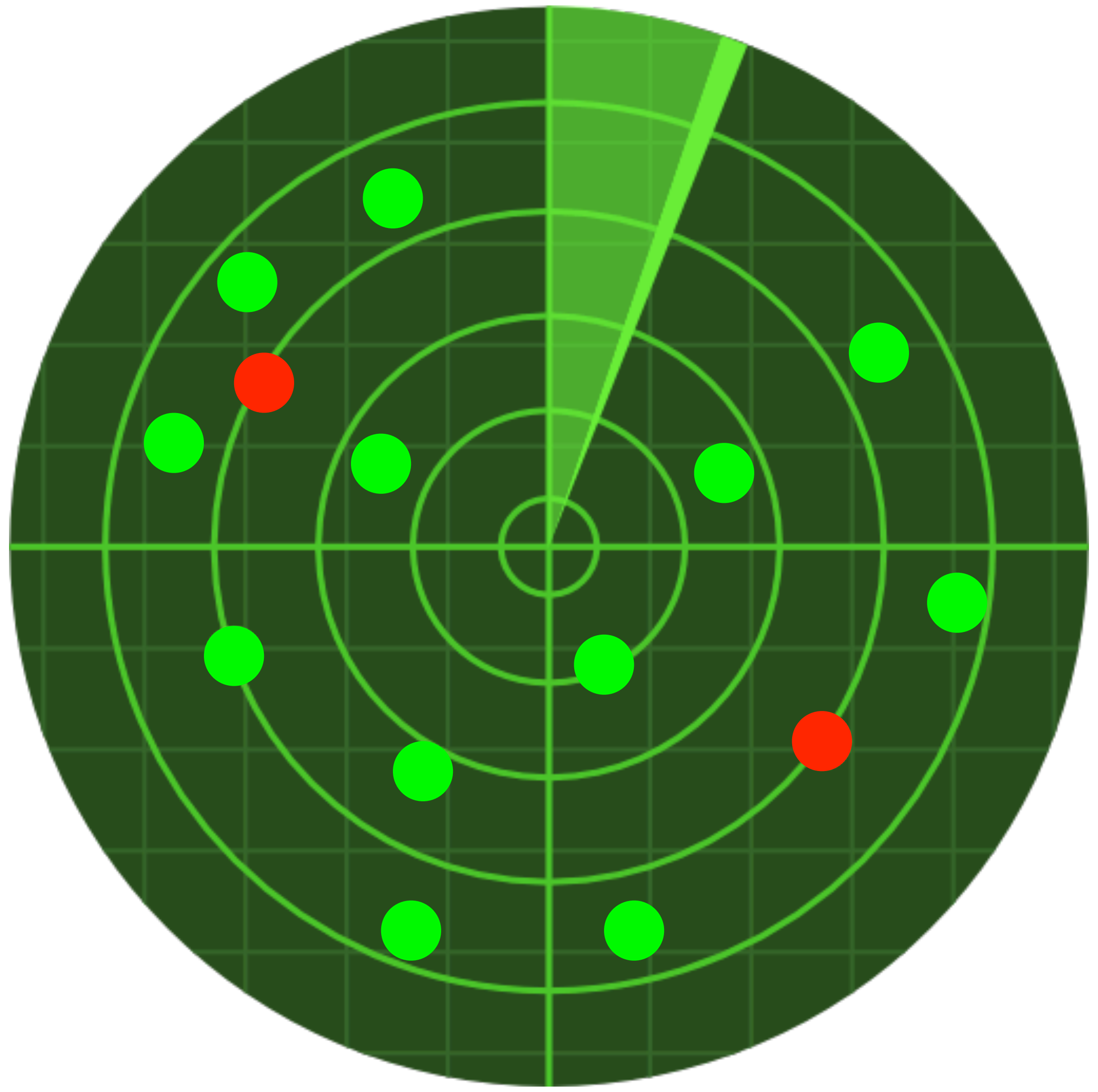
A System to Detect Forged-Origin BGP Hijacks

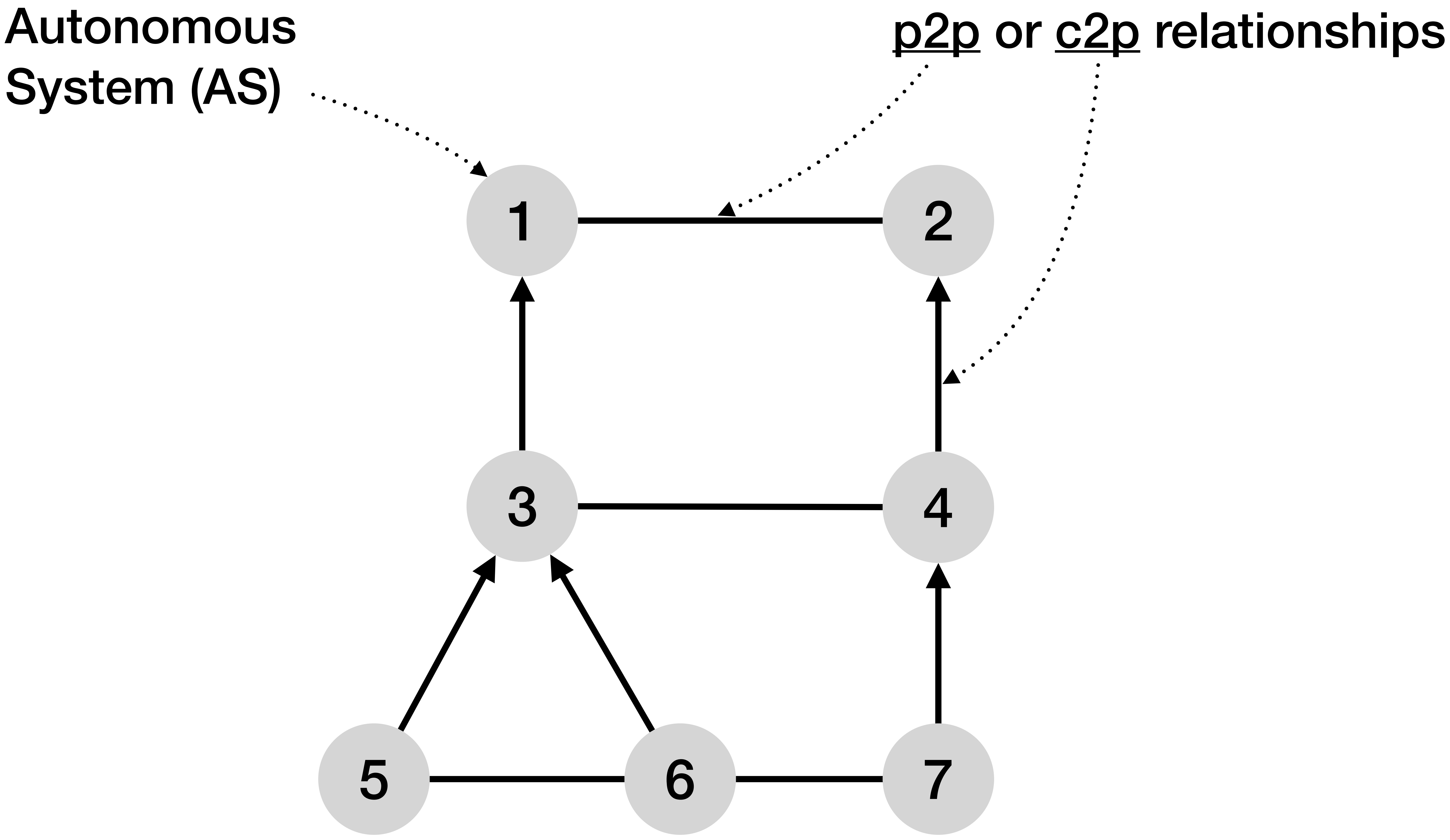
Thomas Holterbach
University of Strasbourg



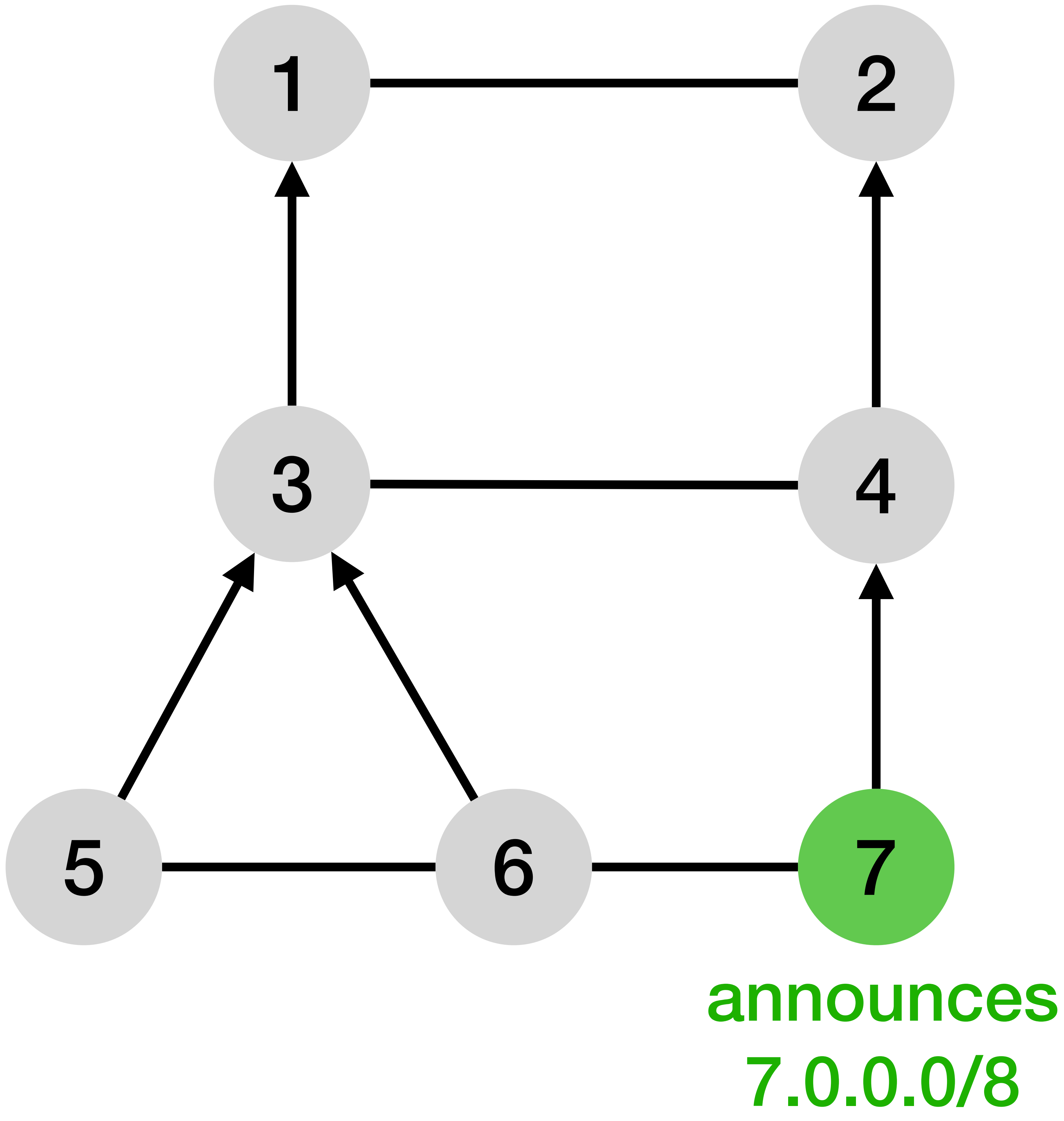
USENIX NSDI'24
Thursday, April 18

Joint work with:
Thomas Alfroy
Amreesh D. Phokeer
Alberto Dainotti
Cristel Pelsser

