# Themis: Accelerating the Detection of Route Origin Hijacking by Distinguishing Legitimate and Illegitimate MOAS

Lancheng Qin[1], Dan Li[1,3], Ruifeng Li[2], Kang Wang[1]

[1]Tsinghua University

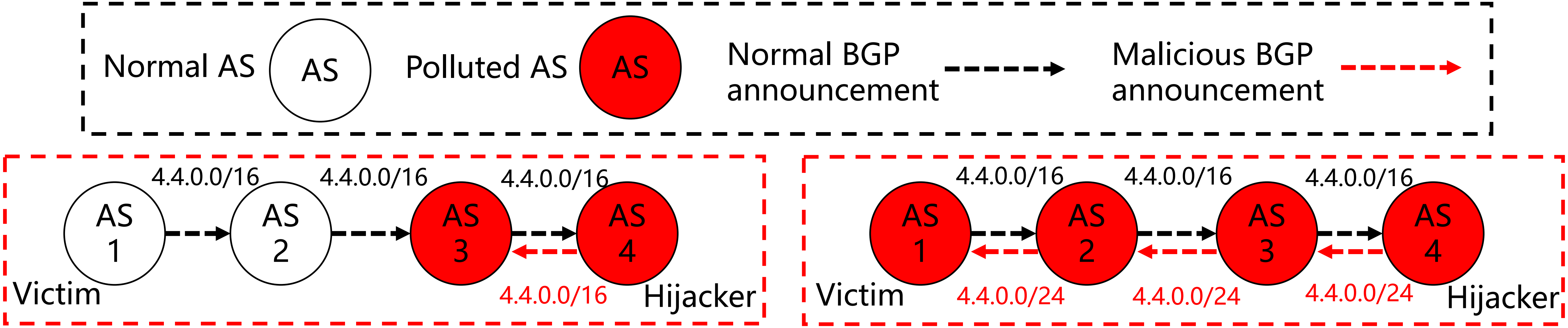
[2]Tsinghua Shenzhen International Graduate School

[3]Zhongguancun Laboratory

# Route Origin Hijacking

**Origin hijacking is the most commonly observed type of BGP hijacking**

- Exact-prefix origin hijacking
  - Hijacker announces the same prefix of victim
- Sub-prefix origin hijacking
  - Hijacker announces a sub-prefix of victim



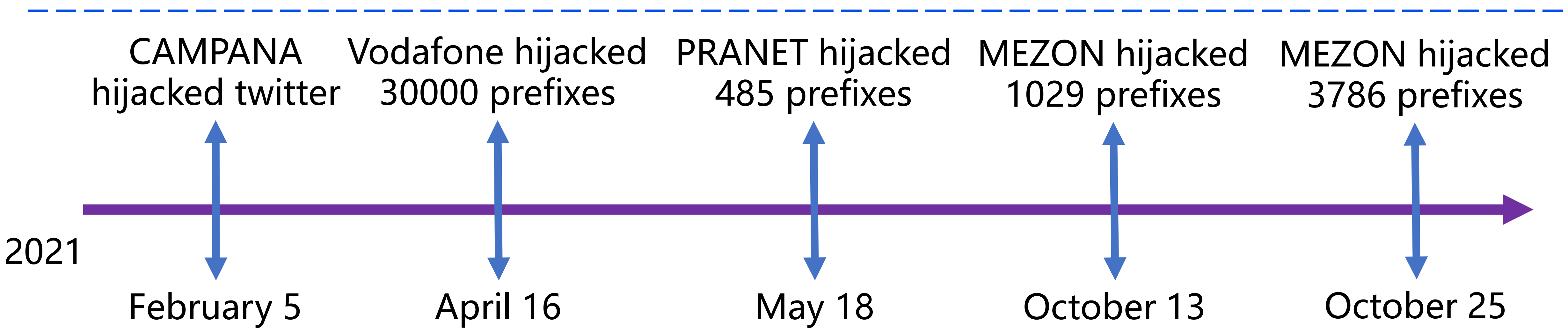
Exact-prefix origin hijacking

Sub-prefix origin hijacking

# Security Threat

□ Origin hijacking usually does not last long, but can cause serious routing and security problems, such as traffic disruption and financial losses

□ In 2021, BGPMON detected 775 major BGP route hijacking incidents, resulting in large-scale network outages



CAMPANA hijacked twitter

Vodafone hijacked 30000 prefixes

PRANET hijacked 485 prefixes

MEZON hijacked 1029 prefixes

MEZON hijacked 3786 prefixes

2021

February 5

April 16

May 18

October 13

October 25

# Existing Defense Mechanisms

- Proactive prevention mechanisms (e.g., RPKI)
  - Use cryptography to authorize all legitimate origin ASes for the prefix in advance to prevent origin hijacking
  - However, these mechanisms are fully effective only when deployed by all ASes, which is a long way to go
- Reactive detection mechanisms (e.g., Argus)
  - Monitor BGP updates from multiple BGP monitors worldwide and raise alarms when detecting origin hijackings

# Reactive Detection Mechanism

- Recent proposals (e.g., Argus) combine both control-plane analysis and data-plane probing to make the detection
  - Control plane: detecting all multiple origin ASes (MOAS) conflicts based on

    address block

  - Data plane: Using traceroutes/pings (and even manual verification) for each MOAS conflict to identify real origin hijackings

**However, most MOAS conflicts are legitimate MOAS**

# Legitimate MOAS

- Legitimate MOAS is a phenomenon in which a prefix is legitimately announced by multiple origin ASes simultaneously
  - business cooperation, IP address transfer/leasing or DDoS protection services
- Current BGP hijacking detection mechanisms cannot effectively distinguish legitimate MOAS from hijacking only by control-plane information

