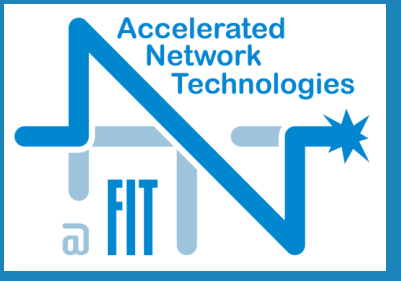
Network Anomaly Detection and Classification via Opportunistic Sampling

Georgios Androulidakis et al.

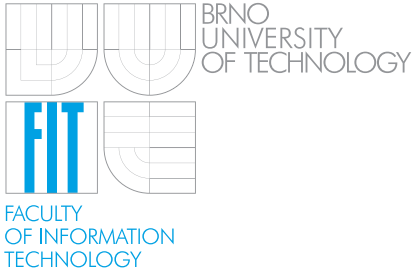
IEEE Network

January/February 2009



Vlastimil Košař

Brno University of Technology, Faculty of Information Technology
Božetěchova 2, 612 00 Brno, CZ
www.fit.vutbr.cz/~ikosar



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Obsah

Úvod

- Obtížné zpracovávání flow na vysokorychlostních sítích
- Dlouhou dobu se myslelo, že vzorkování snižuje úspěšnost detekce anomálí
- Nedávno bylo zjištěno, že detekce založená na entropii vzorkováním příliš netrpí
- Vhodným výběrem vzorkovací metody je možné dosáhnou zlepšení úspěšnosti detekce anomálií
- Poskytují flexibilitu a řízení potřebné pro vyvíjející se obsah

Útoky

- DDOS útok
  - Odepření služby legitimním uživatelům
  - Mnoho zdrojů útoku
  - TCP SYN, ICMP, UDP
- Šíření internetových červů
  - Škodlivý kód, který se sám šíří
  - Snaží se šířit přes specifickou zranitelnost
  - Infikovaný stroj zasílá malé množství paketů velkému množství cílů
- Port scan
  - Sken velkého množství portů na cílovém počítači

Neočekávané legitimní události

- Flash crowd
  - Velké množství legitimních uživatelů se snaží využívat službu najednou
  - Vzestup množství přenášených dat v příchozím i odchozím směru
- Alpha flows
  - Vzestup zatížení linky pouze pomocí několika velkých spojení
  - Typicky přenos velkých souborů, nebo síťové experimenty

Oportunistické vzorkování

- Preferuje nějakou vlastnost při vzorkování
- Používá se na úrovní toků
- Klasická pětice:
  - SRC IP
  - DST IP
  - SRC Port
  - DST Port
  - Protocol
- Pro preferanci používají autoři délku toku
- Autoři navrhují použití 2 algoritmů vzorkování
  - Selektivní vzorkování
  - Chytré vzorkování

Popis

- Preferuje krátké toky
- Mnoho útoků se projevuje krátkými toky
  - DDoS
  - Šíření červů
  - Sken portů

Vzorec

$$p(x) = \begin{cases} c & x \leq z \\ \dfrac{z}{n \cdot x} & x > z \end{cases},$$

Chytré vzorkování (Smart sampling)

Popis

- Preferuje dlouhé toky
- Detekce legálních událostí
  - Flash crowd
  - Alpha flows

Vzorec

$$p(x) = \begin{cases} x/z & x < z \\ 1 & x \geq z \end{cases},$$

Entropie pro detekci a klasifikaci anomálií

- Zkoumáme změny rozložení pravděpodobnoati vybraných vlastností
- Celá skupina přístupů
- Nezávislé na topoligii sítě a jejích charakteristikách
- Používána normalizovaná entropie