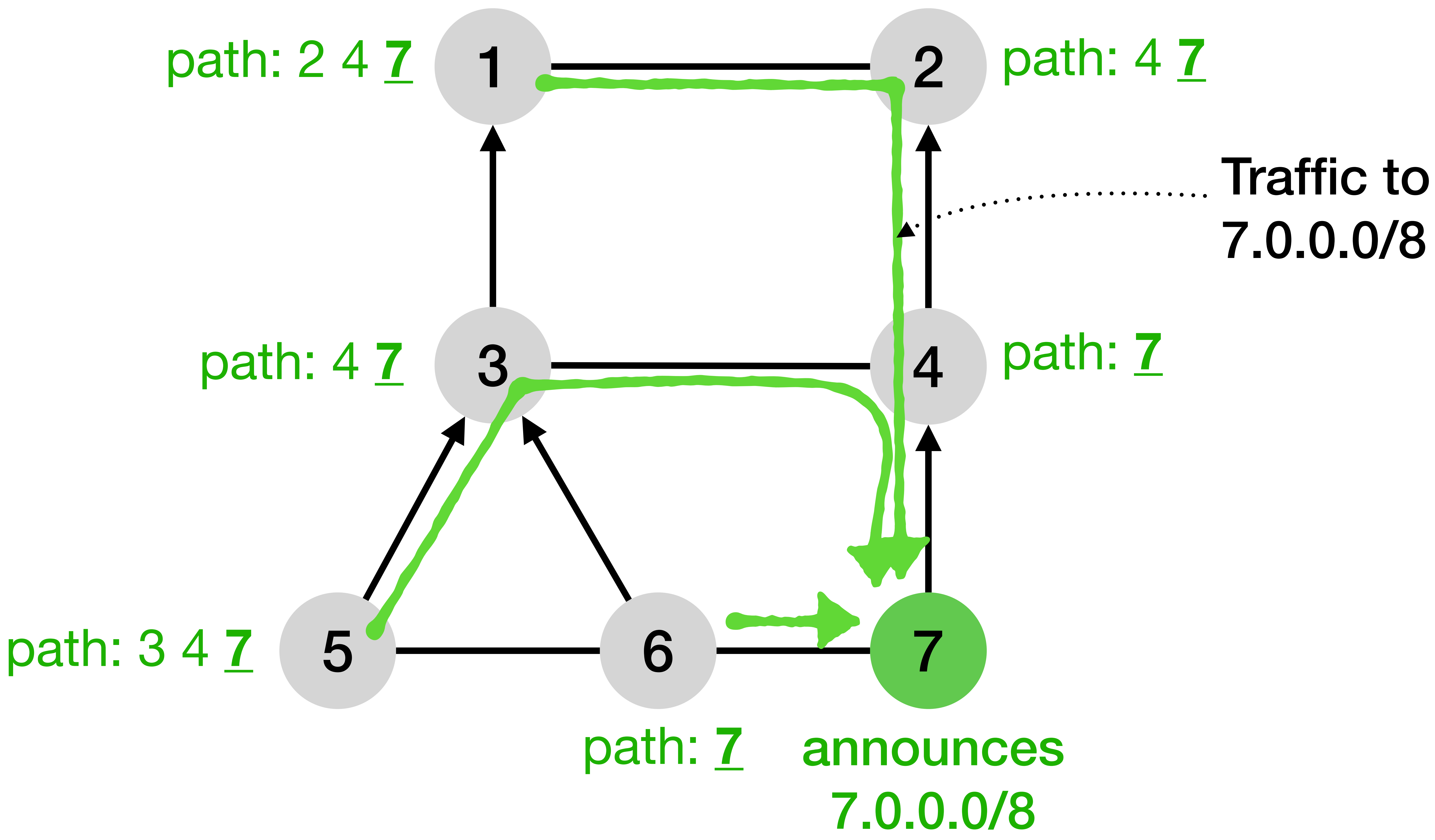
Internet routing (BGP) is vulnerable to traffic hijacking



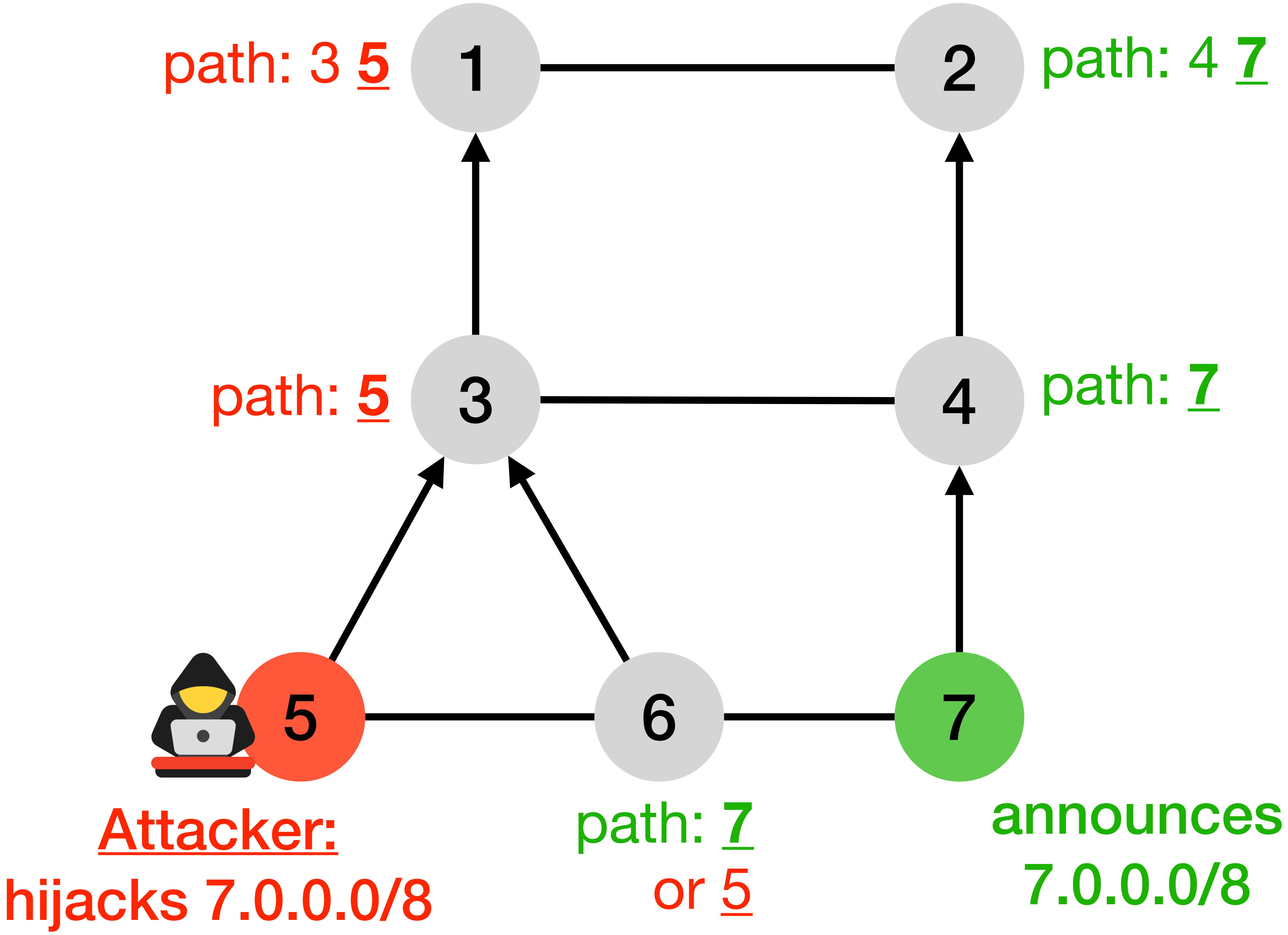
Internet routing (BGP) is vulnerable to traffic hijacking



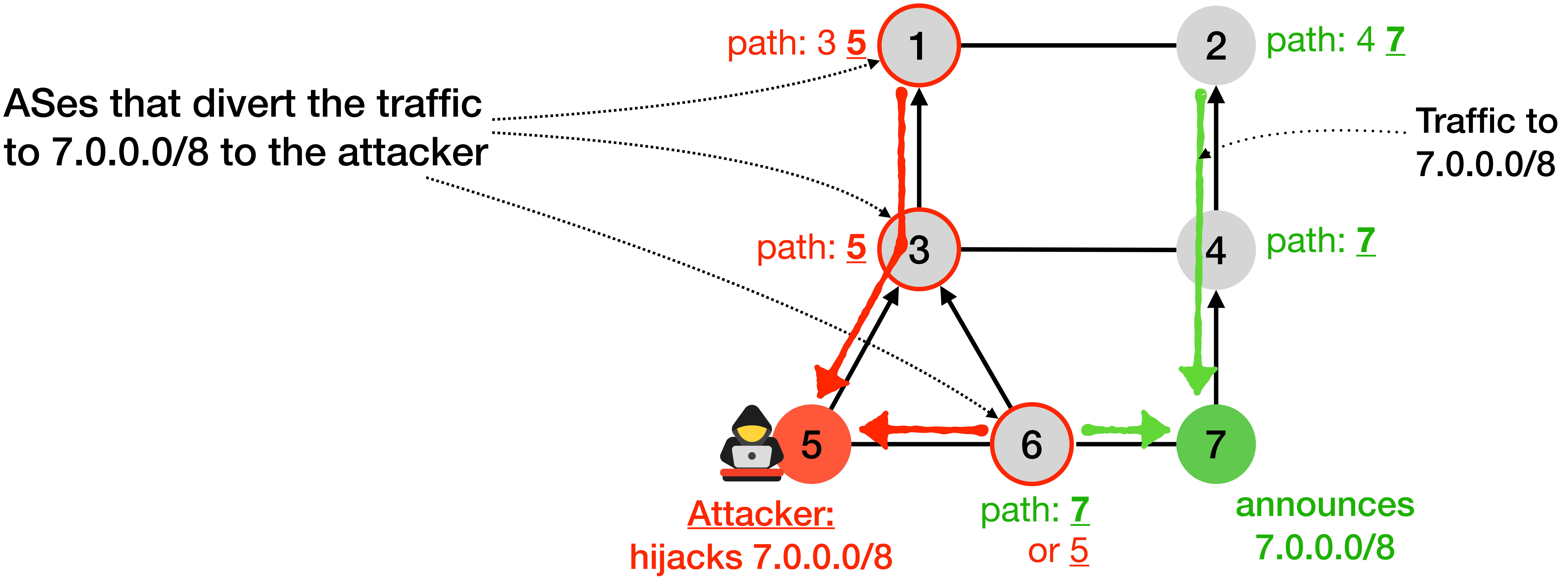
Internet routing (BGP) is vulnerable to traffic hijacking