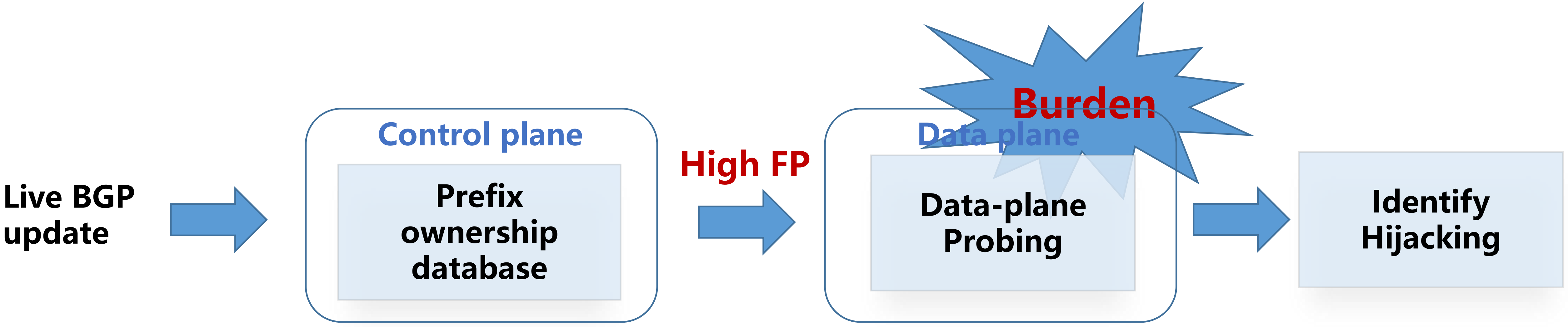
**Data-plane probing or even manual verification is still necessary**

et

the real-time requirements of detection

# High Verification Cost and Latency

- □ Legitimate MOAS happens much more frequently than origin hijacking
- □ Massive false positives from the control plan greatly increase the verification costs and latency

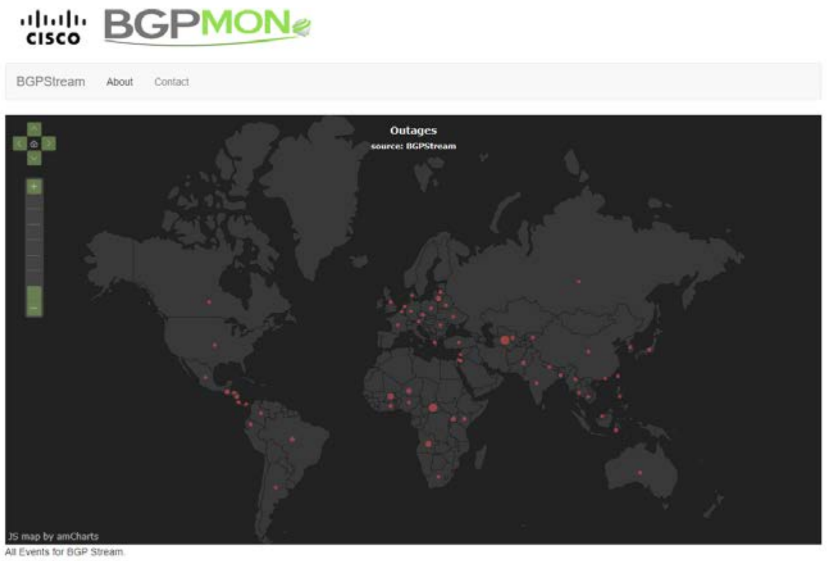


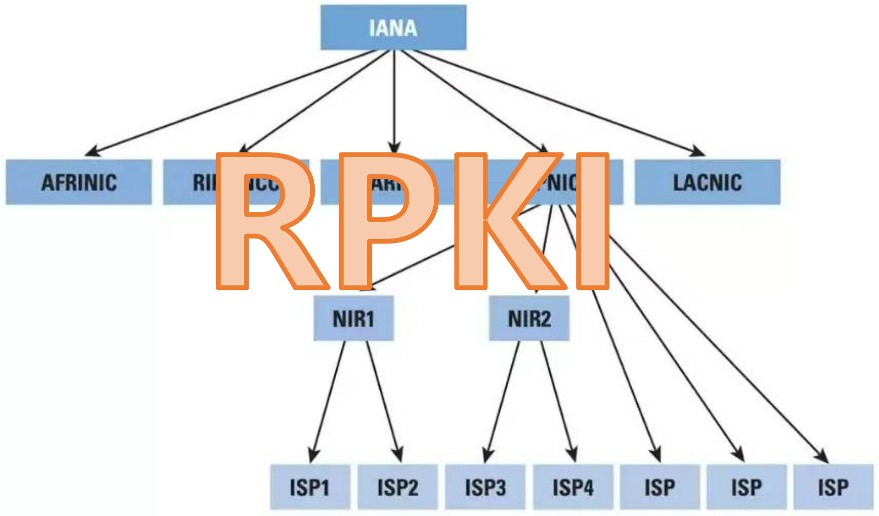
# Research Goals

- Accelerate the detection of route origin hijacking
  - Identify the potential causes of legitimate MOAS conflicts and the behavioral features of real hijackings
  - Train a machine learning classifier to distinguish legitimate MOAS and origin hijacking based on control-plane information
  - Propose a new origin hijacking detection system, i.e., Themis, to save verification cost and latency

# Ground Truth Dataset

- □ Origin hijacking
  - ◆ Collected from BGPmon and validated with ROV, IRR, topology, and MOAS duration
  - ◆ 867 exact-prefix, 476 sub-prefix
  - ◆ Confidence interval: [0.9811, 0.9999]
- □ Legitimate MOAS
  - ◆ Collected from RPKI and validated with BGP data
  - ◆ 499 exact-prefix, 1866 sub-prefix

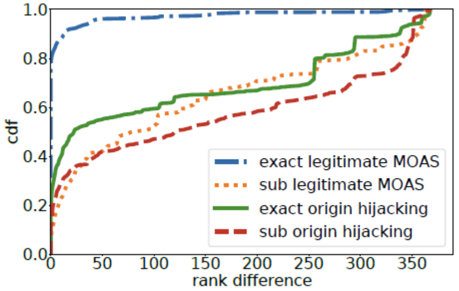


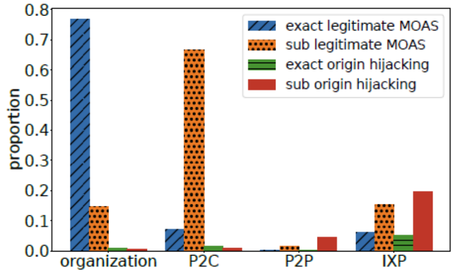


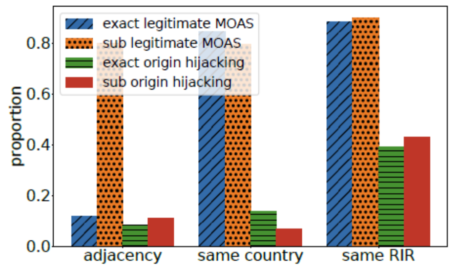
# MOAS Characteristics

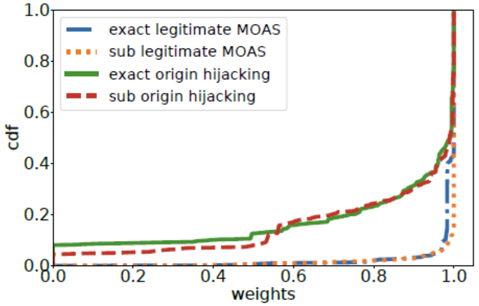
We identify six dominant characteristics:

