Observing Internet Worm and Virus Attacks with a Small Network Telescope

Uli Harder, Matt W. Johnson, Jeremy T. Bradley

and William J. Knottenbelt

Imperial College London

1 Network terminology

- IP-networks can be classified by their IP number range: $A.B.C.D$ where the letters range from 0 to 255.
- A class C network would for instance consists of all IP numbers $A_0.B_0.C_0.0$ to $A_0.B_0.C_0.255$
- Similar definitions for class A and B networks.
- To work out the number of possible networks and IP numbers in a network keep in mind that some ranges are not "routable".
- They are perfectly legal to use but any legal router should not route theses packets from one network to another. (e.g. $10.0.0.z$)
- IP uses "ports" for programs to communicate from one computer to another (e.g. $A.B.C.D.80$ would be the http port of computer $A.B.C.D$)

2 What is a Network Telescope?

- A passive network telescope is an unused part of the IP address space
- For instance our network telescope is a class C network that hasn't seen use for a few years now.
- We use `tcpdump` to monitor the network traffic that is destined for this network.
- Our network telescope does not "respond" to any incoming traffic
- There are other network telescopes which actively respond to incoming traffic, "honeypots"
- Some of these honeypots are very sophisticated and pretend to be almost any OS running any application.
- Due to differences in the implementation of IP (and TCP/UDP) an OS and its patch level can be almost uniquely identified from its response to network traffic (see `nmap`)

3 What events can we possibly see?

- Generally speaking host and portscans
- This is mainly caused by automated programs, a.k.a viruses or worms
- Due to the handshake protocol of TCP we only see the first or incase of spoofed addresses second packets of conversations
- For UDP traffic we see the complete payload, however due to the lack response this usually peters out quickly as well.
- Denial of service (DOS) attacks, In order to disable services like webservers one can spoof IP addresses and flood the comuputer at the TCP level
- Sometimes attackers will use our network telescope fro these attacks and we see the "back-scatter"

4 What does the network one-way traffic look like?

- Distribution of IP numbers and ports chosen by attackers
- We look at the inter arrival rate of packets
- Inter arrival rates of events, i.e. separate attacks
- Draw a 3d picture of 1h of network traffic by plotting
- Look at the autocorrelation of the packet rates of particular subsets of the traffic by computing its power spectrum and using the "detrended fluctuation analysis" (DFA).
- We look in particular at the Sasser worm and a denial of service attack

5 Summary statistics

Type	Frequency in %
S	90.4
UDP	4.8
NBT	2.0
R	1.6
ICMP	1.2

Table 1: Traffic type distribution of the traffic in period 1 and 2.

Rank	Port # per'd 1	cum. frequency	Port # per'd 2	cum. frequency
1	135	0.39	135	0.42
2	445	0.65	445	0.64
3	1433	0.71	1433	0.74
4	1025	0.75	139	0.78
5	80	0.79	1025	0.83
6	139	0.83	38293	0.87
7	38293	0.86	80	0.91
8	26943	0.90	137	0.93
9	137	0.92	ICMP	0.95
10	6129	0.93	6129	0.95
11	2745	0.94	2745	0.96
12	ICMP	0.95	1434	0.96

Table 2: Cumulative frequencies of destination ports in period 1 and 2.