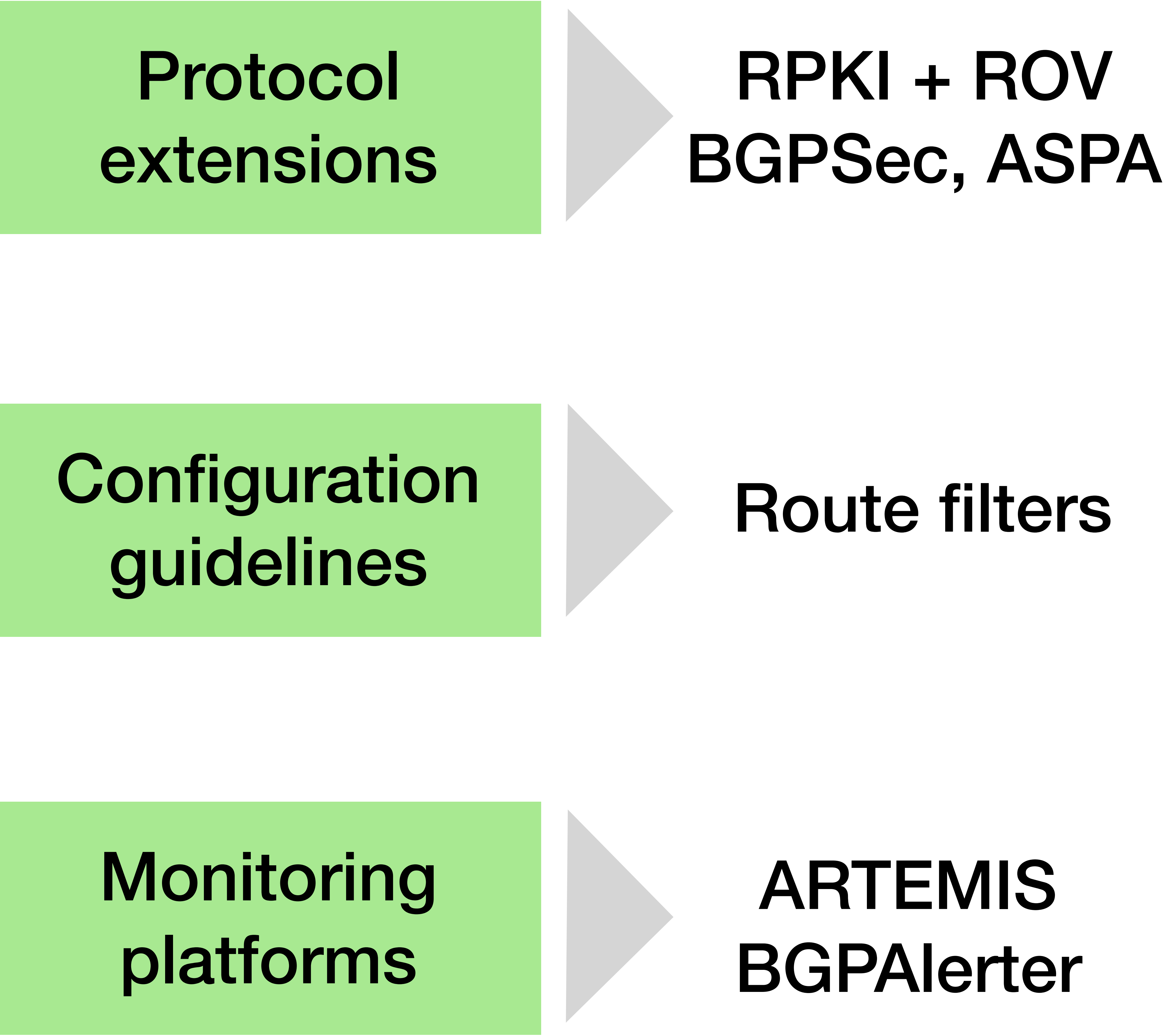
Internet routing (BGP) is vulnerable to traffic hijacking



Internet routing (BGP) is vulnerable to traffic hijacking



Fortunately, there are defenses against BGP hijacking



Protocol extensions

RPKI + ROV
BGPSec, ASPA

Configuration guidelines

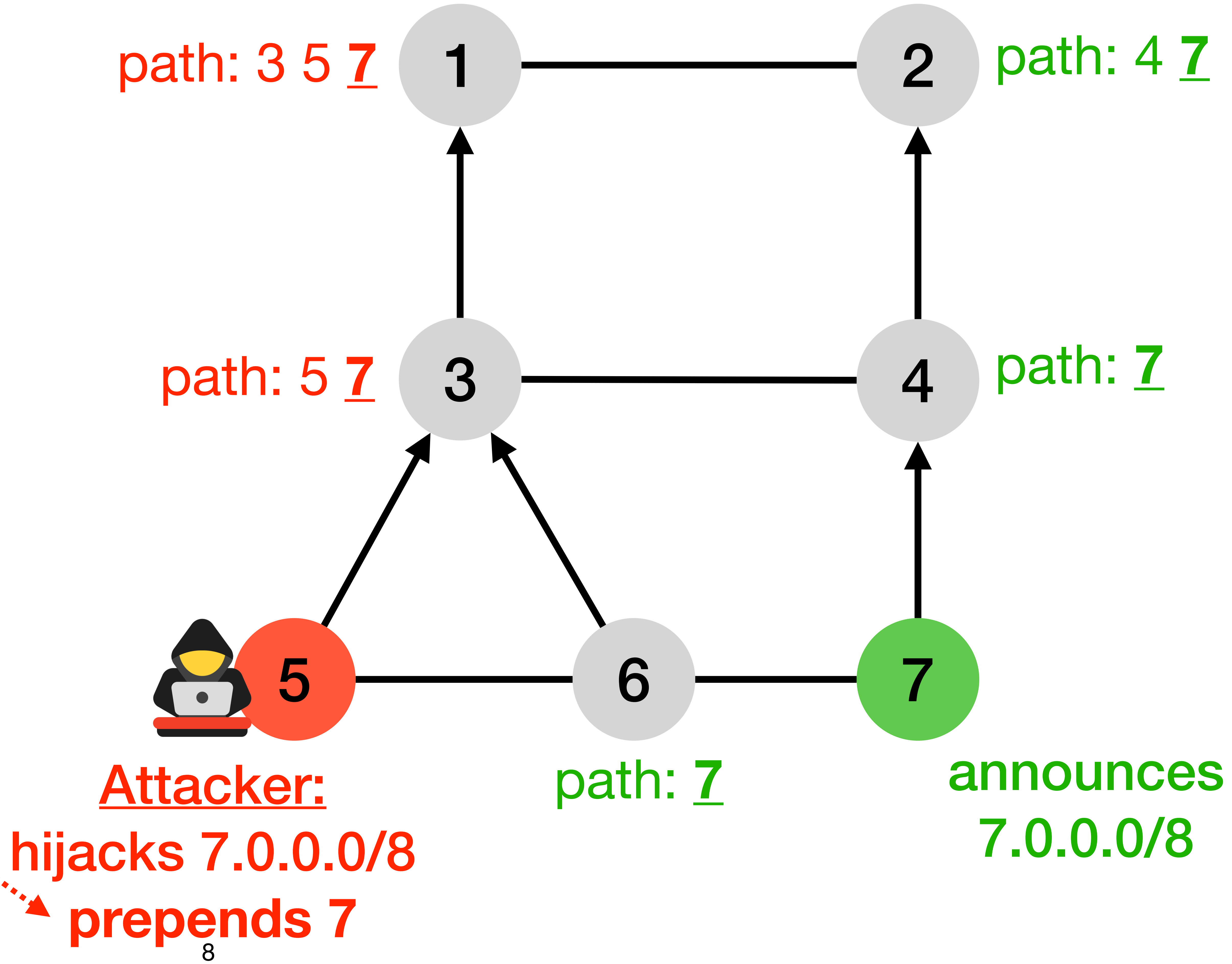
Route filters

Monitoring platforms

ARTEMIS
BGPAlerter

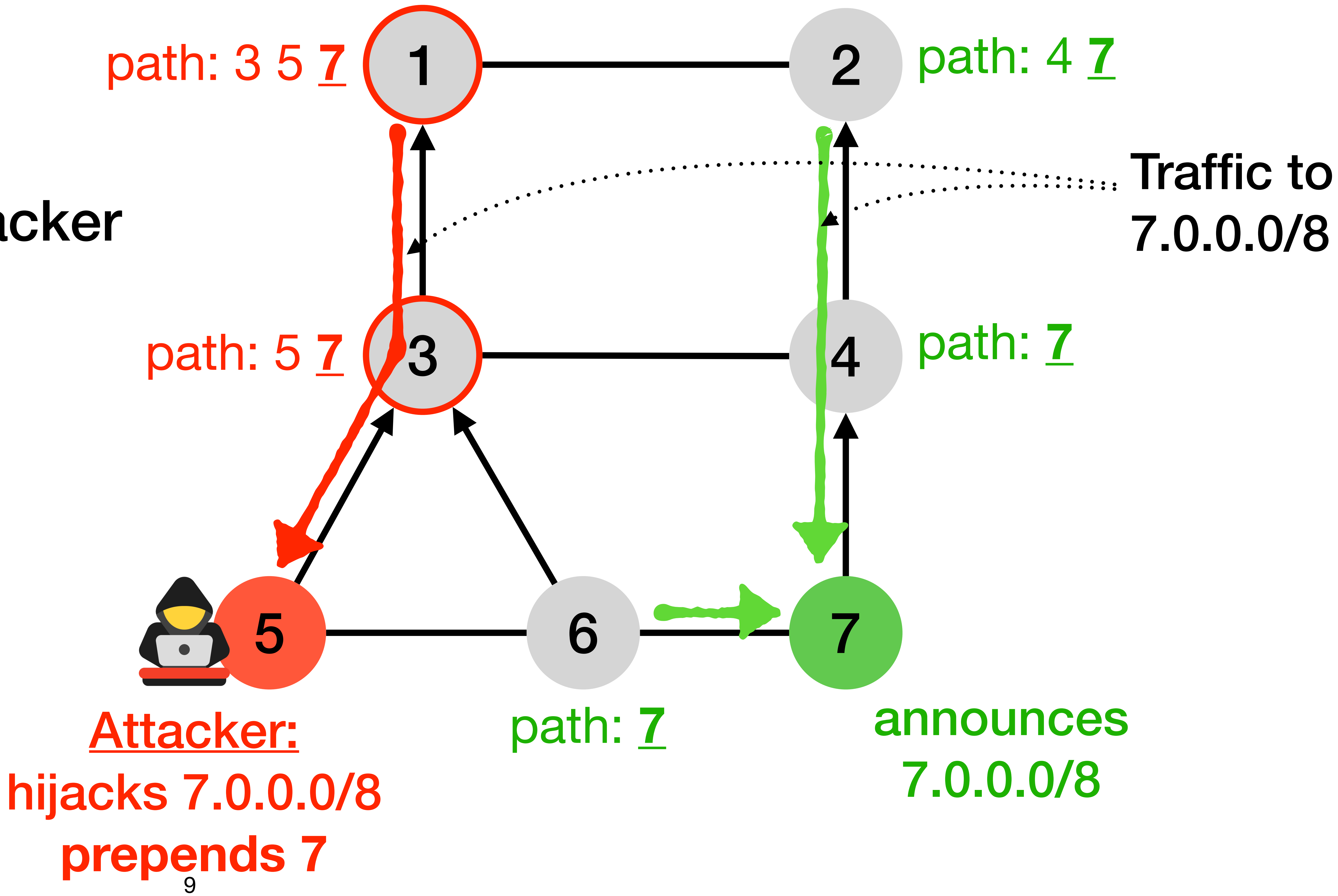
Despite the efforts, BGP is *still* vulnerable to forged-origin hijacks

