our literature survey. These changes are as follow:

1. DNS has been directly mapped into IP addresses in our algorithm instead of router table convention. It means the algorithm first step has reduced due to the conversion of DNS to rDNS (reverse DNS) the implication has reduced [7].
2. Due to large IP block in our literature survey our data are selected in multiple blocks but in algorithm we are dealing with small number of data set. So we are selecting IP multiple not the IP blocks.

3. In our literature survey ,there were use of Entropy computation for selection of dynamic IP block, whenever in our algorithm we are not dealing with IP block but dealing with selected IPs for computation .So we have taken the entropy on the basis of 2(binary).which reduce the complexity of the algorithm.
4. As the analysis of graph 2:

The dataset graph is deviated with standard graph. The deviation can be calculated as:

%tage efficiency for
G(1)= (observed/standard) *100
= (0.193/0.20) *100
=95%

%tage efficiency for G (5)=(0.75/0.8)*100
=93.7%

So we can conclude that our algorithm is average (95+93.75)/2= 188.75 / 2

= 94.3% efficient.
Rest of 5.7 % of data IPs has used for SPAM, BOTNET and FISHING attacks.

FUTURE WORK

This algorithm is basically use for counting accountability of dynamic IPs in our University campus.
For future work, it can be enhanced for tracing the different types of attacks (SPAM, BOTNET and FISHING).If we will look our input trace of DLINK router there might be some continuous IP addresses are missing from the trace.WHERE is these IPs? The answer will be use by other host for another purpose.So in future we will emphasis our focus on to trace these IPs or block which are affected by malicious attack.

It can also solve the problem of unauthorized IPs or website to not access in any university campus (content filtering).

It can be used as software firewall, which will work as anti-virus, anti-spam software. Means in future there were no needs of anti-virus or anti-spam software in our campus.

CONCLUSIONS

This paper concludes that dynamic IP addresses can provide the security for any network (Wi-Fi). Dynamic IP generation can solve the problem of IP conflict problem. Means any user can analyze the static IP in the system but can not analyze the dynamic IP address in the system.
This algorithm would automatically generate IP addresses and adjust the lease time for dynamic IP addresses. This detailed small-scale dynamic IP generation showed that majority of IP addresses are owned by the various user which can access our WI-FI network, and hence some of the IP addresses are missing ,which produced the knowledge that there must be possibility of attacks. In our observation that only for one router there was 94.3% usage of dynamic IP address for accessing the internet. Rests of 5.7% of IP are used by the attackers.

REFERENCES:

[1] Cisco Network Register user’s Guide. May 2011.

[2] Y.Hie, F.Yu, K. Achan, E. Gillum, M.Goldszmidt and T.Wobber. “How Dynamic are IP Addresses”. In Proc. Of ACM SIGCOMM, 2007.

[3] Filippo Geraci and Roberto Grossi February 2011,”Distilling Router Data Analysis for faster and simpler Dynamic IP lookup Algorithms”.

[4] Yu Jin, Esam Sharafuddin, Zhi-Li Zhang “Identifying Dynamic IP Address Blocks Serendipitously through Background Scanning Traffic”. University of Minnesota. April 2011.

[5] R.K. Ahuja , T.L. Magnanti, and J.B Orlin“Network Flows: Theory. Algorithms and Application “Printice Hall ,1993.

[6] Sanjay Darisi “Analysis of IP Lookup algorithms using Controlled Prefix Expansion “, Computer Science and Engineering Department Arizona State University. June 2011.

[7] S.Foo, S.C .Hui, S.W. Yip, and Y.He, “Approaches for resolving Dynamic IP Addressing. Internet Research: Electronic Networking Applications and policy”. April 2011.

[8] , Murali Kodialam and T.V. Lakshman “Integrated Dynamic IP and wavelength Routing in IP over WDM network.”, Bell Laboratories .may 1997.

[9] S.Bhattacharya and K.Xu . Profiling internet Backbone Traffic: Behavior models and Applications .july 2011.

[10] D.Comer and D.Stevans “internetworking with TCP/IP” Prientice Hall, Englewood Chiffs July 2011.