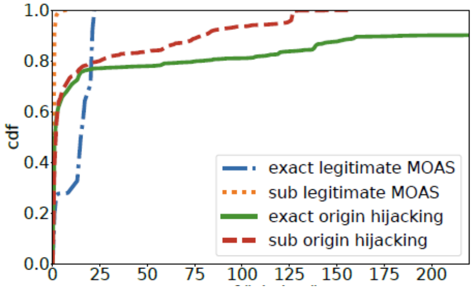
- □ Exact prefix or sub prefix
- □ Rank difference
- □ Business relationship
- □ Geographical relationship
- □ Announcement stability
- □ Hijacking activity

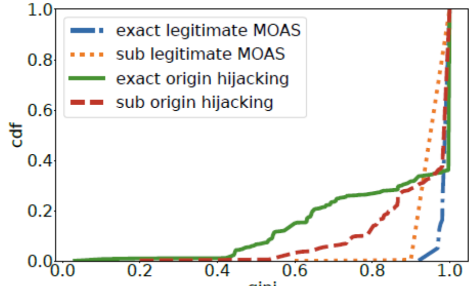


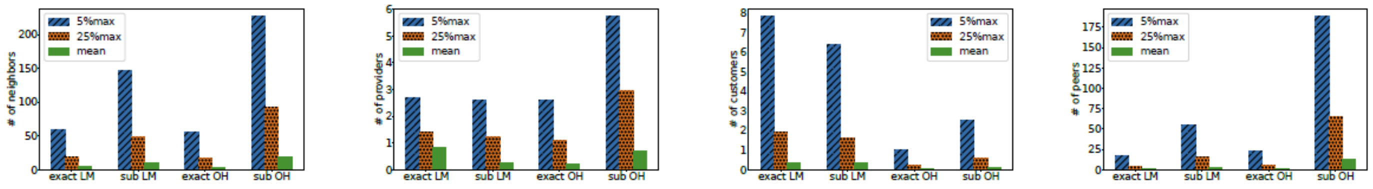












(a) number of common neighbors (b) number of common providers (c) number of common customers (d) number of common peers

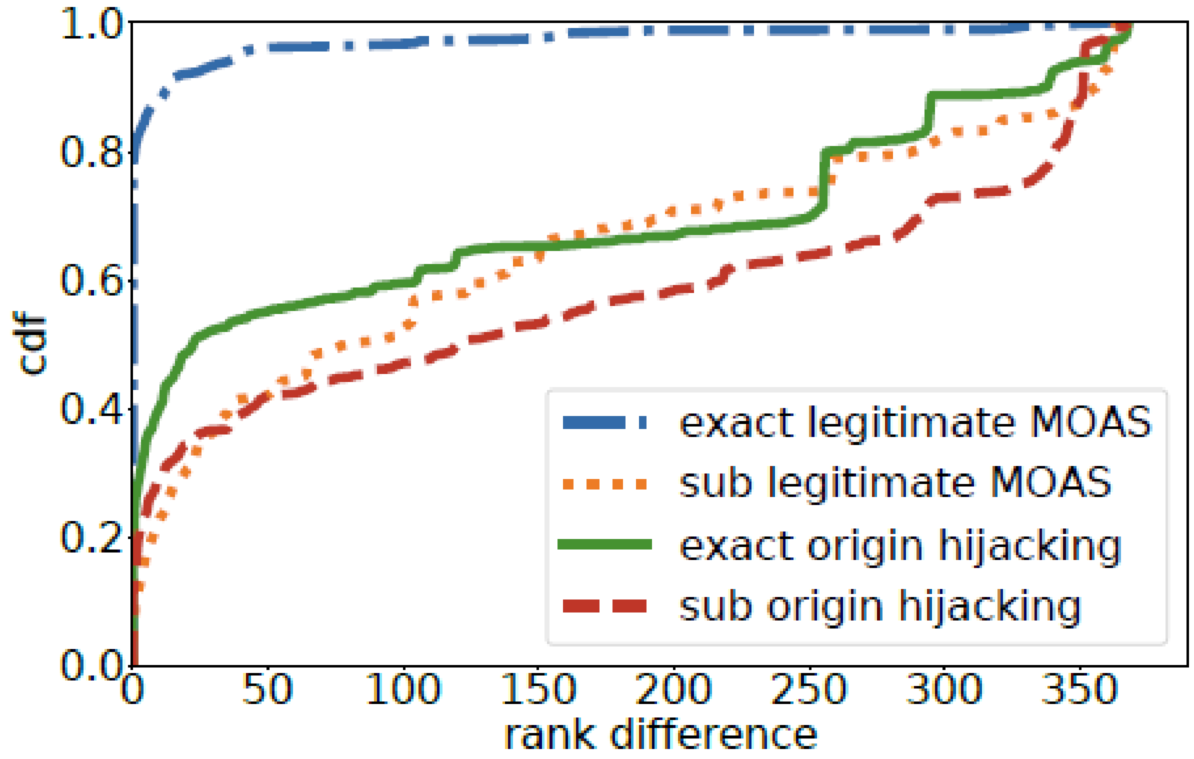
# Exact prefix or sub prefix

- Exact-prefix legitimate MOAS and sub-prefix legitimate MOAS are significantly different in some characteristics

- We intuitively assume that exact-prefix legitimate MOAS and sub-prefix legitimate MOAS are caused by different operations