Vzorec

$$H_n(X) = -\frac{\sum_{i=1}^{N} p_i \log_2(p_i)}{\log_2 N}$$

Některé výhodné charakteristiky

- SRC IP
- DST IP
- SRC Port
- DST Port
- Délka toku

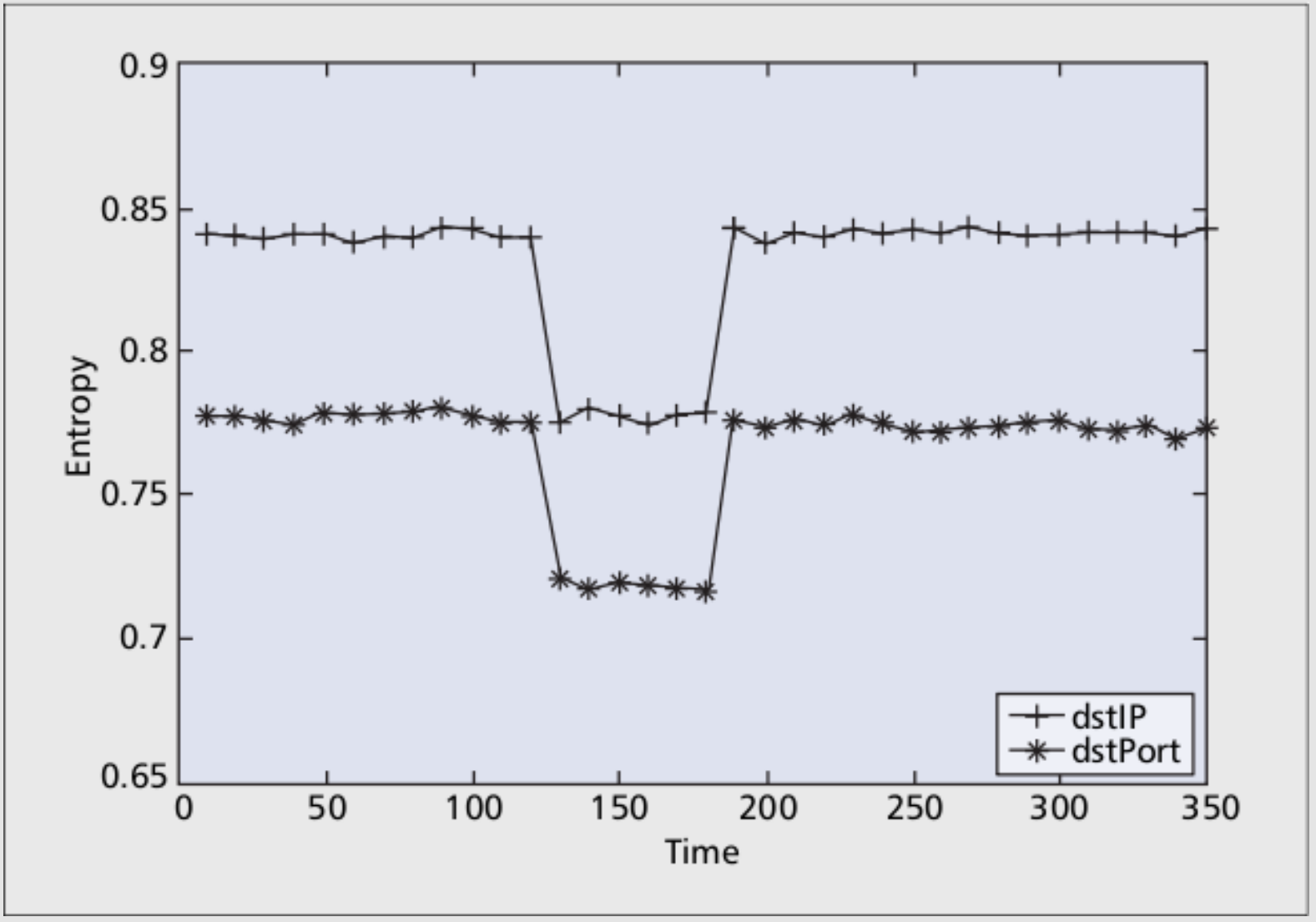
Projevy útoků na entropii

Anomaly	Description	Entropy change
Distributed denial of service (DDoS) attack	An attack on a specific service, making the resource unavailable to its users	Significant decrease in dstIP and dstPort. Almost no change in srcIP, srcPort, and flow-size.
Worm propagation	A self-replicating program that tries to infect other machines by exploiting a specific vulnerability	Significant decrease in srcIP and dstPort. Slight increase in dstIP and srcPort. Slight decrease in flow-size.
Portscan	Sending probe packets to a wide range of ports in a specific host to check which services are available	Significant decrease in srcIP, dstIP, and srcPort. Slight increase in dstPort. Slight decrease in flow-size.
Flash crowd	A large demand for a specific service (i.e., many clients downloading a specific file from an HTTP/FTP server)	Slight decrease in srcIP, dstIP, srcPort, dstPort, and flow-size.
Alpha flows	A small number of flows that have a very large quantity of packets (data transferred between two specific hosts)	Slight decrease in srcIP and dstIP. Almost no change in srcPort, dstPort, and flow-size.

