Classifiers Unclassified: An Efficient Approach to Revealing IP Traffic Classification Rules

Fangfan Li, Arash Molavi Kakhki, David Choffnes, Alan Mislove, Northeastern University, Phillipa Gill, Stony Brook University

Motivation

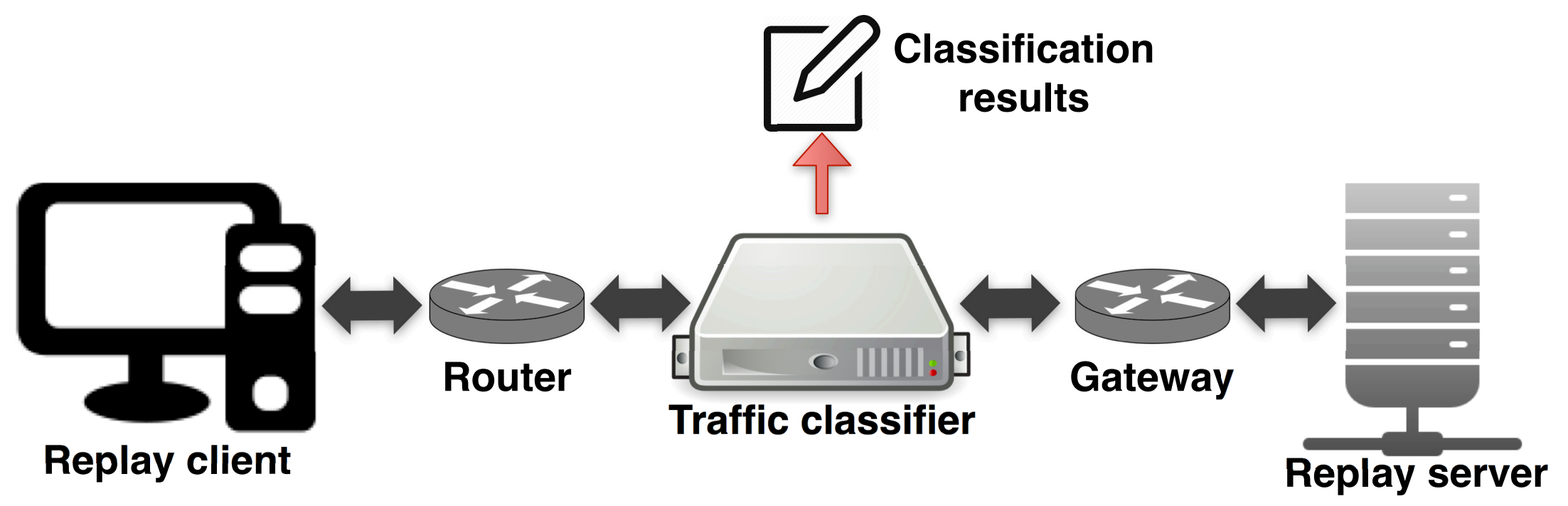
- Network providers use differentiation to enact network policies
- Such policies need a classifier to first assign Internet traffic to a category
- Little is known about implementations

Key questions

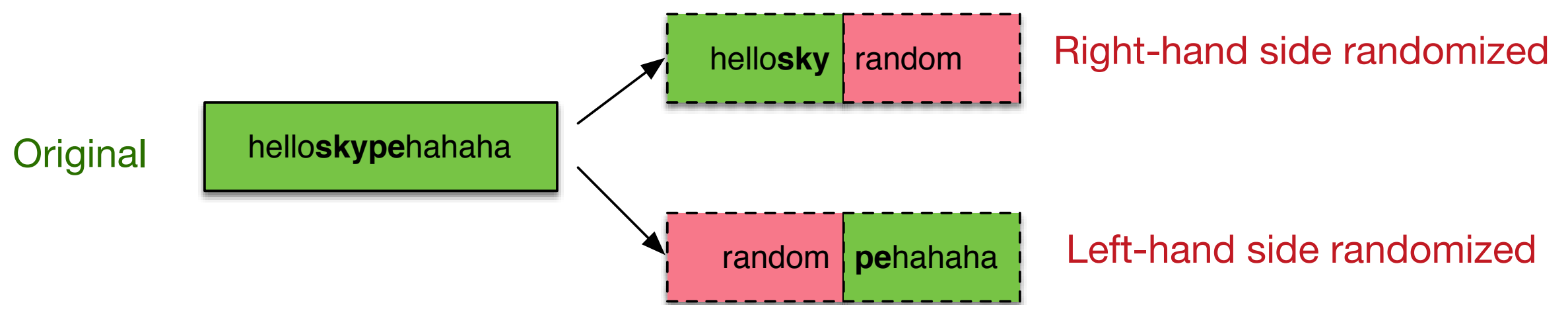
- How do classifiers detect applications?
- How do we extract classifier rules efficiently?