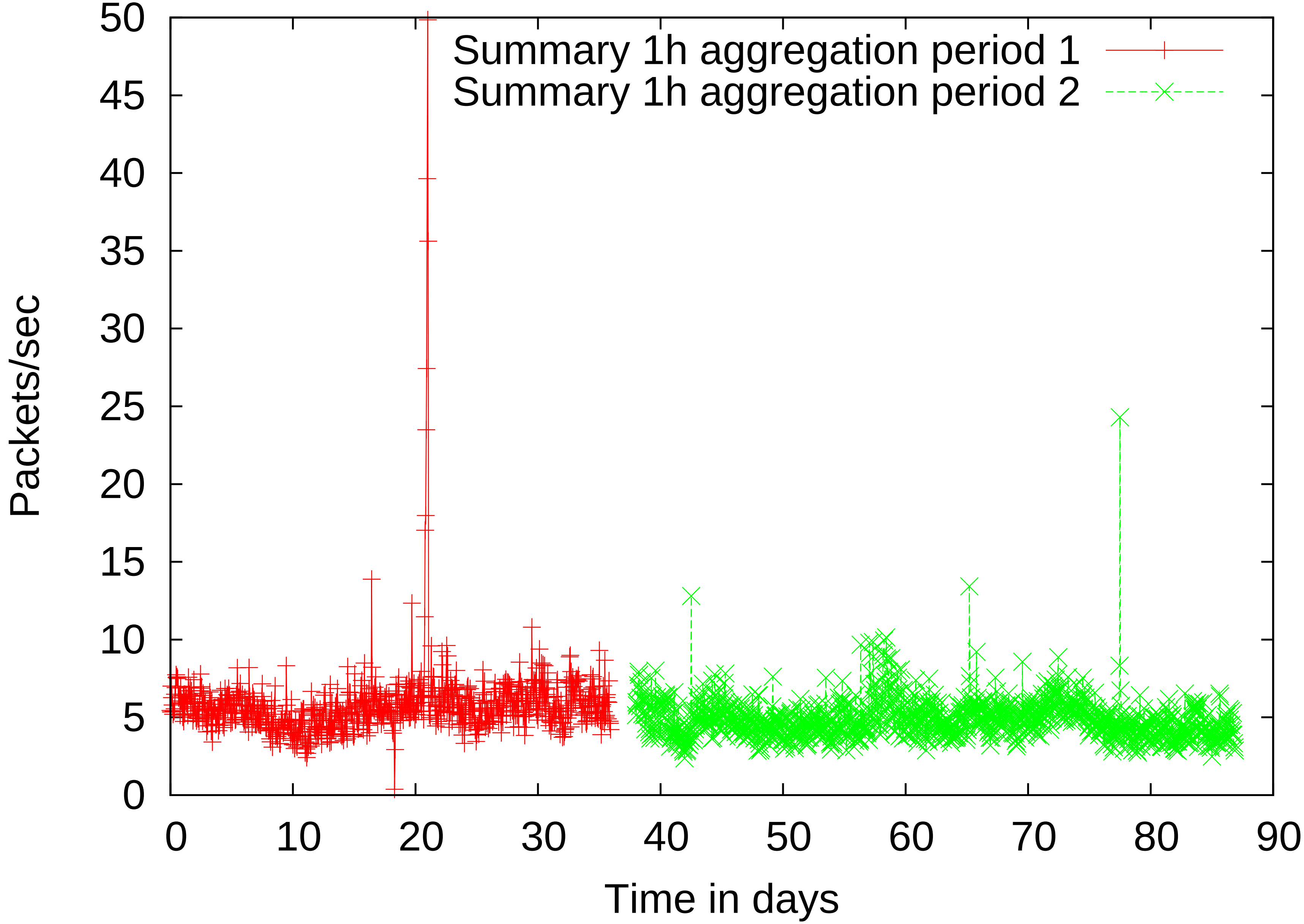


Figure 1: Summary of the observed network traffic.