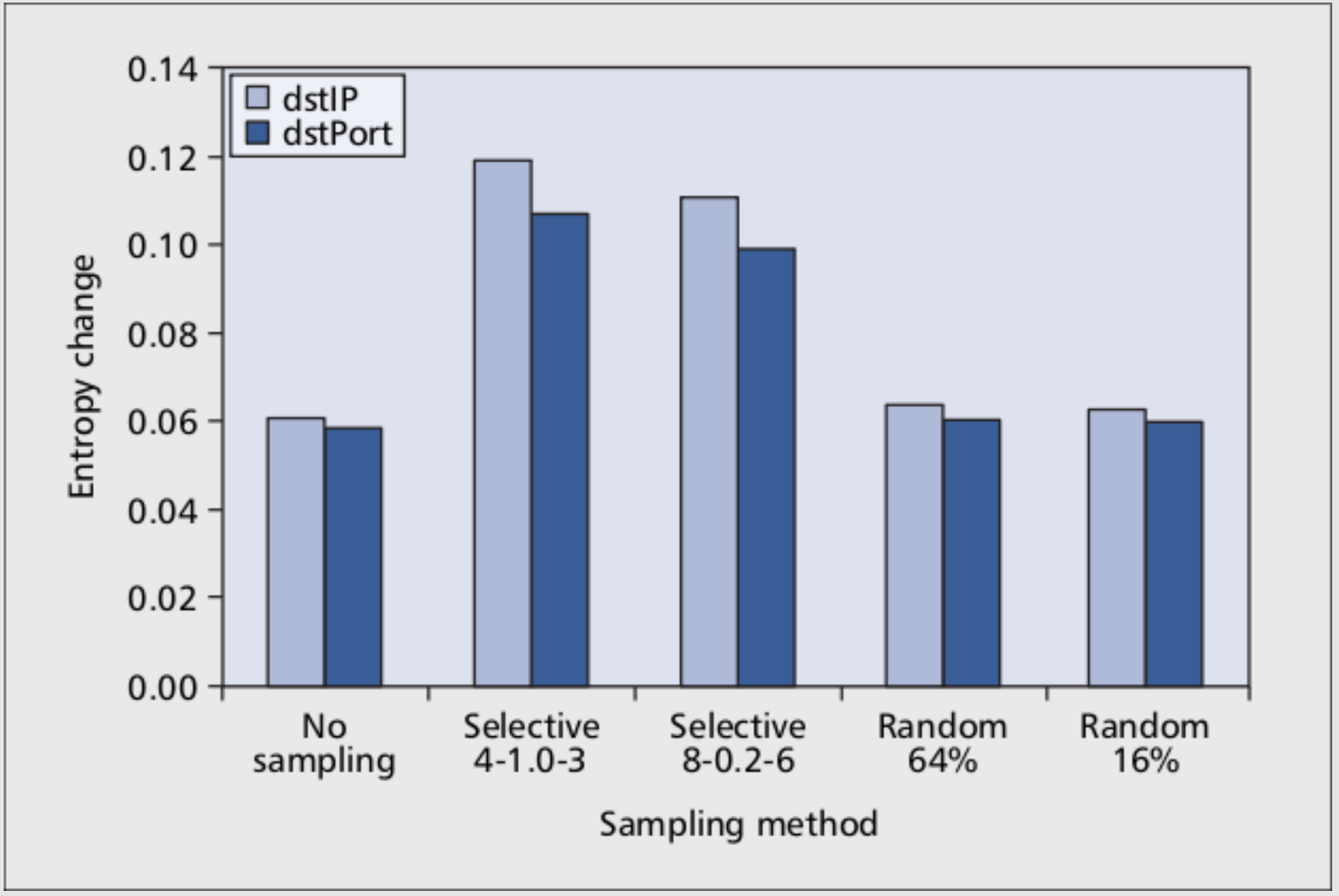
Framework

- Získána data z linky spojující NTUA a GRNET
- Celkem cca 4000 PC
- 250 MBit/s
- Pro experimenty byly do těchto dat zanášeny studované útoky
- Vyhodnocení probíhalo oproti nevzorkovaným flow datům a náhodně vzorkovaným flow datům