Methodology

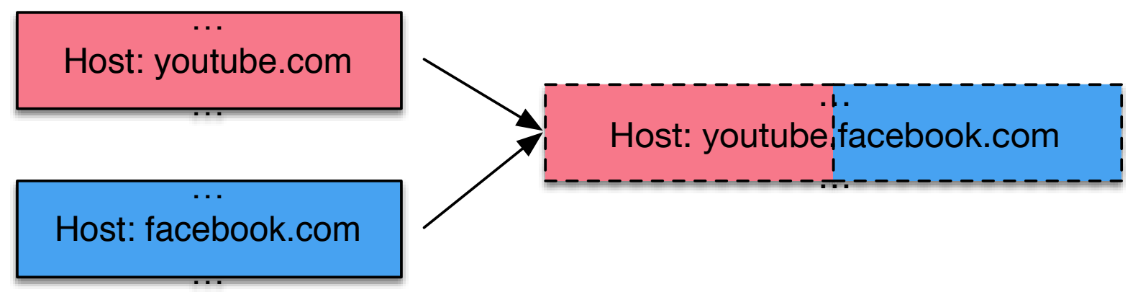
- Record and replay targeted applications



- Binary search for matching fields



- Construct *frankenflow* for precise rules



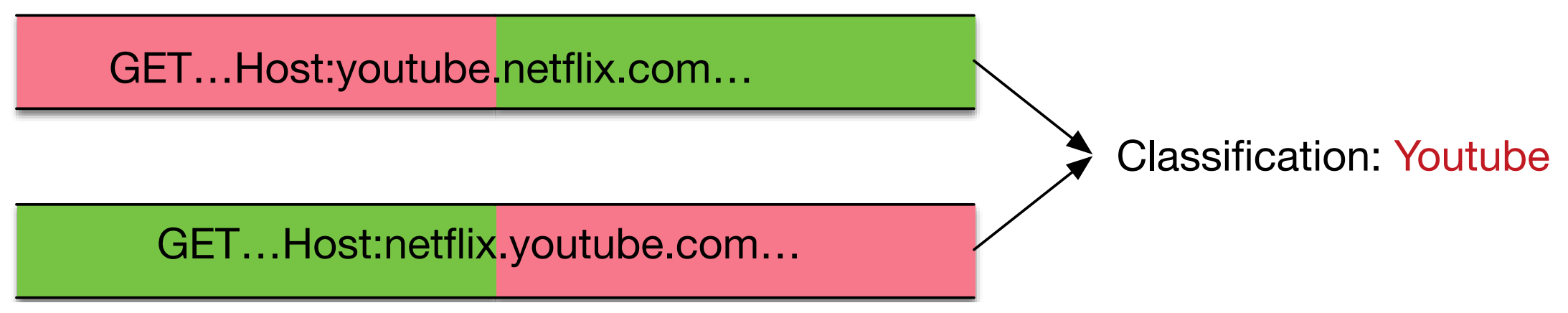
Key Findings

- ***Matching fields***
- First two packets in HTTP/S flows
- For HTTP traffic, the classifiers generally focus on URI, Host field, User Agent field and Content Type field
- For HTTPS traffic, the classifiers match on fields in TLS handshake such as SNI and Certificate.

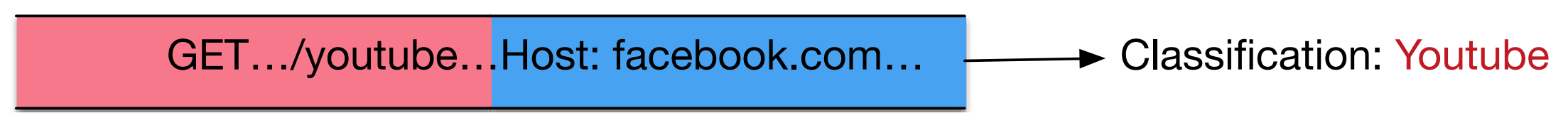
- ***Precise matching rules***

Header	Example Value	Application
URI	site.js?h={...}-**nbcsports**-com	NBC Sports
Host	Host: www.**netflix**.com	Netflix
User-Agent	User-Agent: **Pandora** 5.0.1 {...}	Pandora
Content-Type	Content-Type: video/**quicktime**	QuickTime