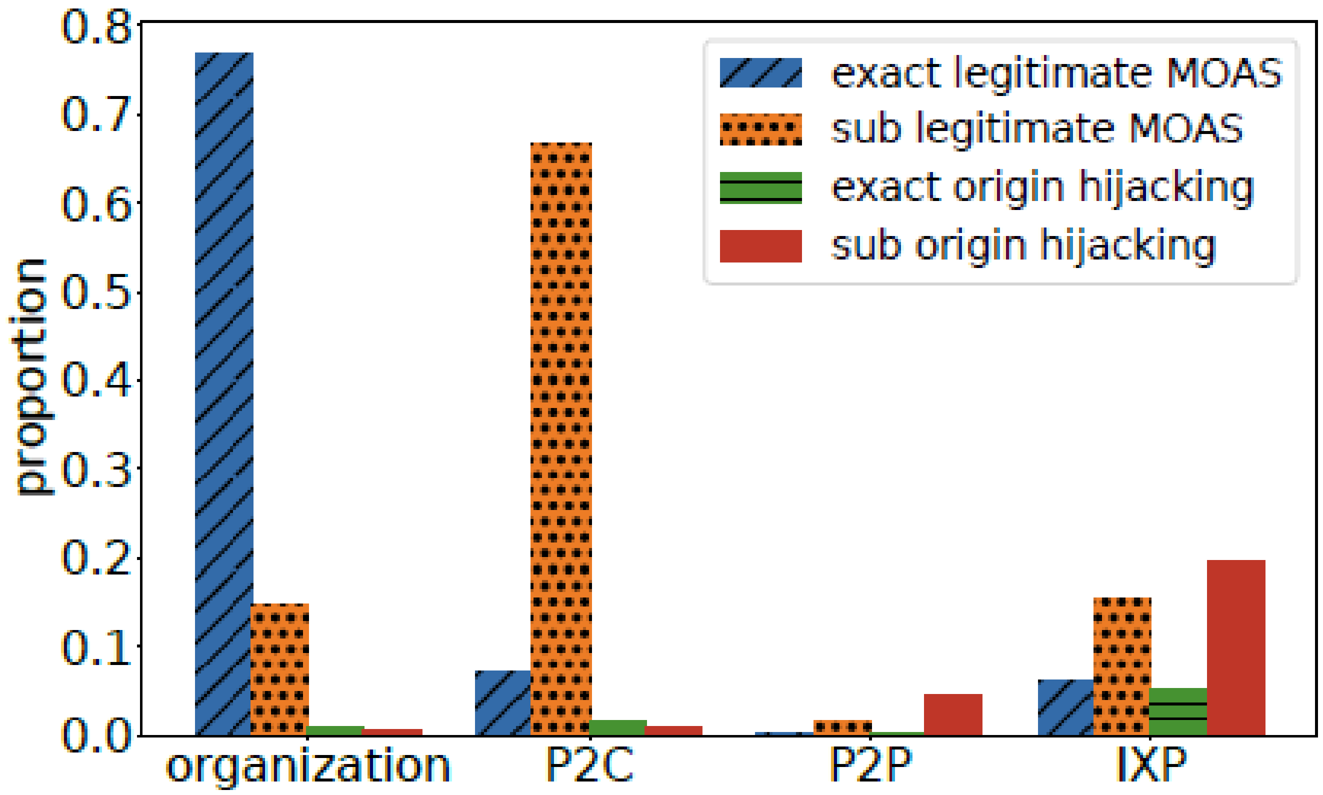
# MOAS Characteristics: Rank difference

- The overall rank difference of exact-prefix legitimate MOAS is much smaller than the other three MOAS types



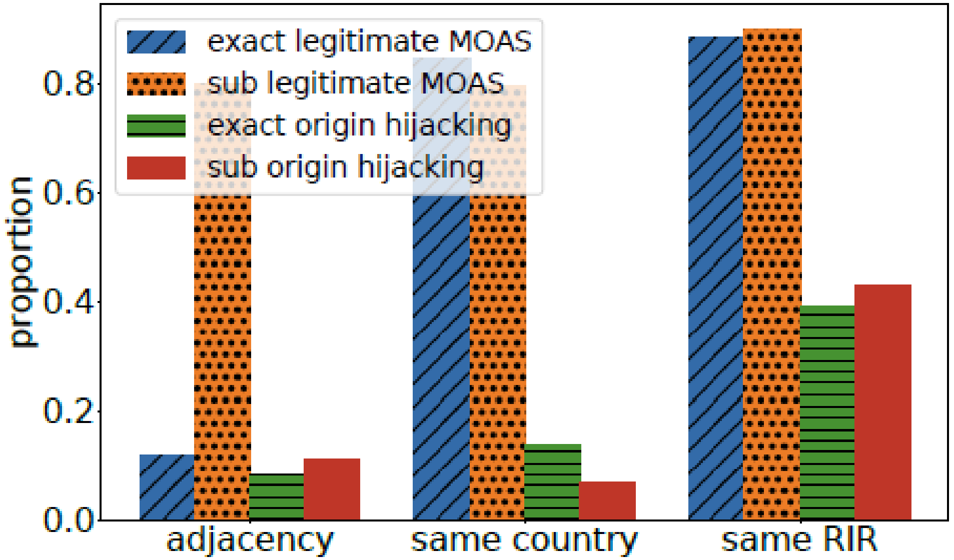
# Business relationship

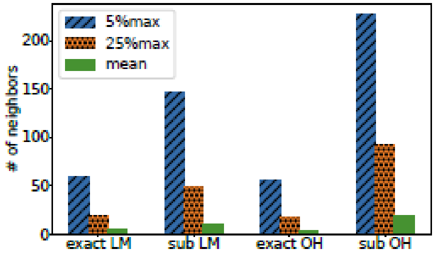
- We measure the proportion of different business relationships in individual MOAS types



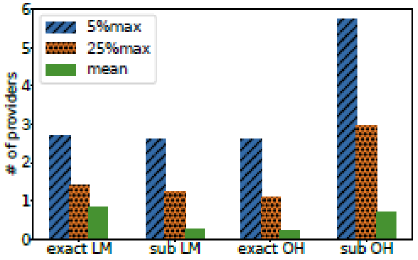
# Geographical relationship

- We measure the proportion of different geographical relationships in individual MOAS types

