Darknet experiment at SINET

(Sept. 2006 ~)

Kensuke FUKUDA
National Institute of Informatics, Japan
kensuke@nii.ac.jp
Goal of (my) study

• Effective monitoring for unwanted traffic detection
  • for smaller and distributed address blocks
• Prediction of traffic pattern by using spatial and temporal knowledge of anomaly

As a first step, we try to statistically quantify darknet traffic
Darknet

- Darknet is routed subnet, but with no hosts (network telescope, network sensor system,...)
- Coming packets to Darknet is something wrong
- portscan, DDoS, worm, misconfiguration
- Experimentally, we run /18 subnet darknet (=16384 addrs) in our network
Weekly darknet traffic

- /18 (16384 addrs) blocks
- mean: 19kbps, max: 200kbps
- dumpfile: 100MB/day
TCP Dport (24h)
UDP Dport (24h)
Source addr breakdown (12h)

(IP addr -> ASN -> Country)

- TCP SIP
  - EU(11451), CN(9754), KR(7566), JP(4456), US(4449), TW(1651), DE(528), ZA(399), NL(328), AU(159)

- UDP SIP
  - CN(21422), US(2948), EU(2640), DE(795), PE(729), JP(722), ID(575), CA(410), HK(371), KR(349)

- ICMP SIP
  - US(7391), KR(124), EU(105), CN(51), TH(9), IN(8), NL(5), JP(5), FR(5), TW(4)

- Is there any geographical difference??
Temporal correlation of traffic time series
Scaling analysis

- DFA (Detrended Fluctuation Analysis) [Peng98]
- Detection of LRD in a given time series
- Estimated scaling exponent: $\beta$
  - $\beta = 0.5$: random walk
  - $0.5 < \beta \leq 1.0$: LRD (= Hurst parameter)
  - $\beta > 1$: non-stationary time series
- Reconstruct /24 block time series (bin = 1 min.) from 1-day trace, then apply DFA to the time series
Scaling exponent (TCP)

- Weaker temporal correlation (≠ random fluctuation)
- Possibility of prediction (?)
Scaling exponent (UDP)

- Most values are around 0.5: random fluctuation
- More than 1.0, fluctuation is non-stationary (= anomaly)
Raw time series (/24)

- TCP: correlated fluctuation
Raw time series (/24)

- UDP: random fluctuation
Raw time series (/24)

- UDP: non-stationary fluctuation
Results

• TCP:
  • Time series is LRD
  • Possibility of prediction by AR model(?)

• UDP:
  • Time series is random
  • Anomaly can be found by DFA

• Further analysis
  • different block size time series (/18 <-> /32)
  • Port-level time series
Spatial correlation between two time series of address block
per-address packets (12h)

- Difference between 1st and 2nd /24s
- No widely-spread icmp probes?
Spatial correlation

- Investigate the similarity of temporal traffic pattern
- Correlation coefficient between two time series of /24 address block apart from distance D
  - $-1 \leq \gamma < 0$: anti-correlated
  - $\gamma = 0$: non-correlated
  - $0 < \gamma \leq 1$: correlated
Spatial correlation

- Correlation between two /24 block time series
- TCP: no correlation apart from 20 blocks (6144 addrs)
- UDP: larger correlation and some synchronized blocks
Results

- TCP:
  - No correlation apart from 20 blocks (6144 addrs)
  - Periodic assignment of monitoring blocks(?)

- UDP:
  - Larger correlation and some synchronized blocks
  - Existence of important/unimportant blocks(?)

- Further analysis
  - Dependency of block size (/17 -> /32)
  - Port-level analysis
Concluding remarks

- Temporal and spatial correlation of darknet traffic time series
- TCP is weak LRD, UDP is random walk
- Spatial correlation lasts to only 20 /24-blocks for TCP, and some synchronization of blocks is appeared in UDP

Future work

- Port-level and smaller address block analysis
- Possibility of comparison with CAIDA data? (problem: our measurement started from sept. 2006)

- Geographical and IP addr space differences?