6 Port and IP frequency plot

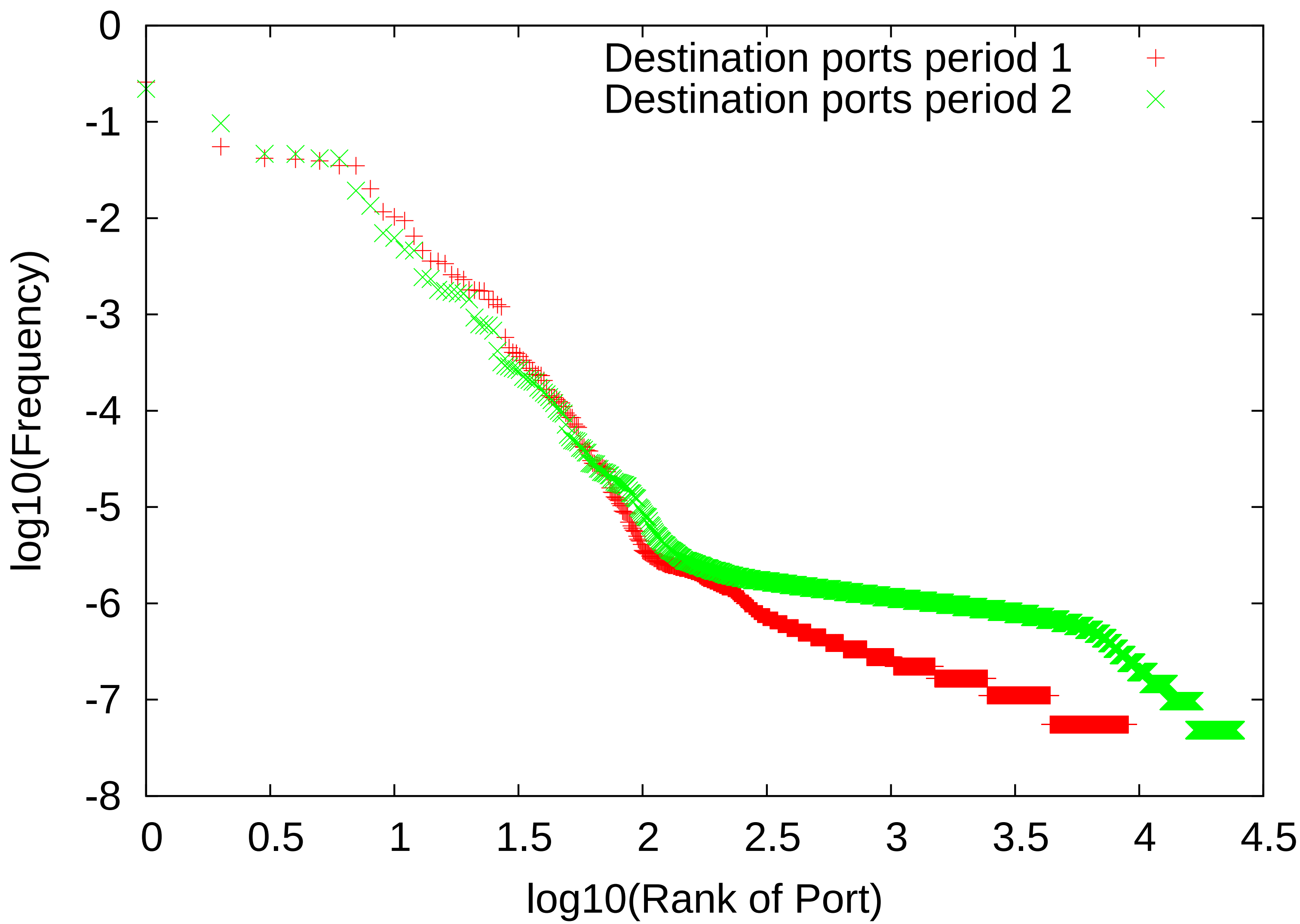