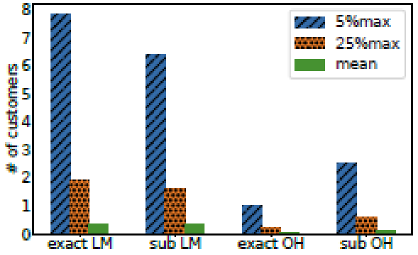




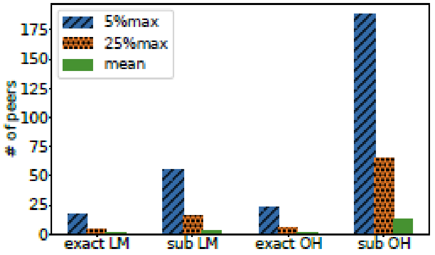
(a) number of common neighbors



(b) number of common providers



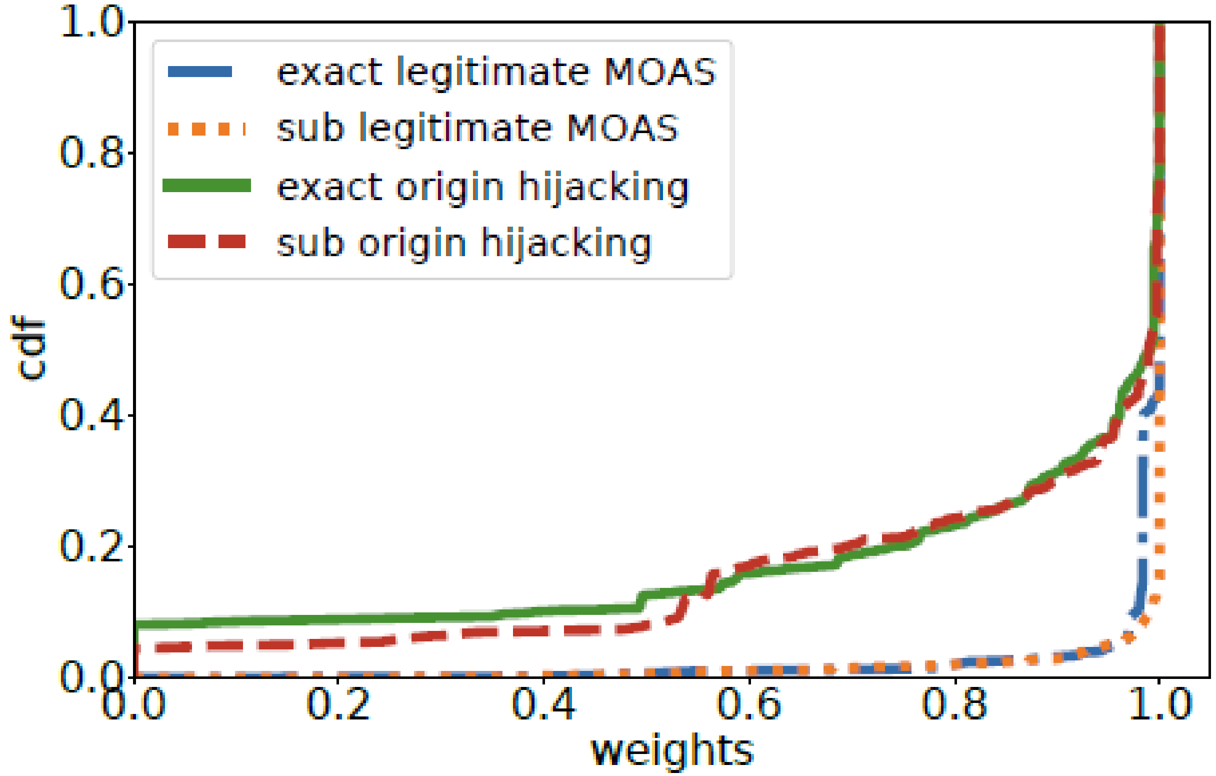
(c) number of common customers



(d) number of common peers

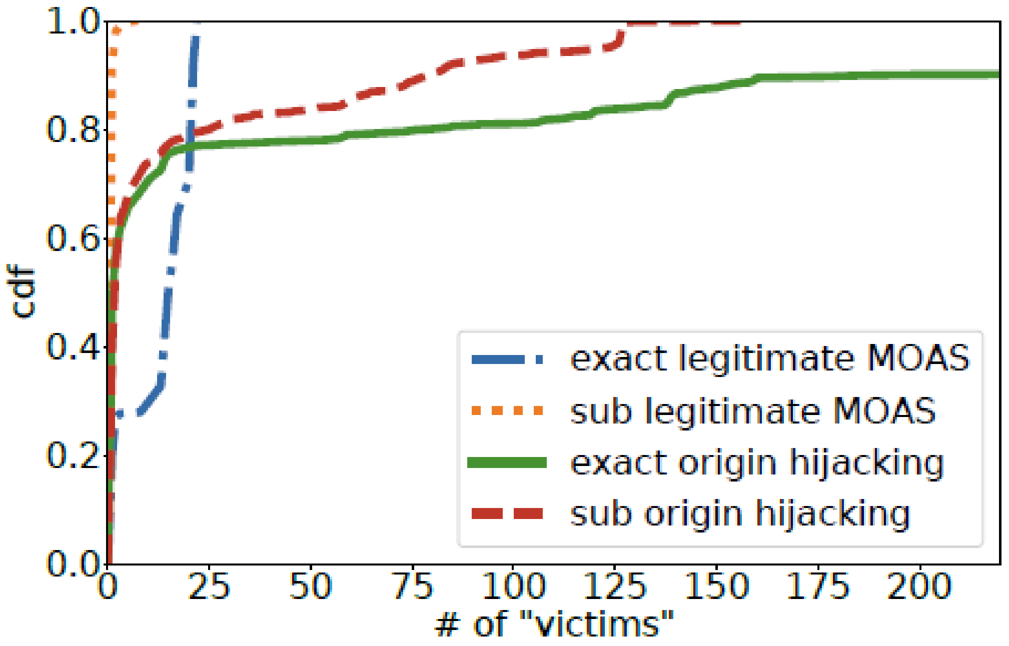
# Announcement stability

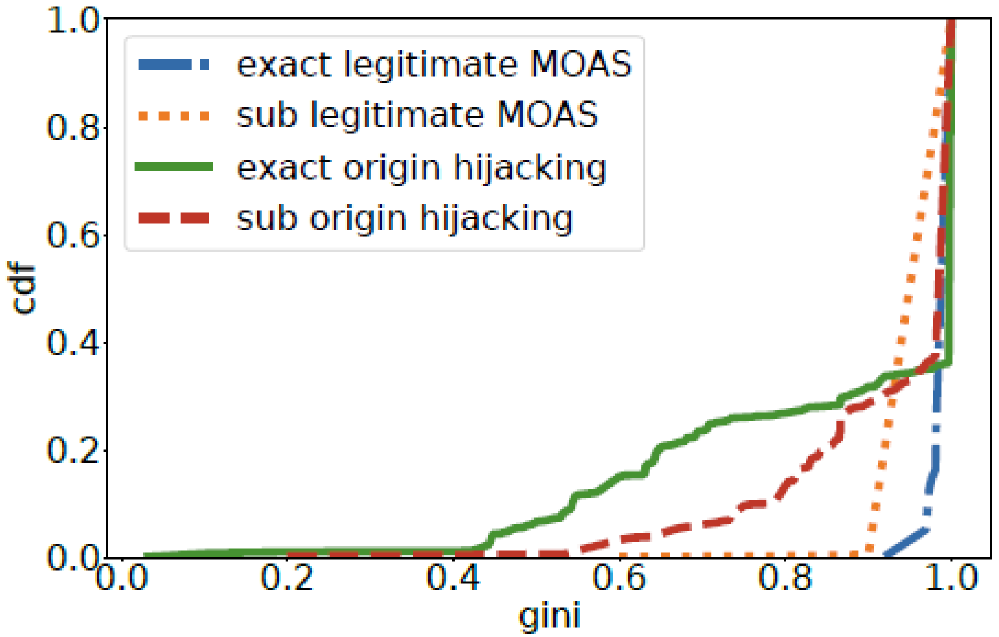
- Some hijackers do not appear in global routing tables for a long time before hijacking, while legitimate ASes usually announce prefixes in daily BGP updates