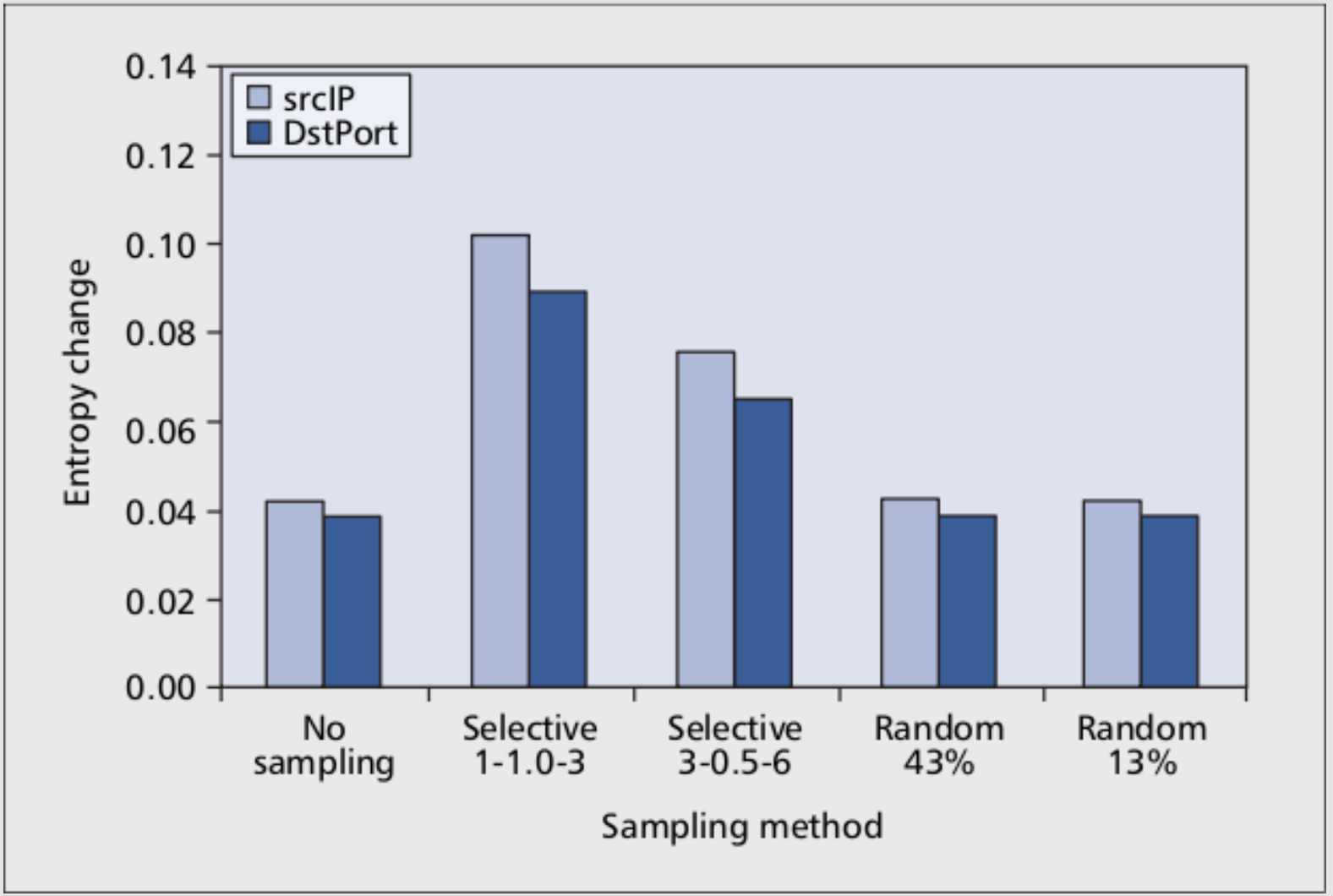
Entropie cílové IP adresy a portu



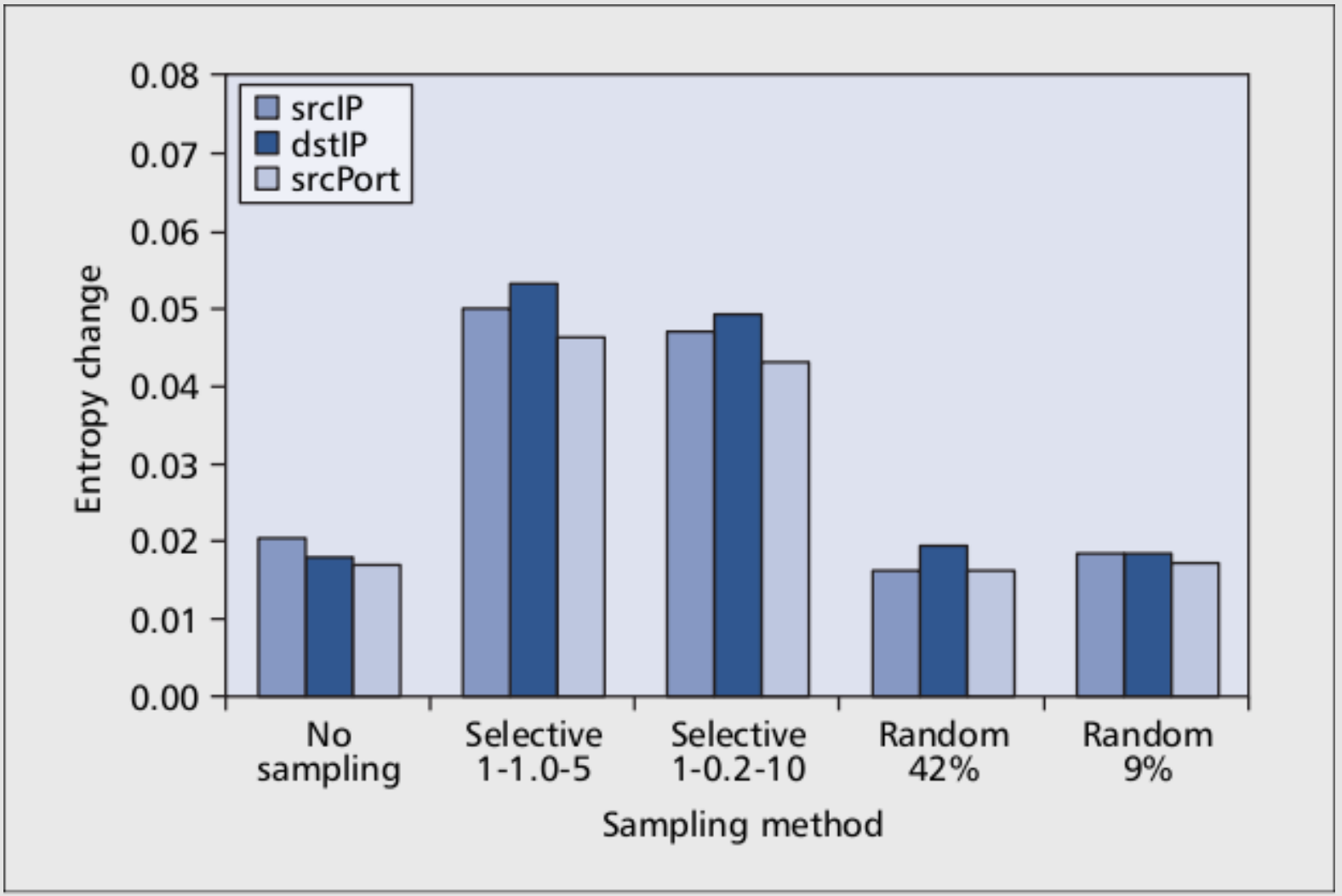
Porovnání


dstIP
dstPort
Entropy change
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
No sampling
Selective 4-1.0-3
Selective 8-0.2-6
Random 64%
Random 16%
Sampling method