The attacker prepends the legitimate AS number to the AS path



Despite the efforts, BGP is *still* vulnerable to forged-origin hijacks

Less but still a significant fraction of the traffic is diverted to the attacker

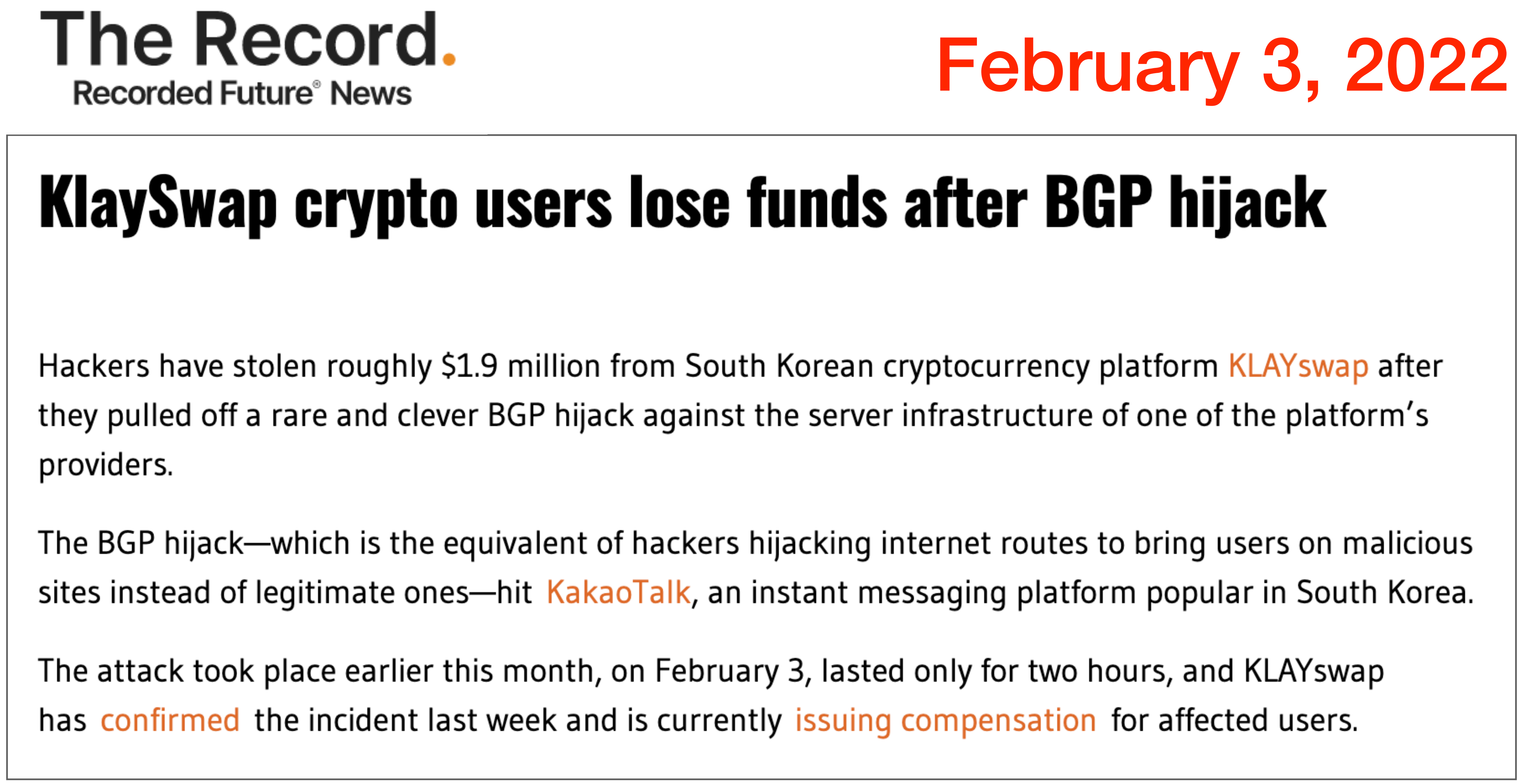


Existing defenses poorly neutralise forged-origin hijacks

Protocol extensions	RPKI + ROV BGPSec, ASPA	RPKI+ROV can’t detect forged-origin hijacks BGPSec and ASPA will take years to be widely deployed
Configuration guidelines	Route filters	Often missing and inaccurate as they are constructed based on the IRR
Monitoring platforms	ARTEMIS BGPAlerter	Narrowly focused as they detect hijacks that only pertain to the AS deploying it

Forged-origin hijacks are actively used by attackers

February 3, 2022



The Record.
Recorded Future® News

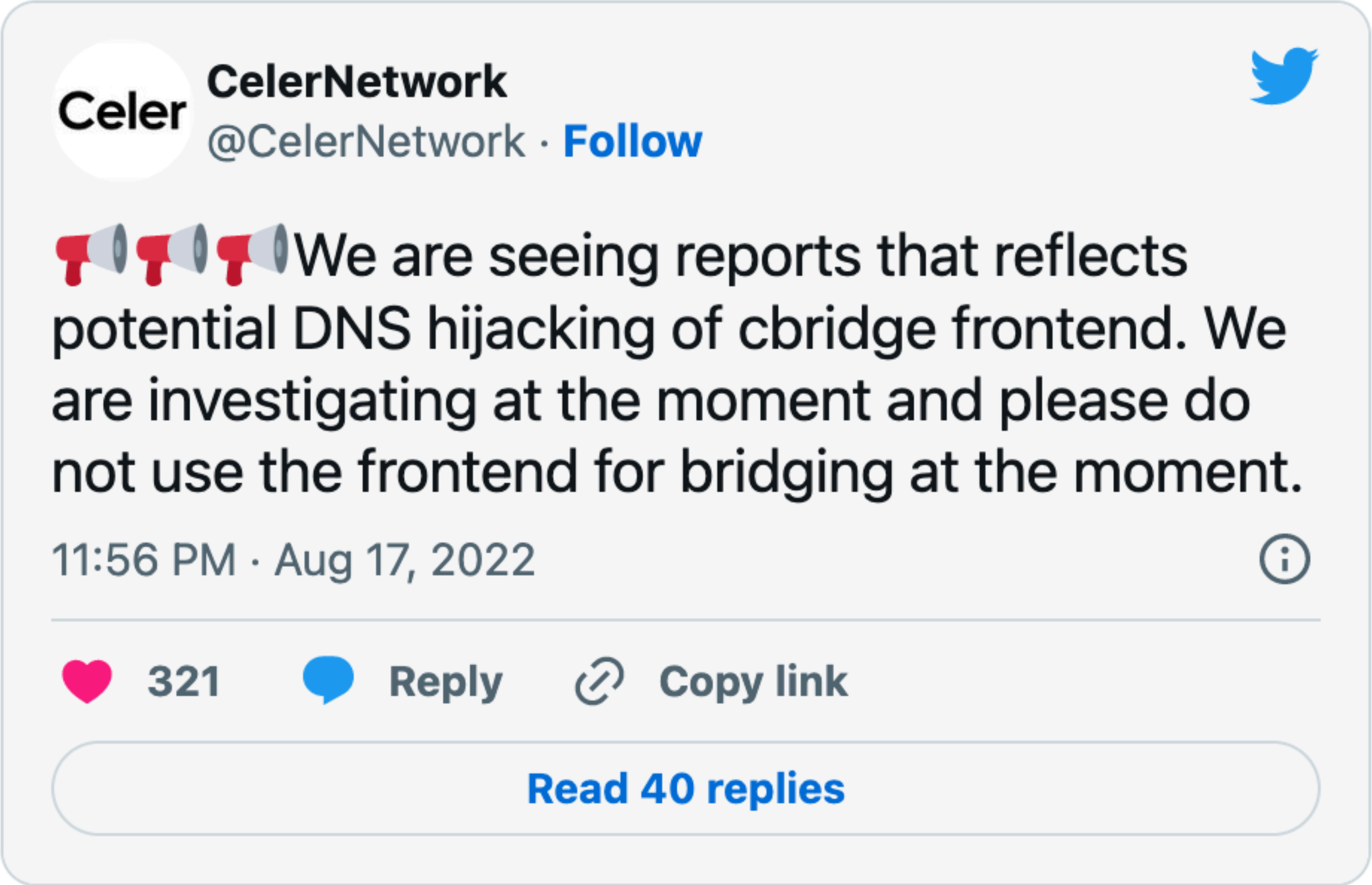
KlaySwap crypto users lose funds after BGP hijack

Hackers have stolen roughly $1.9 million from South Korean cryptocurrency platform KLAYswap after they pulled off a rare and clever BGP hijack against the server infrastructure of one of the platform's providers.

The BGP hijack—which is the equivalent of hackers hijacking internet routes to bring users on malicious sites instead of legitimate ones—hit KakaoTalk, an instant messaging platform popular in South Korea.

The attack took place earlier this month, on February 3, lasted only for two hours, and KLAYswap has confirmed the incident last week and is currently issuing compensation for affected users.

August 17, 2022



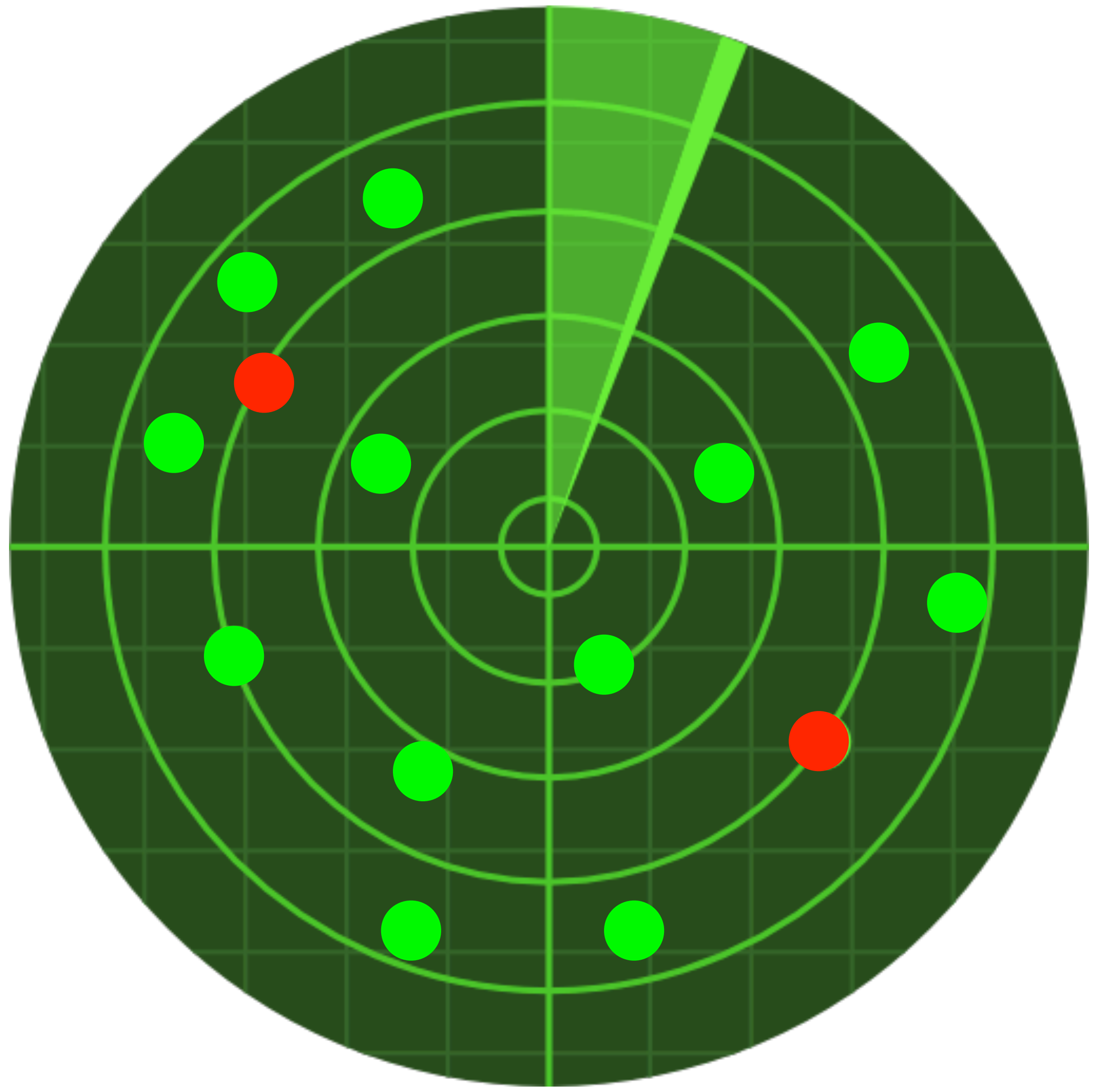
Celer
CelerNetwork
@CelerNetwork · Follow

📢📢📢We are seeing reports that reflects potential DNS hijacking of cbridge frontend. We are investigating at the moment and please do not use the frontend for bridging at the moment.