# Hijacking activity

- The hijacker in a large hijacking incident may hijack hundreds of victims (located in different RIRs) at the same time

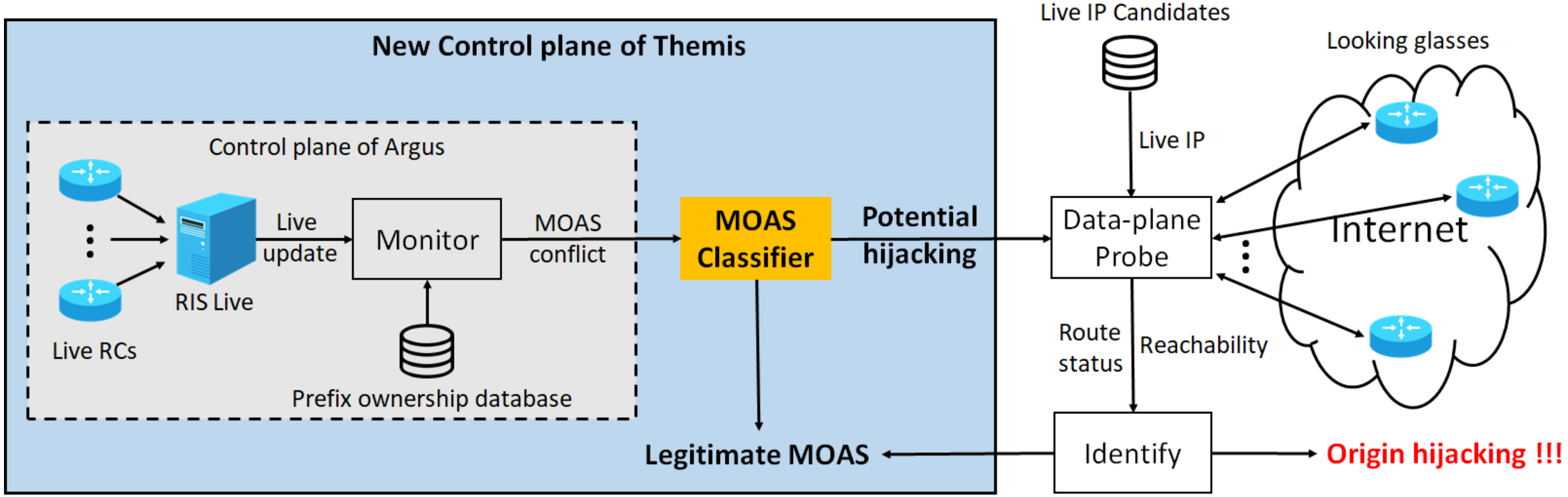




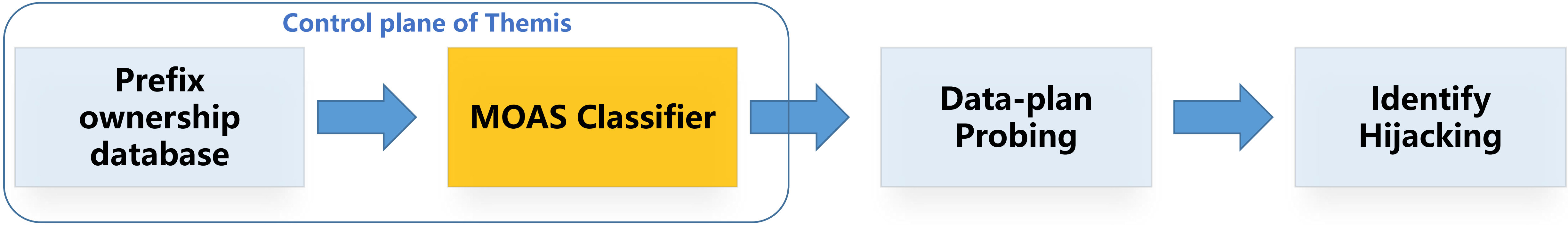
# MOAS Classifier

- □ Tree-based classifier
  - ◆ Use bootstrapping samples in the training phase of individual trees and compute the Out-Of-Bag (OOB) score
- □ 26 features that capture the six MOAS characteristics
  - ◆ All features add positive OOB accuracy
- □ 95.49% precision and 99.20% recall
  - ◆ Significantly reduce the false positives of existing control-plane analysis