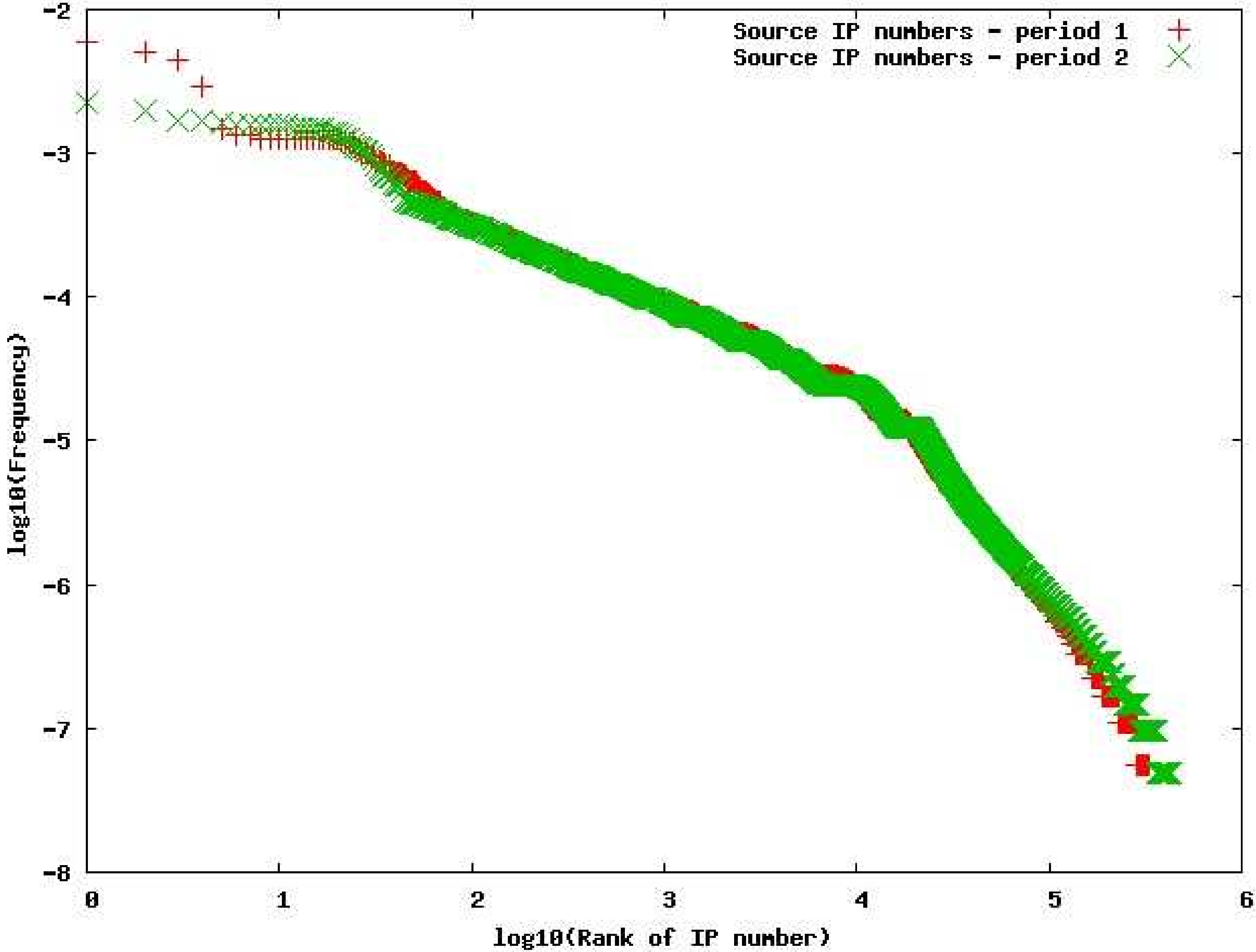
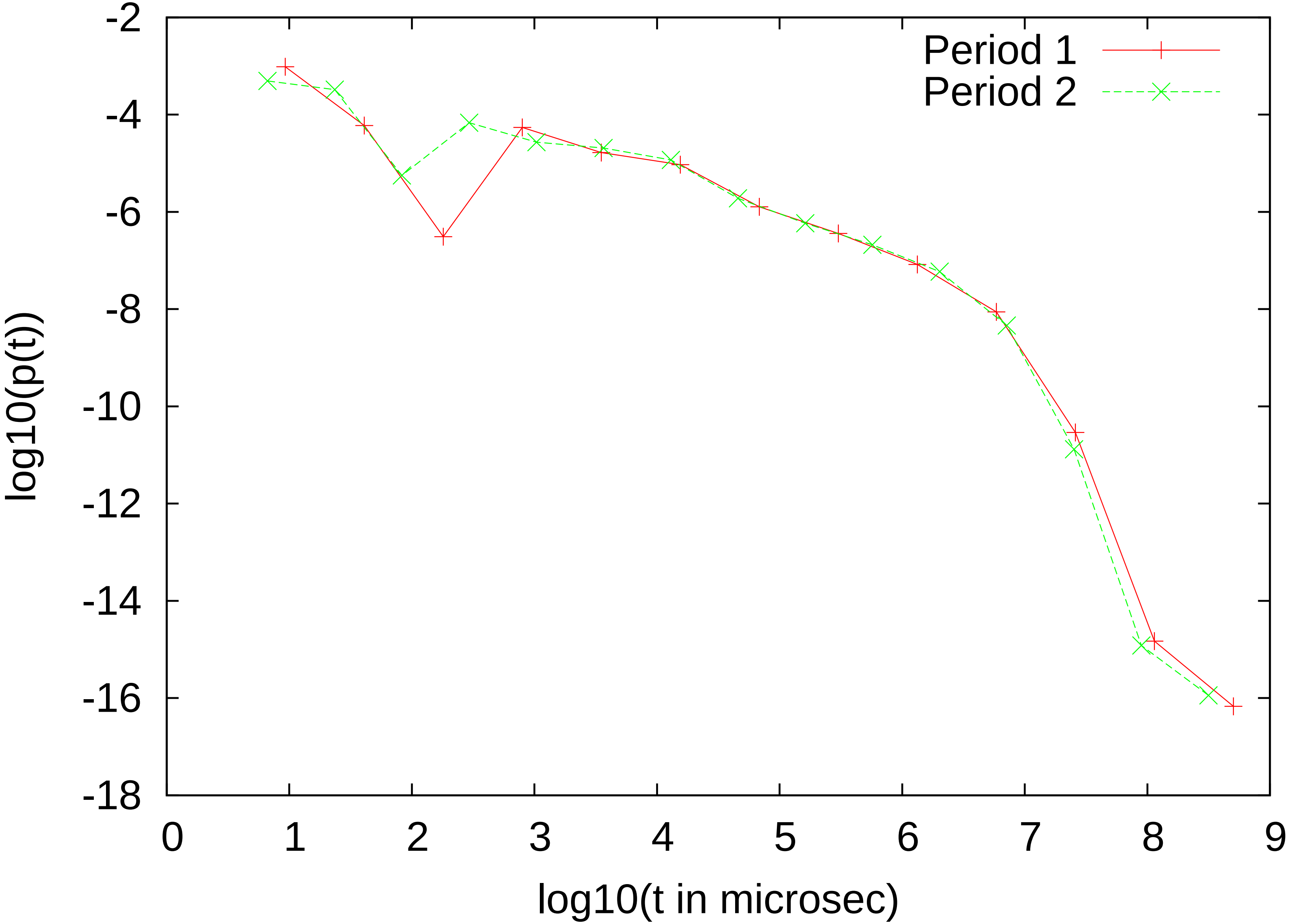