11:56 PM · Aug 17, 2022

Both attacks are the result of a forged-origin hijack

DFOH: A System to Detect Forged-Origin BGP Hijacks on the Whole Internet



Outline

DFOH’s main challenge

DFOH’s inference pipeline

DFOH’s inferences are accurate

DFOH is up and running

Outline

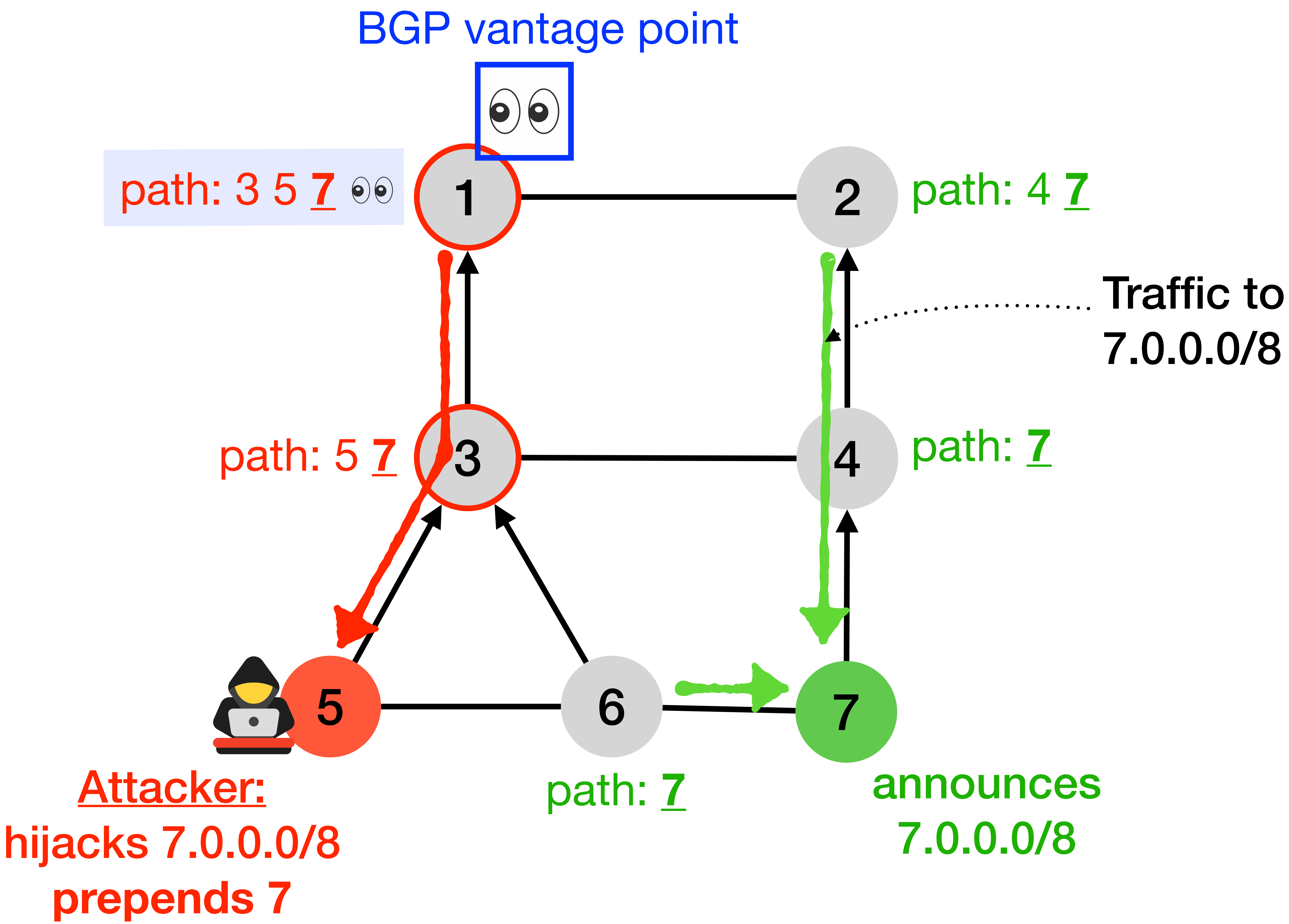
DFOH’s main challenge is to detect fake AS links

DFOH’s inference pipeline

DFOH’s inferences are accurate

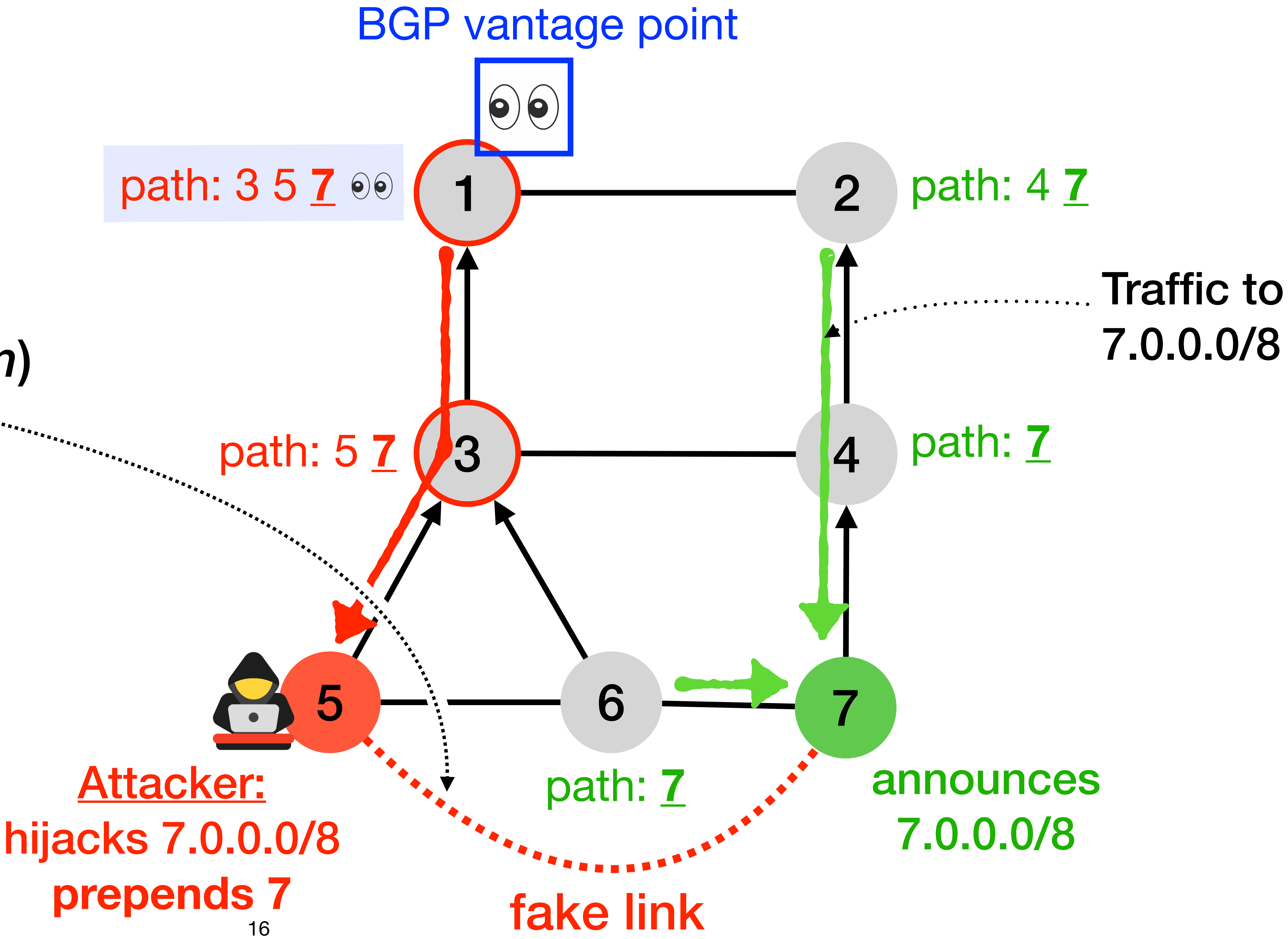
DFOH is up and running

DFOH aims to detect the fake AS links induced by forged-origin hijacks



DFOH aims to detect the fake AS links induced by forged-origin hijacks

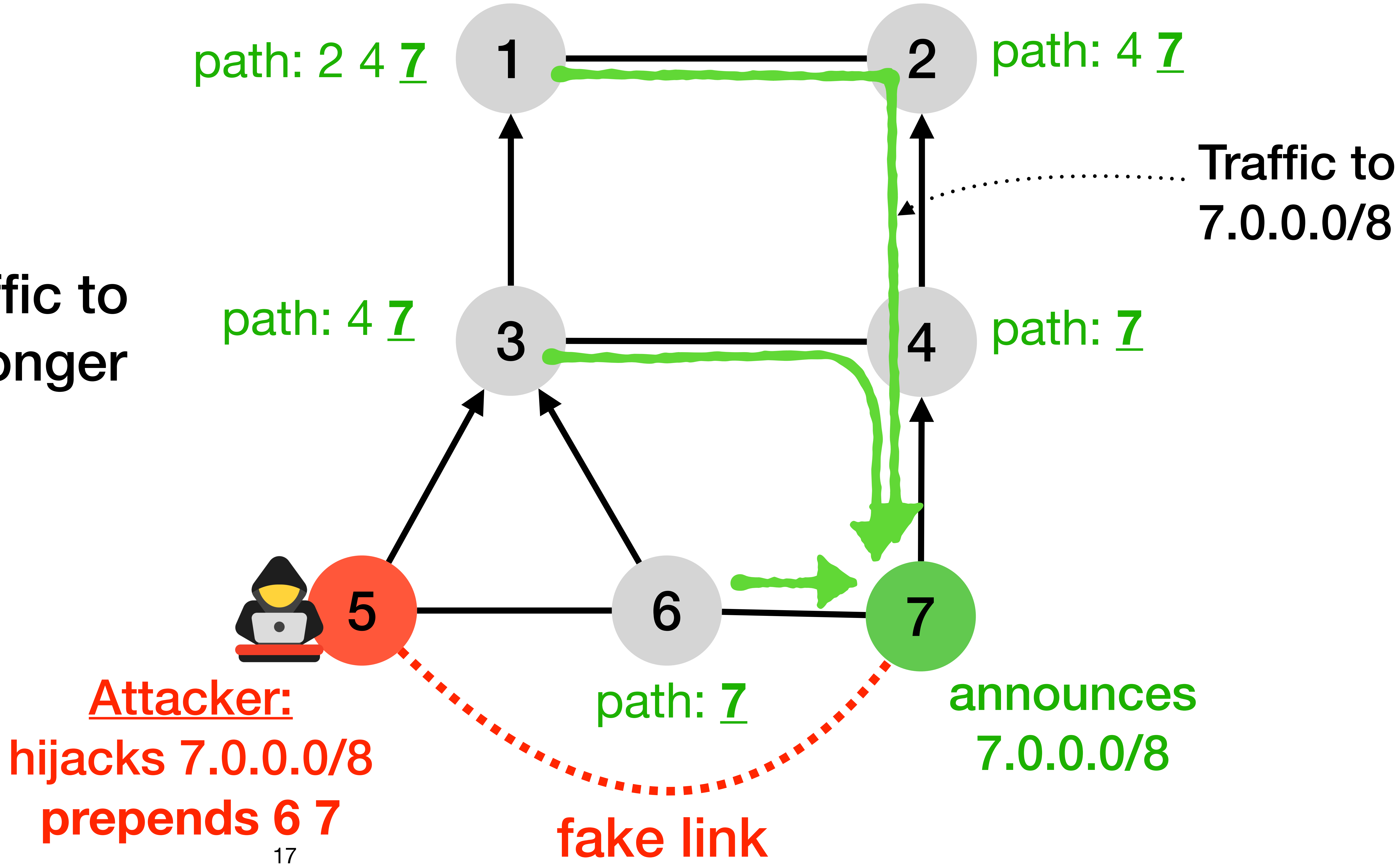
<ins>Upon the attack:</ins>
AS5 (*attacker*) and AS7 (*victim*) appear directly connected