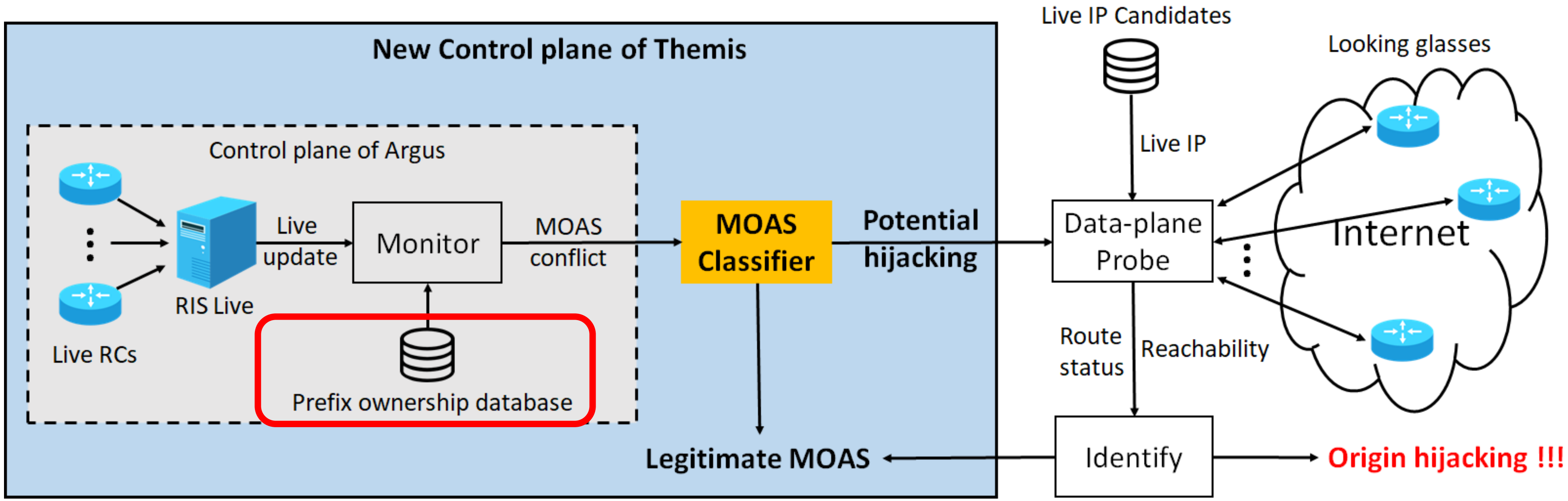
# Themis Architecture



## Themis workflow



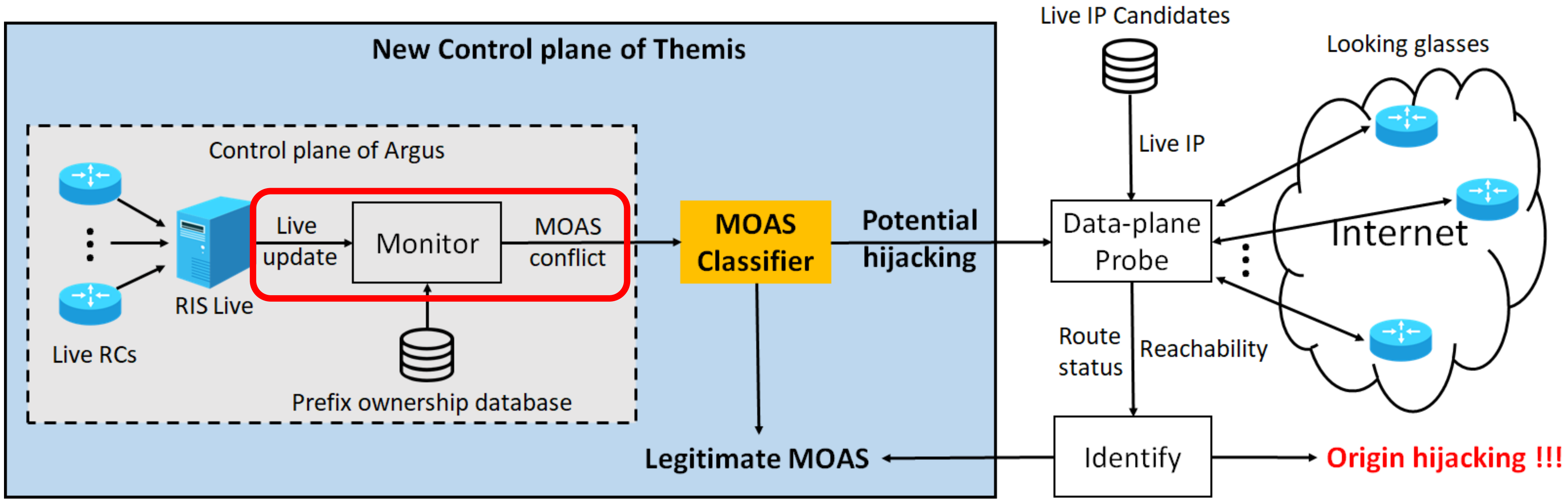
# Themis Architecture



## □ Building Prefix Ownership Database

- ◆ In addition to historical BGP data, Themis also collects more prefix ownership information from RPKI and IRR