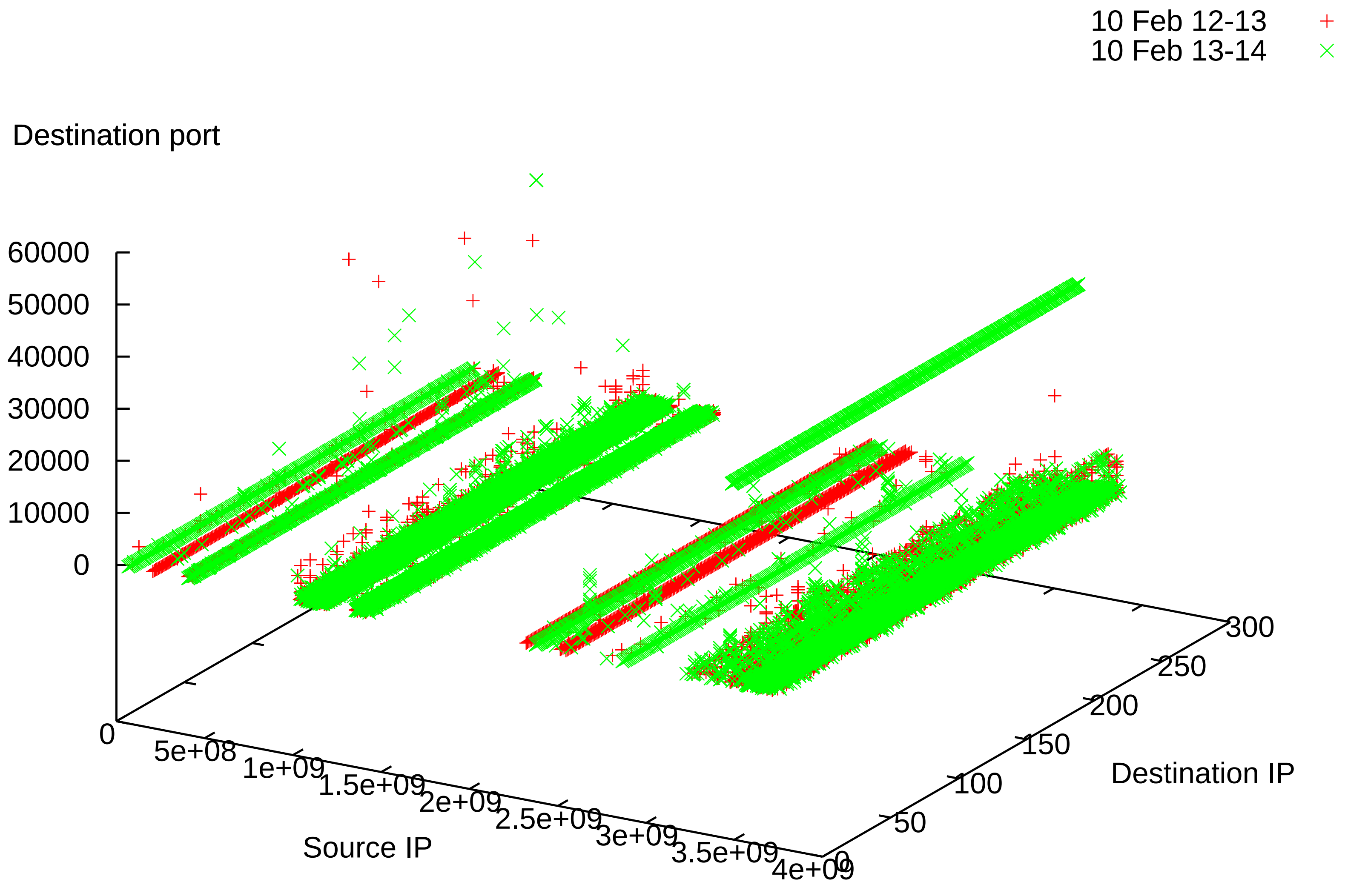
Figure 2: Rank-frequency plot for the destination ports and the IP numbers in period 1 and 2.