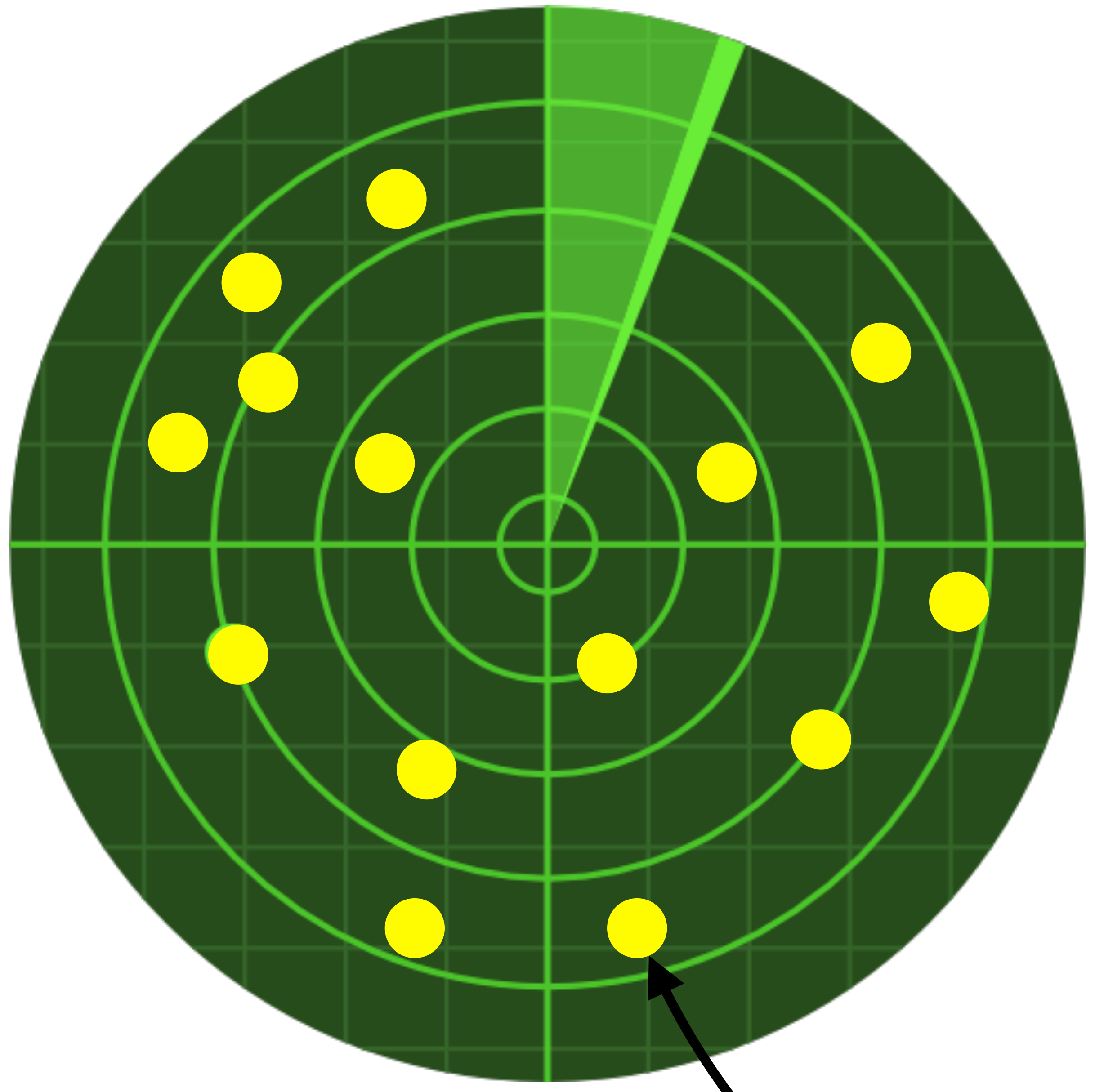
An attacker cannot escape from creating a fake AS link without hampering the effectiveness of its attack

There is no new AS link if the attacker prepends **6 7**

But none of the ASes divert traffic to the attacker as the AS path is longer



Problem: There are many new AS links every day but no simple property that tells whether they are real or fake

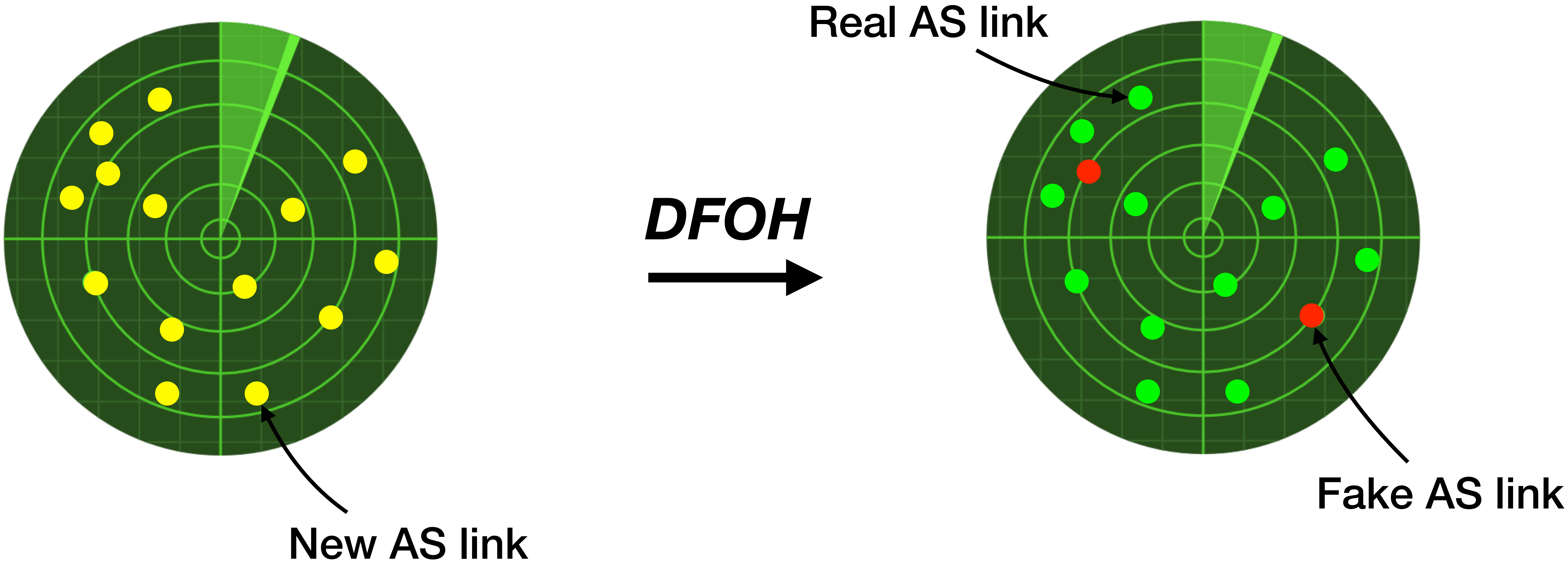


New AS link

We find 166 new AS links every day (median) and the vast majority are likely legitimate

Using the BGP data from 200 RIS and RouteViews peers and collected during ten months in 2022

<u>Problem:</u> There are many new AS links every day
but no simple property that tells whether they are real or fake



New AS link

DFOH

Real AS link

Fake AS link

Outline

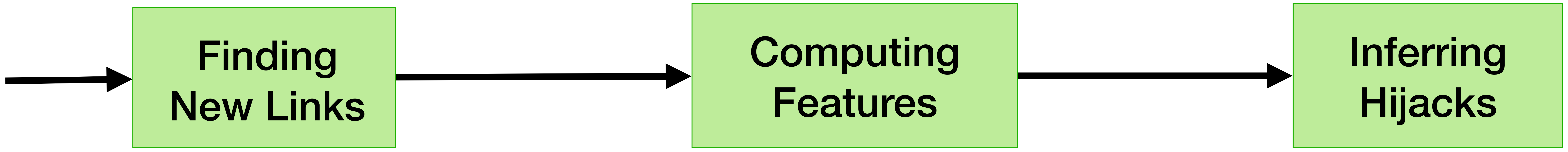
DFOH's main challenge is to detect fake AS links

DFOH's inference pipeline relies on domain-specific knowledge and a tailored link prediction framework

DFOH's inferences are accurate

DFOH is up and running

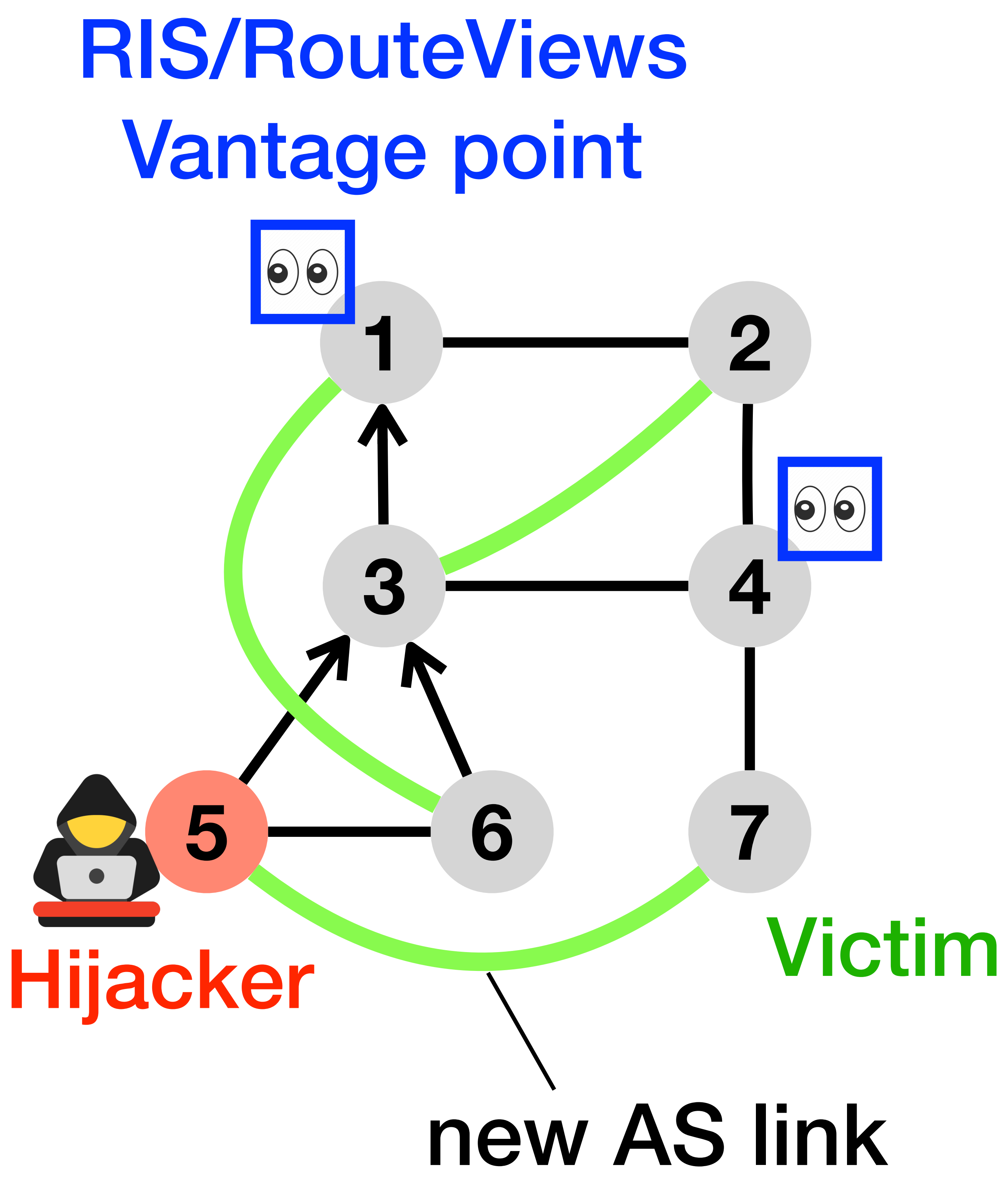
DFOH's fake AS links inference algorithm comprises three steps