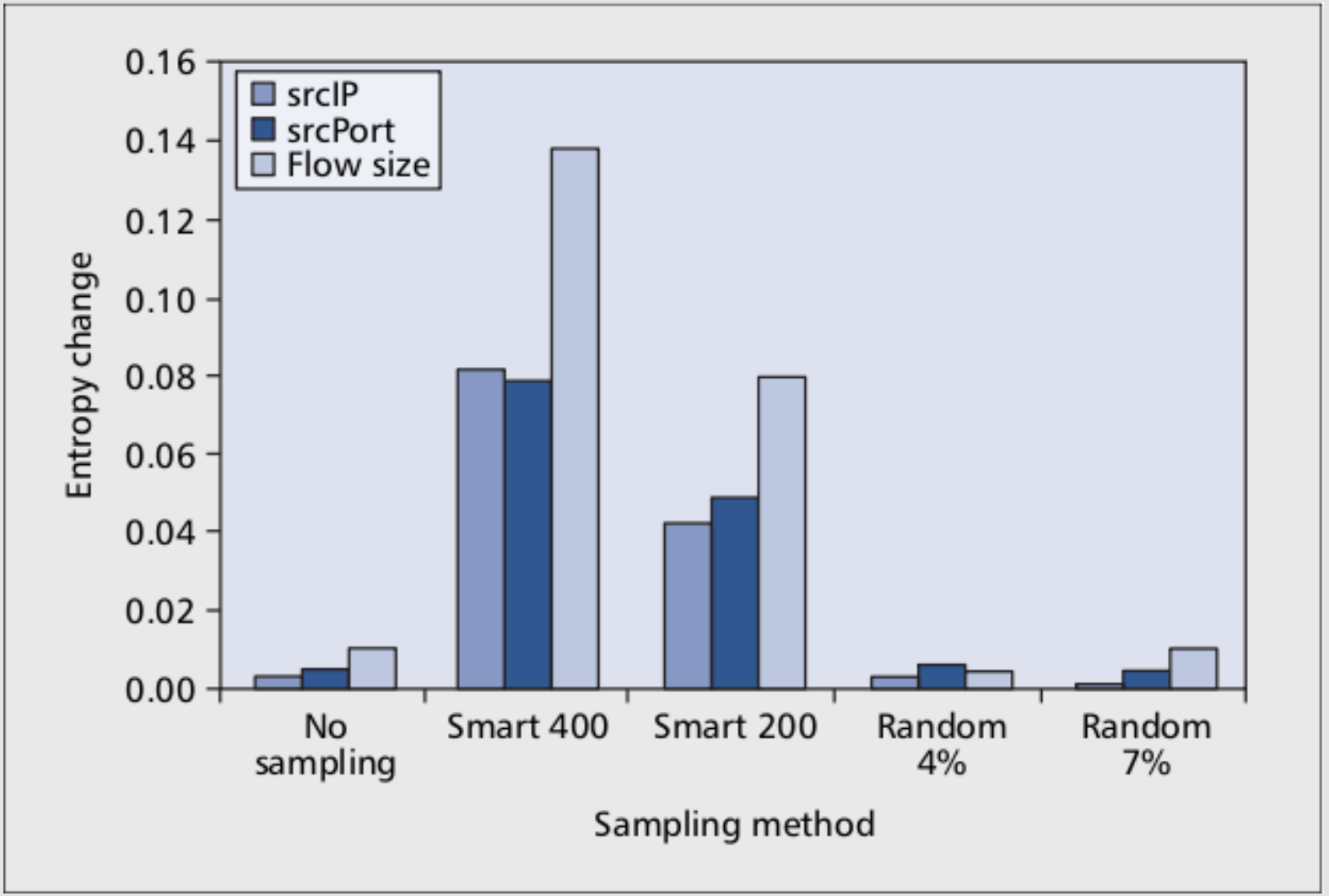
Porovnání


srcIP
DstPort
Entropy change
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
No sampling
Selective 1-1.0-3
Selective 3-0.5-6
Random 43%
Random 13%
Sampling method