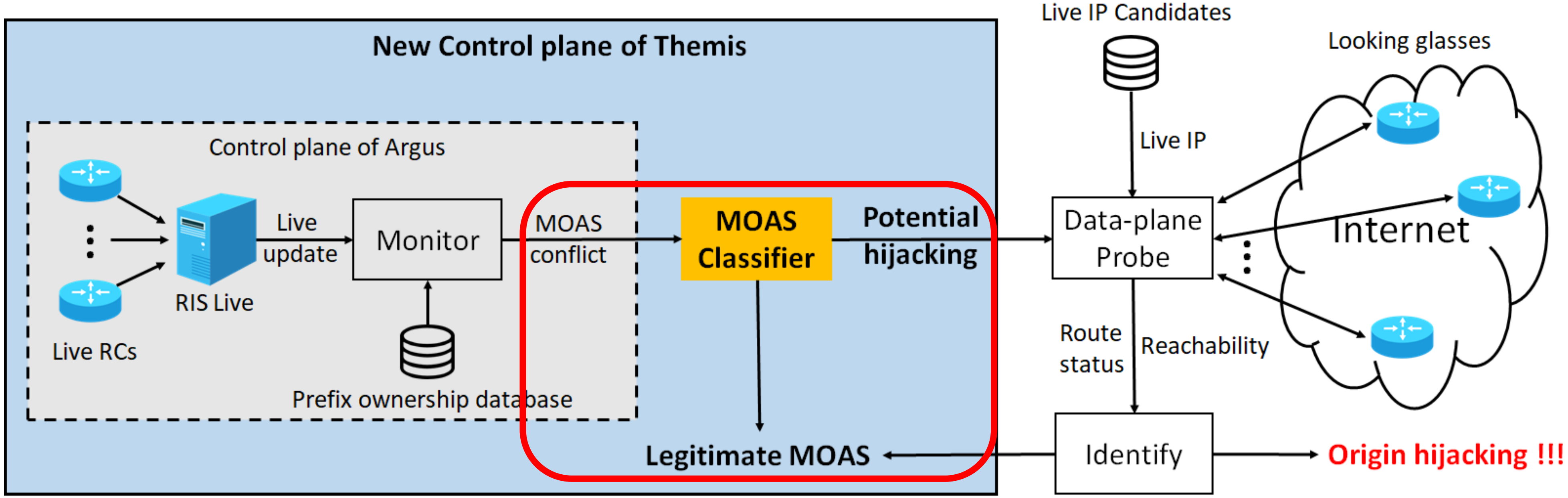
# Themis Architecture



□ Monitoring MOAS Conflicts

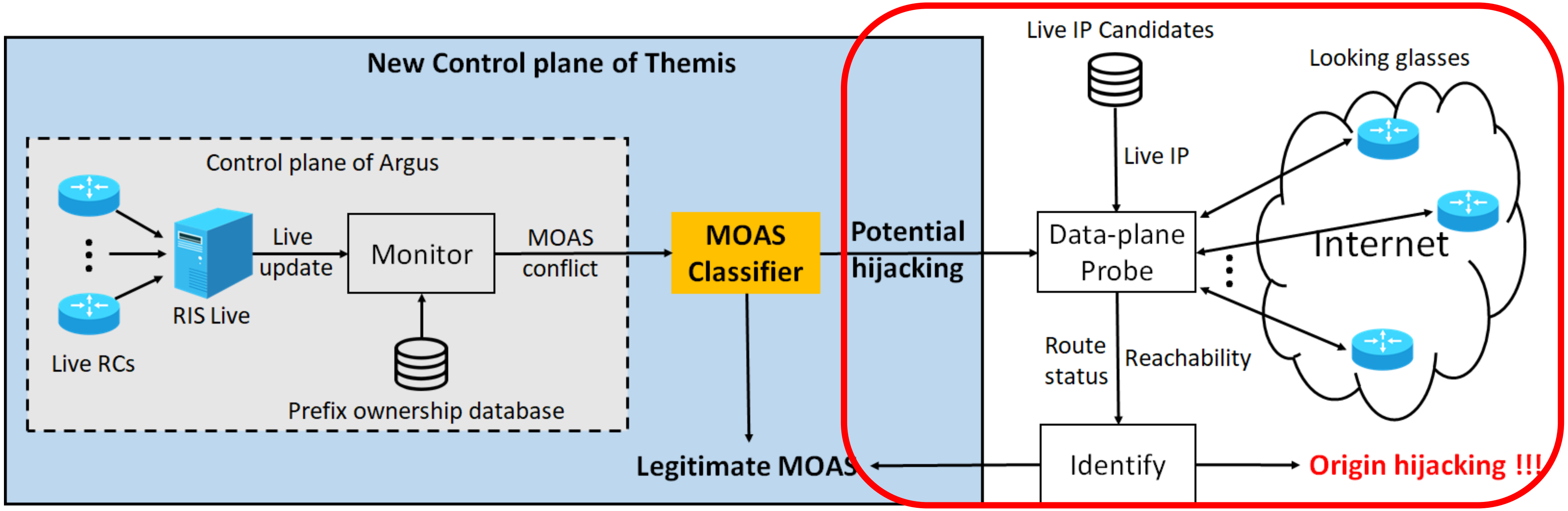
◆ Themis optimizes the monitor of Argus by combing historical BGP validation, ROV, and IRR validation to identify as many legitimate BGP updates as possible

# Themis Architecture



- □ Filtering Legitimate MOAS (which is the core of Themis)
  - ◆ Themis classifies MOAS conflicts into potential hijackings and legitimate MOAS conflicts by using MOAS classifier
  - ◆ Only potential hijackings will be sent to the data plane

# Themis Architecture

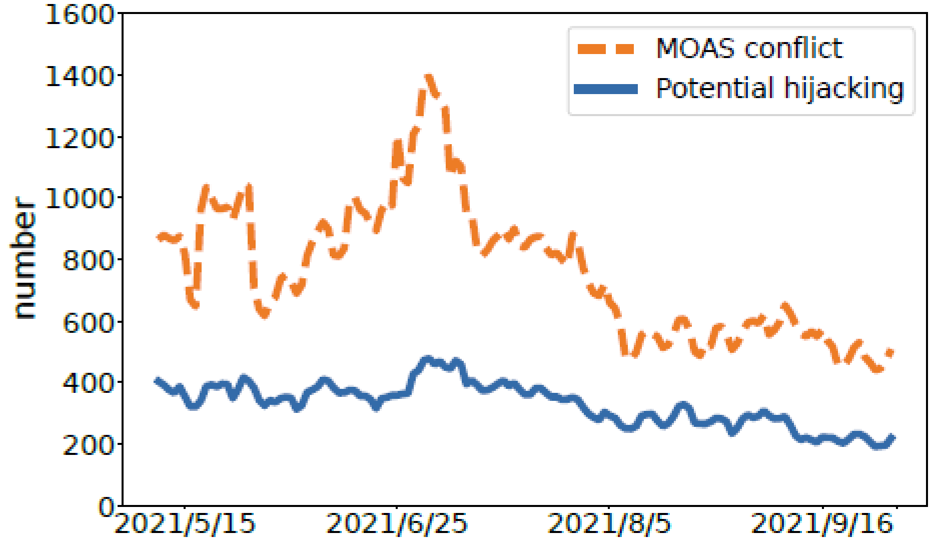


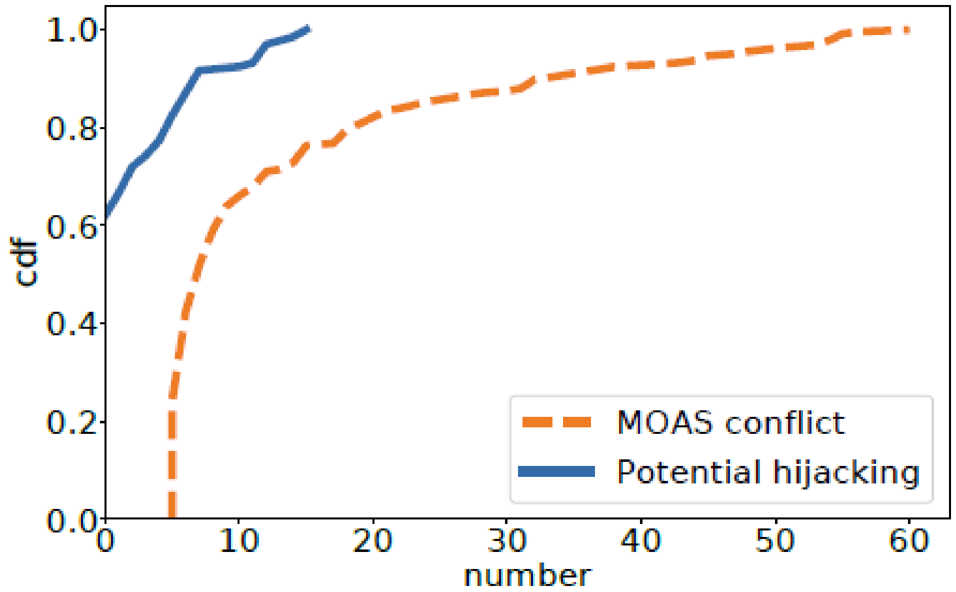
□ Probing and Identifying Origin Hijacking

◆ Themis uses the same data-plane probing and identification method as Argus to ensure detection accuracy

# Evaluation

- Themis achieves almost the same accuracy as Argus, and
  - Reduces 56.69% of verification costs than existing detection systems
  - Significantly accelerates the detection when many concurrent MOAS conflicts occur





# Thank you!

**Lancheng Qin**

**qlc19@mails.tsinghua.edu.cn**

**August 2022**