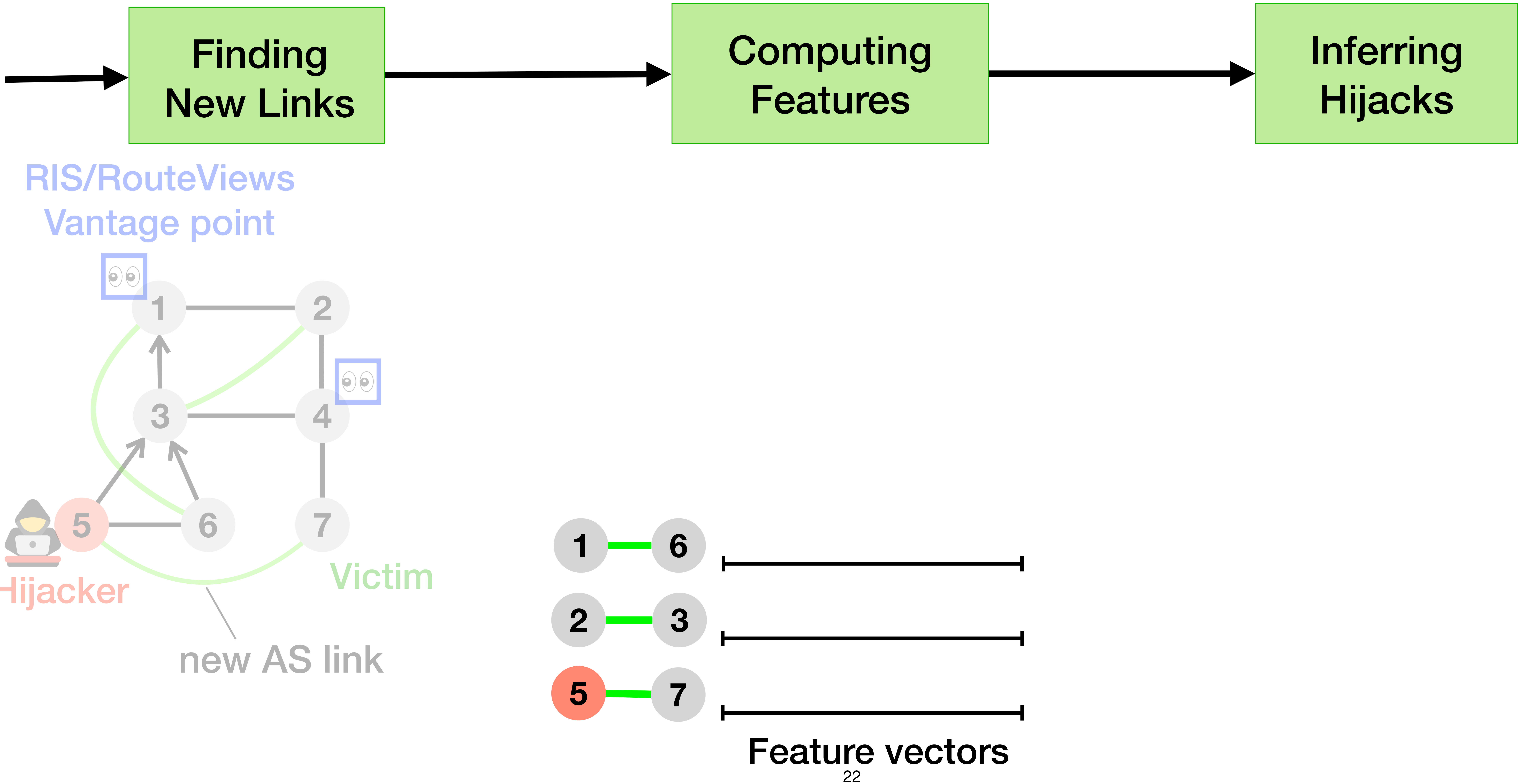
Finding New Links

Computing Features

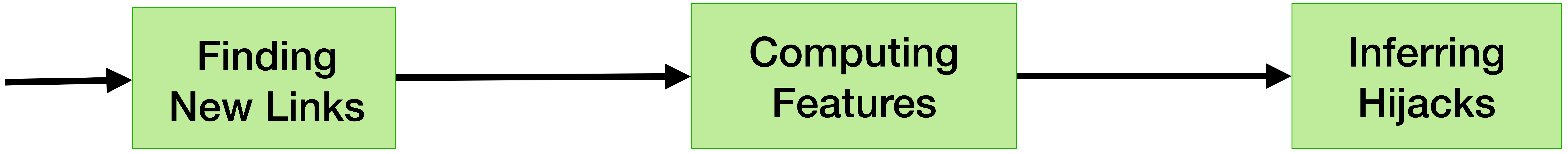
Inferring Hijacks

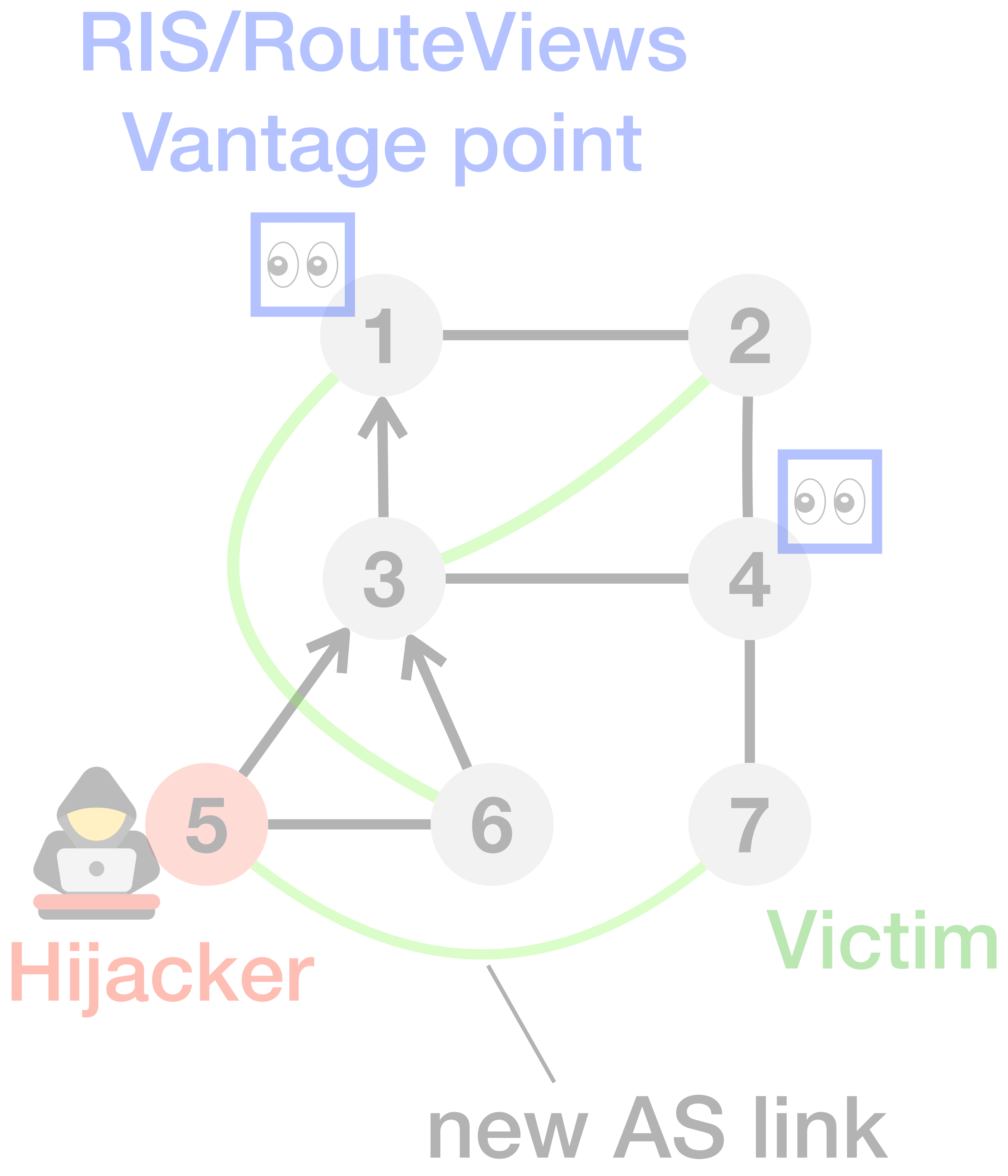


DFOH's fake AS links inference algorithm comprises three steps

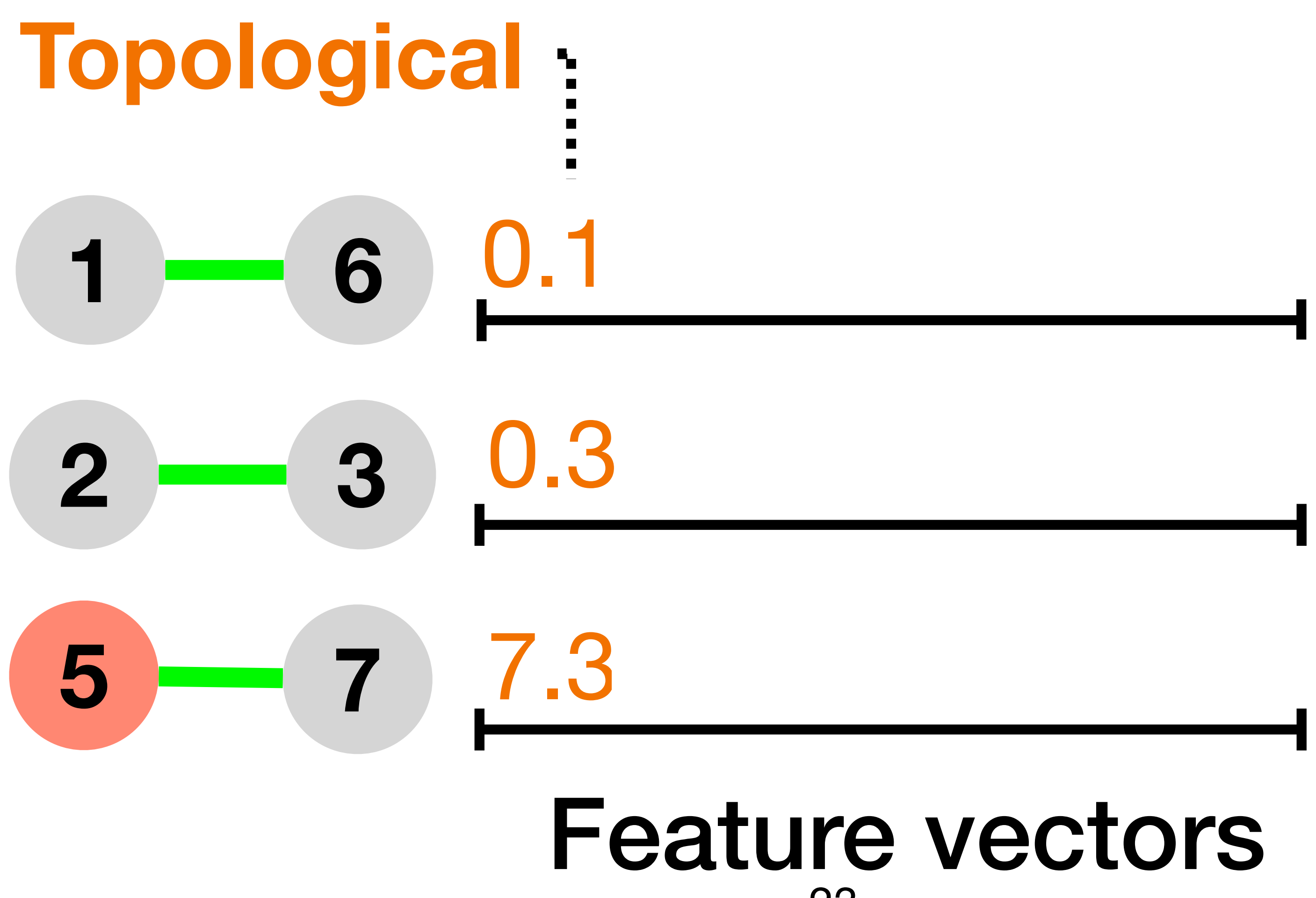


DFOH's fake AS links inference algorithm comprises three steps