Source IP numbers - period 1
Source IP numbers - period 2
log10(Frequency)
log10(Rank of IP number)
-2
-3
-4
-5
-6
-7
-8
0
1
2
3
4
5
6

7 Inter arrival times of packets



8 A 3d picture



9 Correlated traffic?

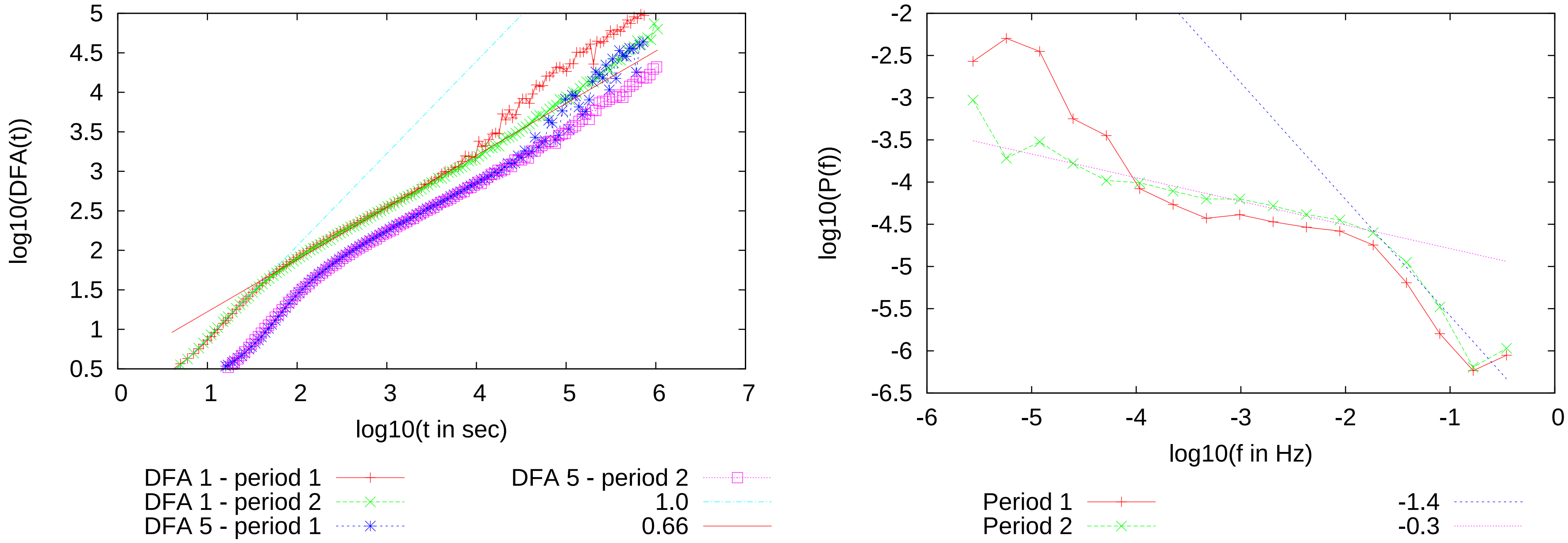
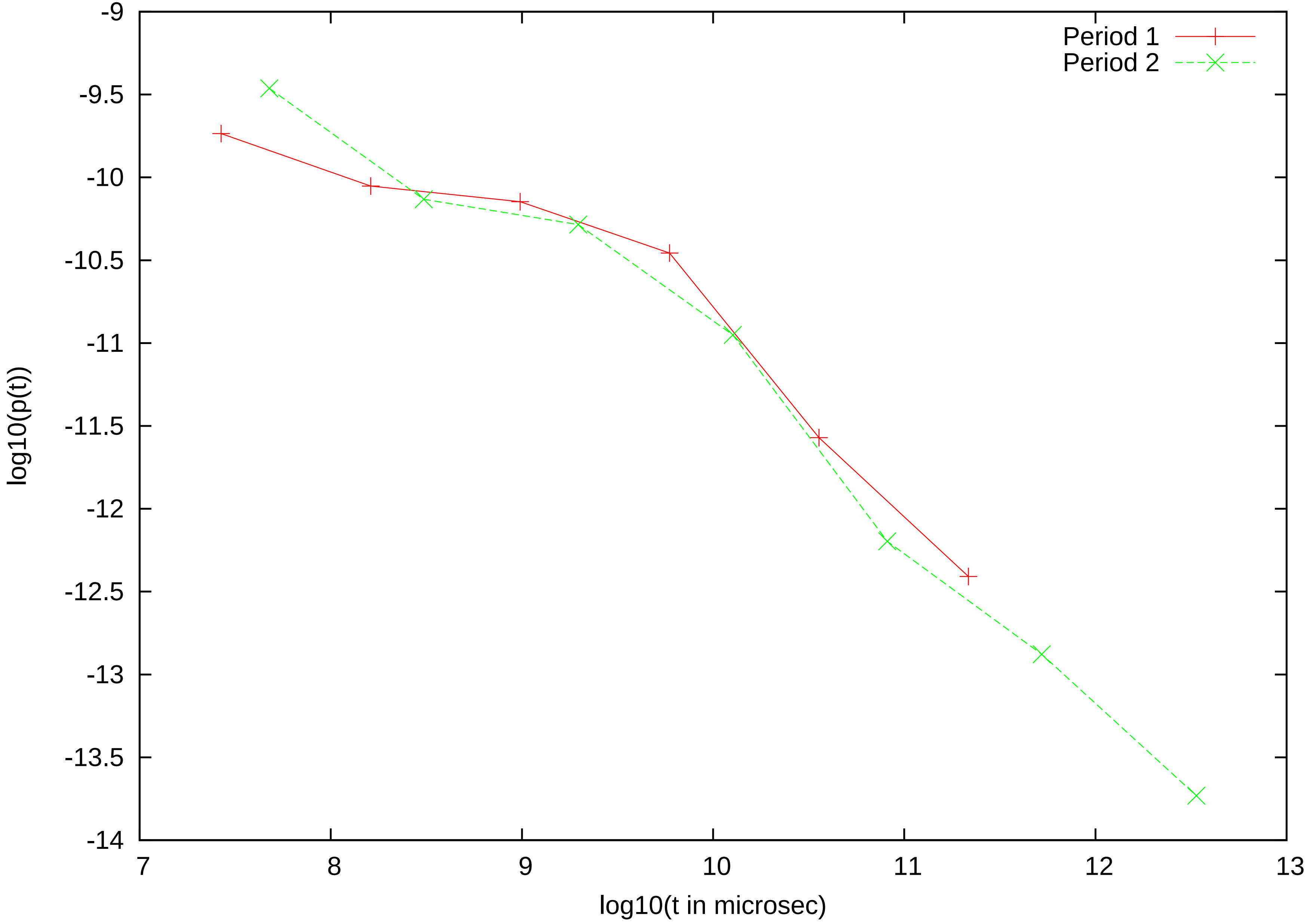


Figure 6: Looking for long range dependence using the DFA (left) and power spectrum method (right).