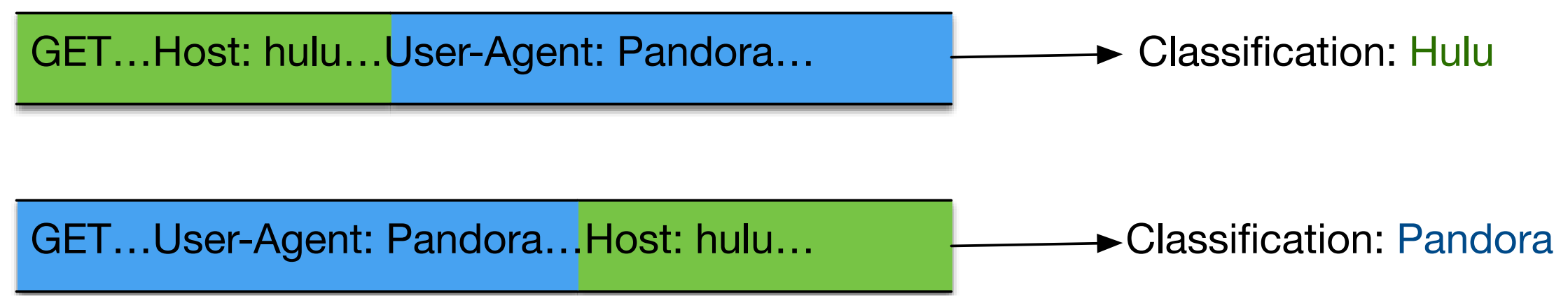
- ***Priority of different rules***
- Within the same field



- Across different fields



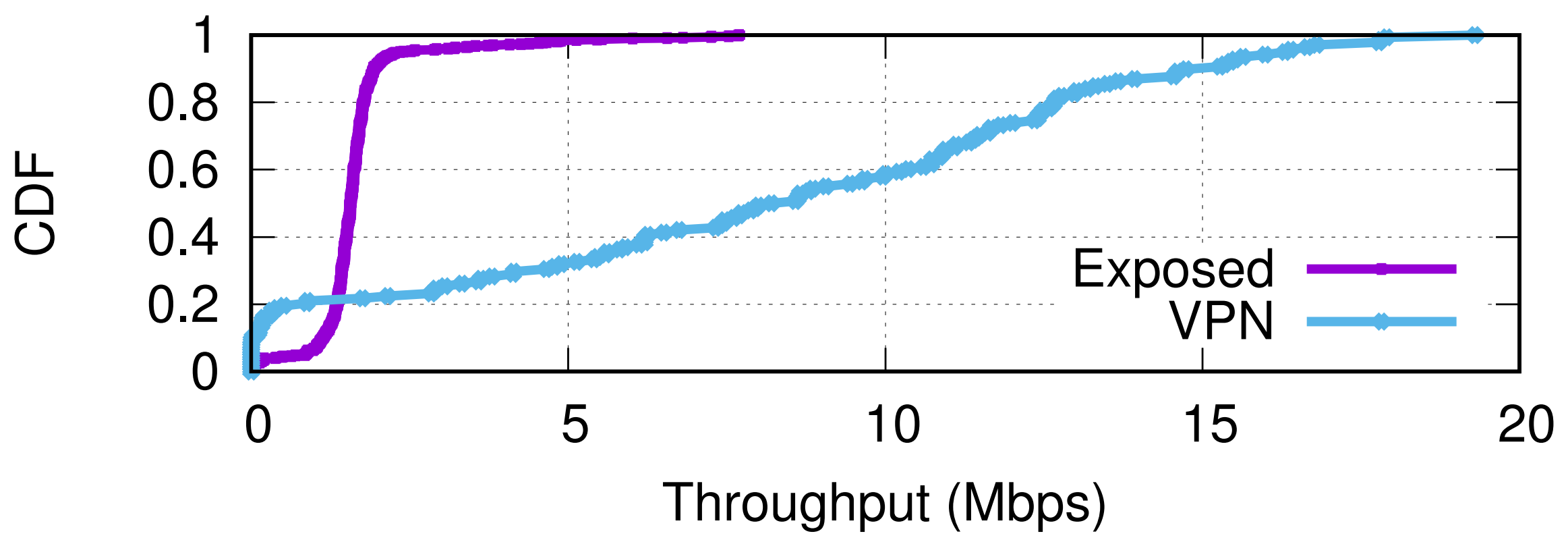
- The order of the fields appeared in the packet


GET…Host: hulu…User-Agent: Pandora…
Classification: Hulu
GET…User-Agent: Pandora…Host: hulu…
Classification: Pandora

Case study: T-Mobile's Binge On

Free video, but throttled to 1.5Mbps

- **Performance**



- **Binge On Implementation**
- Uses Host, Content Type, SNI
  - Host: determines whether flow is zero-rated
  - Content Type: check whether flow is throttled
  - SNI: determines whether flow is zero-rated and throttled at the same time

Future work

- Traffic that are not HTTP/S
- Deployment outside of the US
- Mobile app to allow anyone to test
- Automated circumvention

Full paper to appear in IMC 2016

More info: http://dd.meddle.mobi/