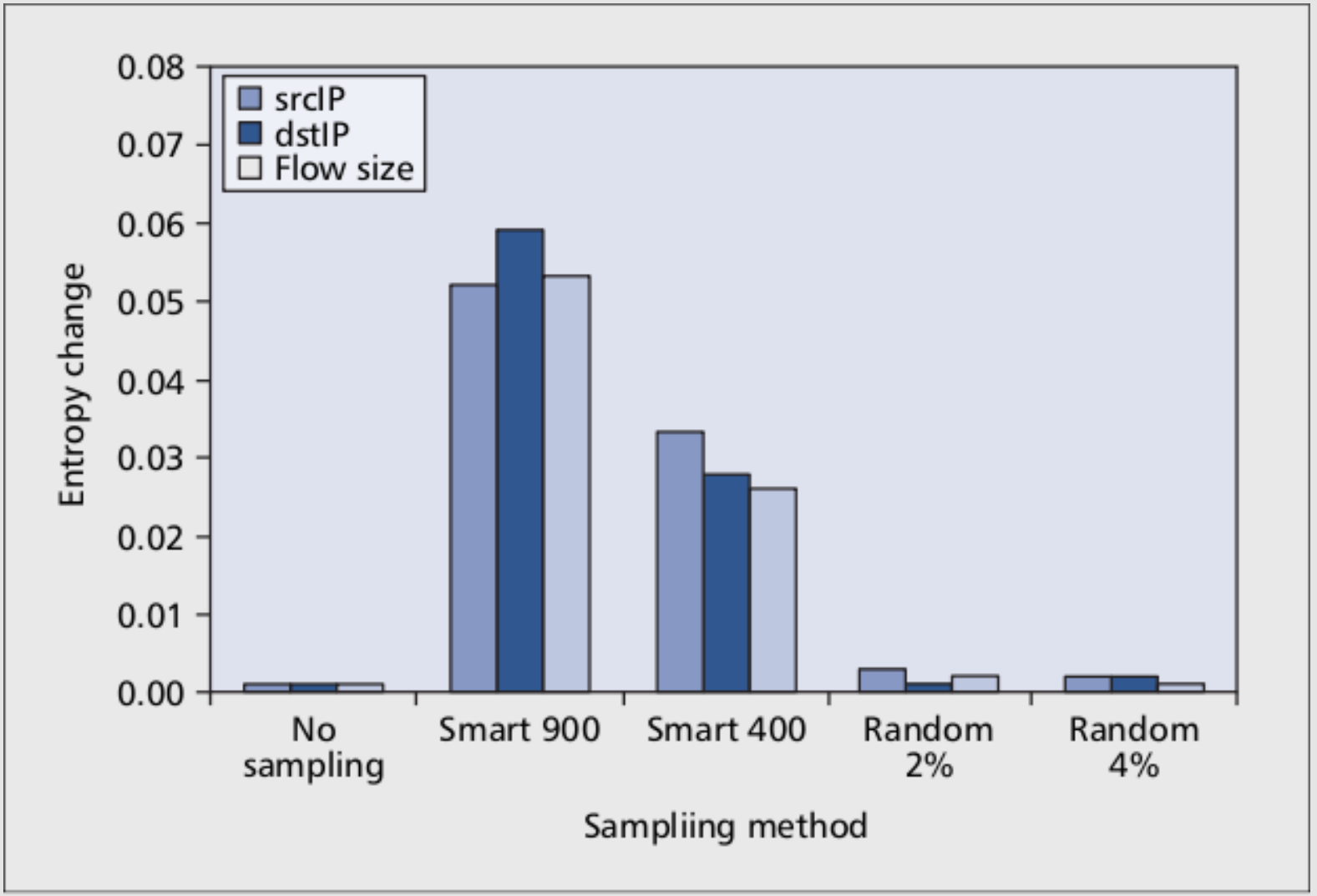
Porovnání


srcIP
dstIP
srcPort
Entropy change
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
No sampling
Selective 1-1.0-5
Selective 1-0.2-10
Random 42%
Random 9%
Sampling method

Porovnání


srcIP
srcPort
Flow size
Entropy change
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
No sampling
Smart 400
Smart 200
Random 4%
Random 7%
Sampling method

Porovnání


srcIP
dstIP
Flow size
Entropy change
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
No sampling
Smart 900
Smart 400
Random 2%
Random 4%
Sampliing method

Výzvy

- Flow sampling na 1G, 10G
- Dvoustupňové vzorkování pro snížení redukce
- IPv6

A nyní diskuze!