10 How does Sasser work?

- Attacks services on port 445
- 3 simple rules to choose a class C network for the next attack. If the infected host has got the IP number A.B.C.D
  1. a random number with probability 1/8
  2. A.x.y.0 randomly with probability 1/2
  3. A.B.x.0 randomly with probability 3/8
- Backs off when a machine in chosen network is already infected

11 Time between attacks



12 Are they correlated?

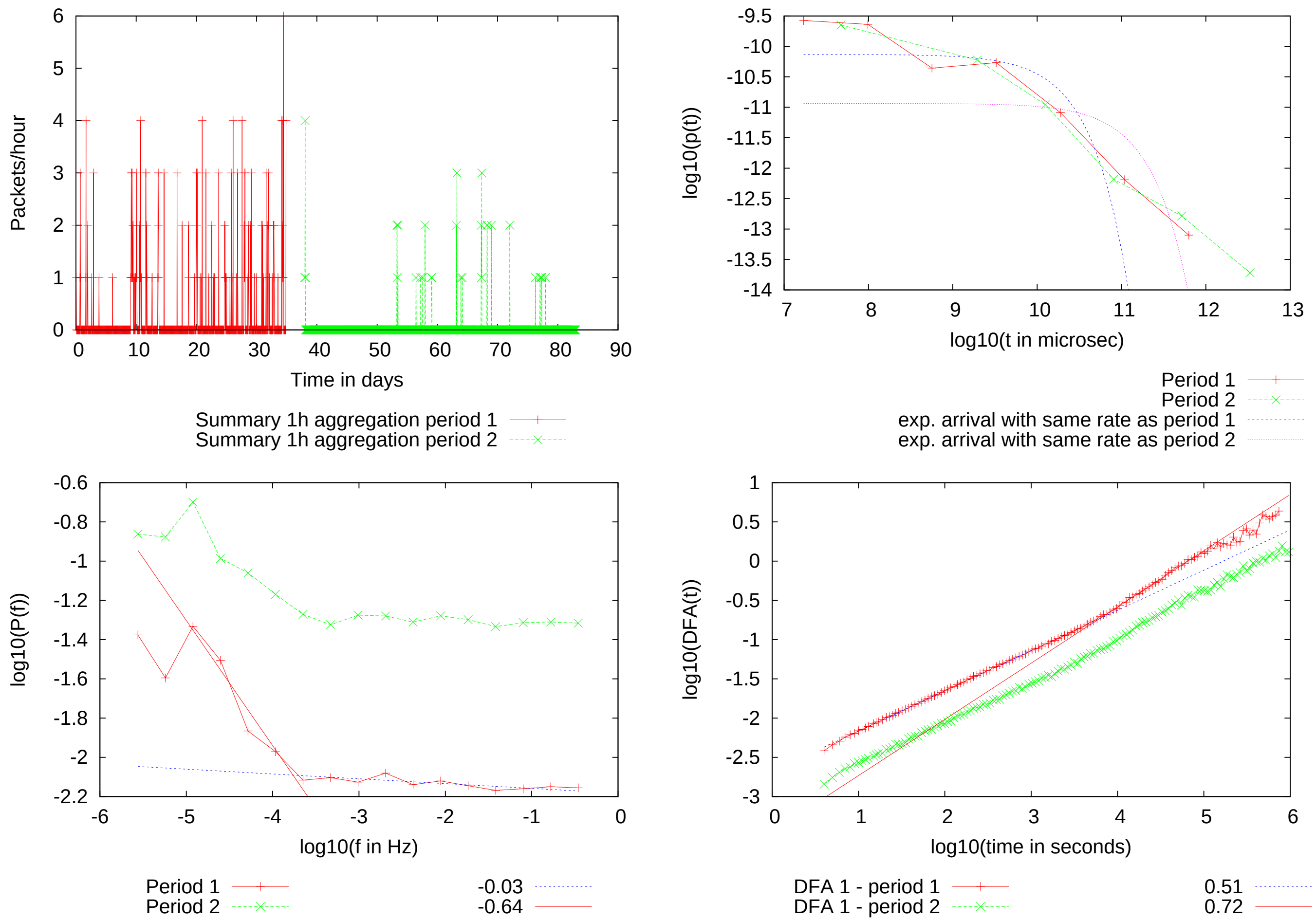


Figure 8: Details of the behaviour of one particular attacking IP number. In the top row the packet rate plot and the inter-arrival time distribution of all packets from this attacker. In the bottom row the power spectrum and the DFA plot.

13 A denial of service attack

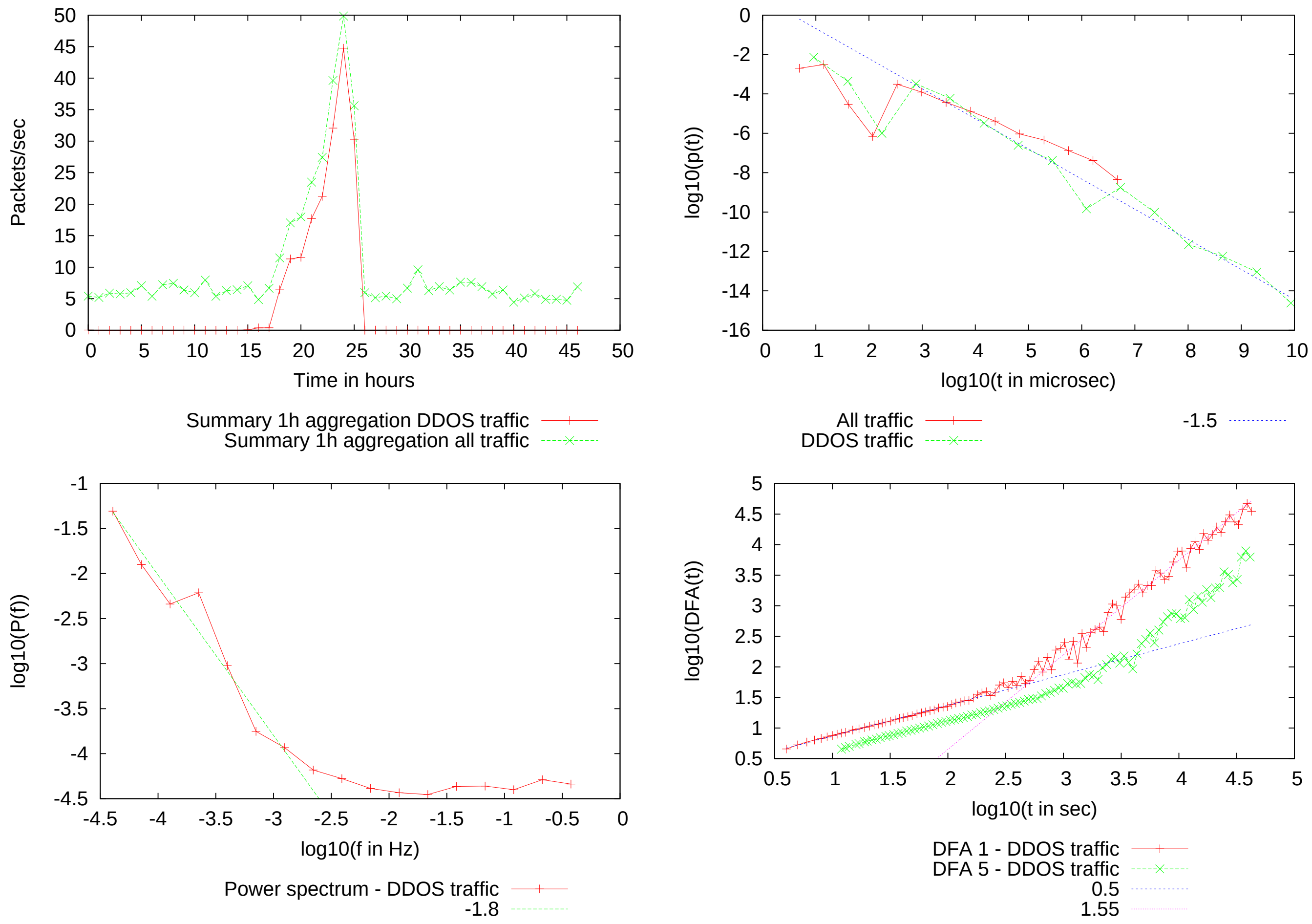


Figure 9: Details of the backscatter from a denial-of-service attack. First the packet rate summary and the inter-arrival times. In the bottom row the power spectrum and DFA.

14 What attack rates would we expect?

- Attack rate of an infected machine is fairly well known.
- $$\lambda_A = N_C \lambda_O \tag{1}$$
- For one particular attacker IP address we observe a rate of 5 attacks per day in the first and 1 attack per day in the second period.
- Given that there are 254^3 class C networks, this amounts to $\lambda_A = 1200$ attacks/sec or $\lambda_A = 300$ attacks/sec.
- Sasser variants scan 510 to 40,960 IP addresses per second, meaning anything between 2 to 160 attacks per second.
- Notably, the observation in the first period is much higher than anything suggested in [**?**].
- At closer inspection we find that the attacker IP address appears to be part of

an address pool that is assigned via DHCP to ADSL subscribers.

- The total number of observation of attacks in the first period is 223 and in the second is 45.

15 Conclusions

- No long range correlation
- Data to verify models of attack?
- Ratio of infection the same for A, B and C class networks
- Rank-frequency plots show power-law
- Rate of 6 packets a second