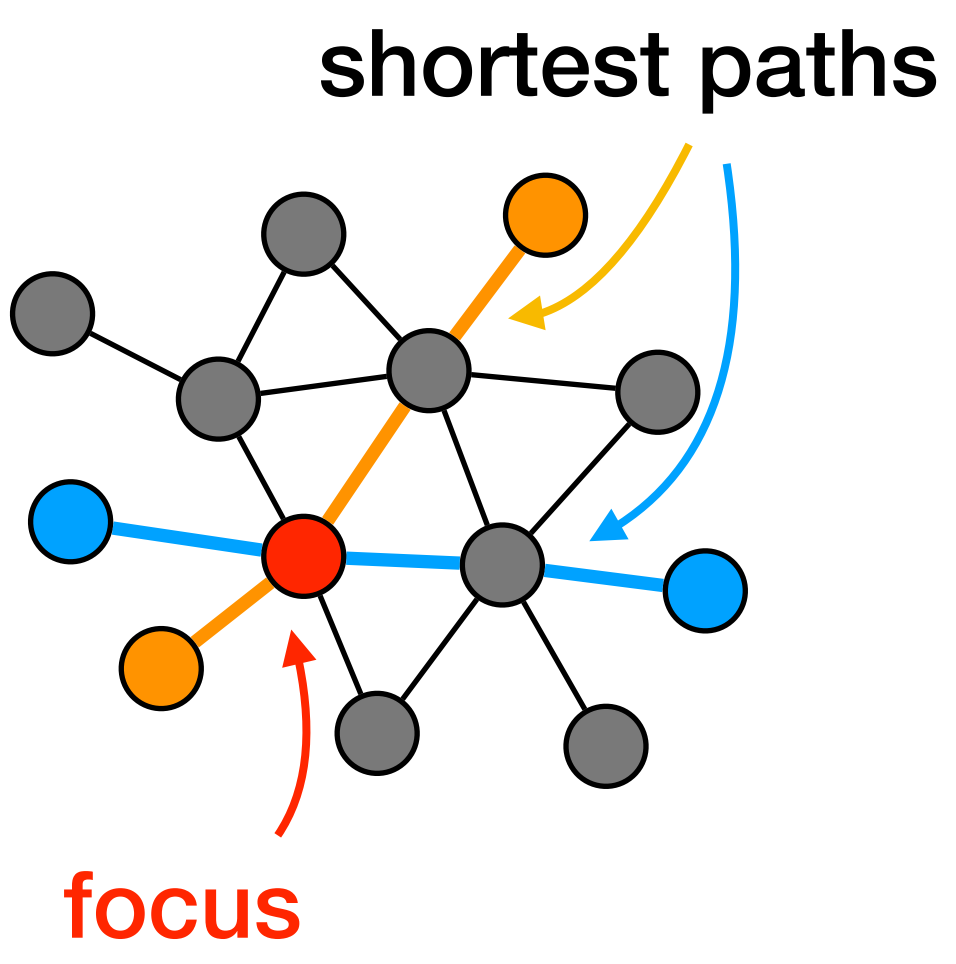


Feature categories:

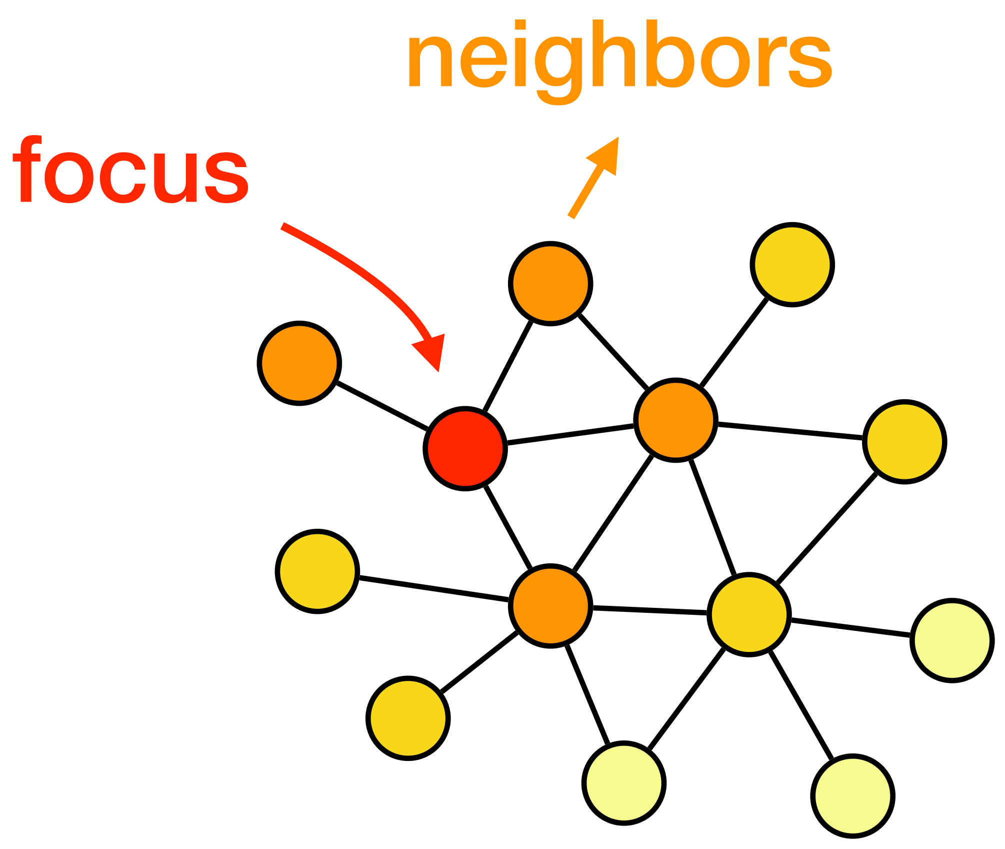


DFOH uses a total of 11 topological features that can be divided into four categories

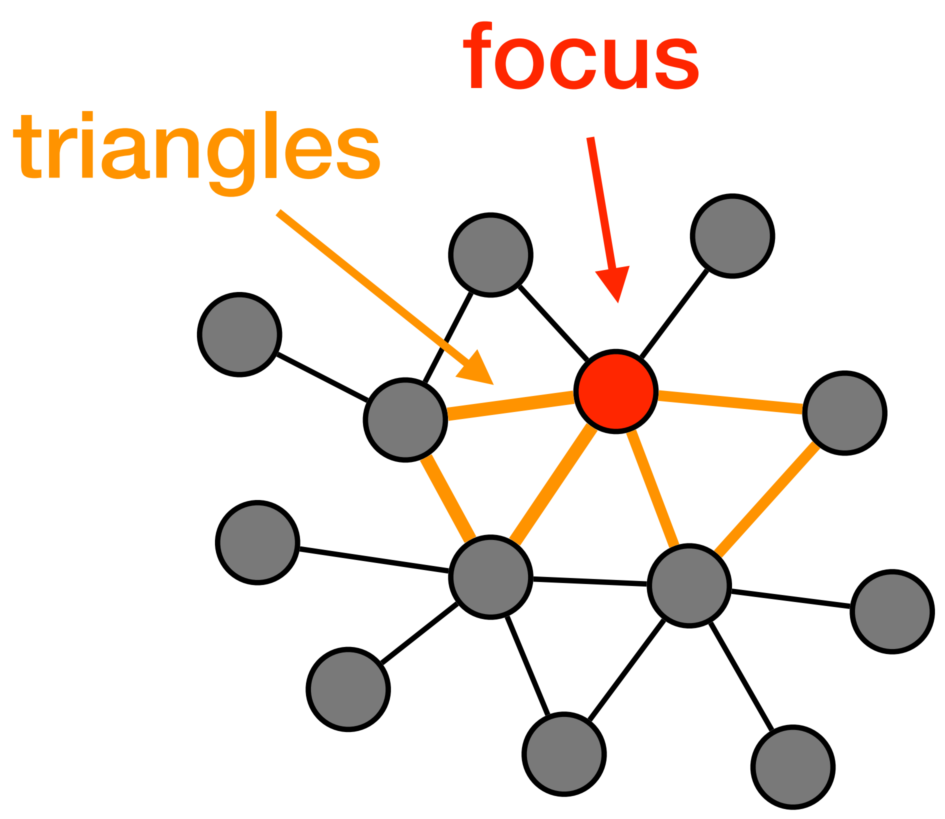
Node centrality



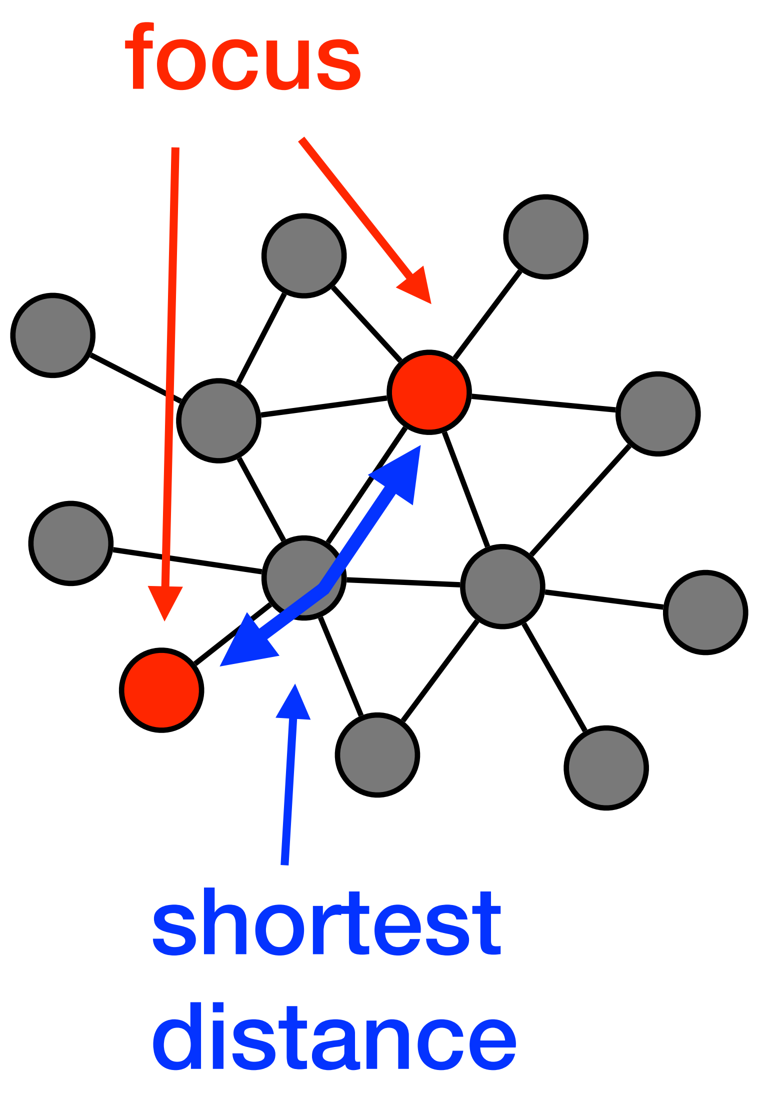
Neighborhood richness



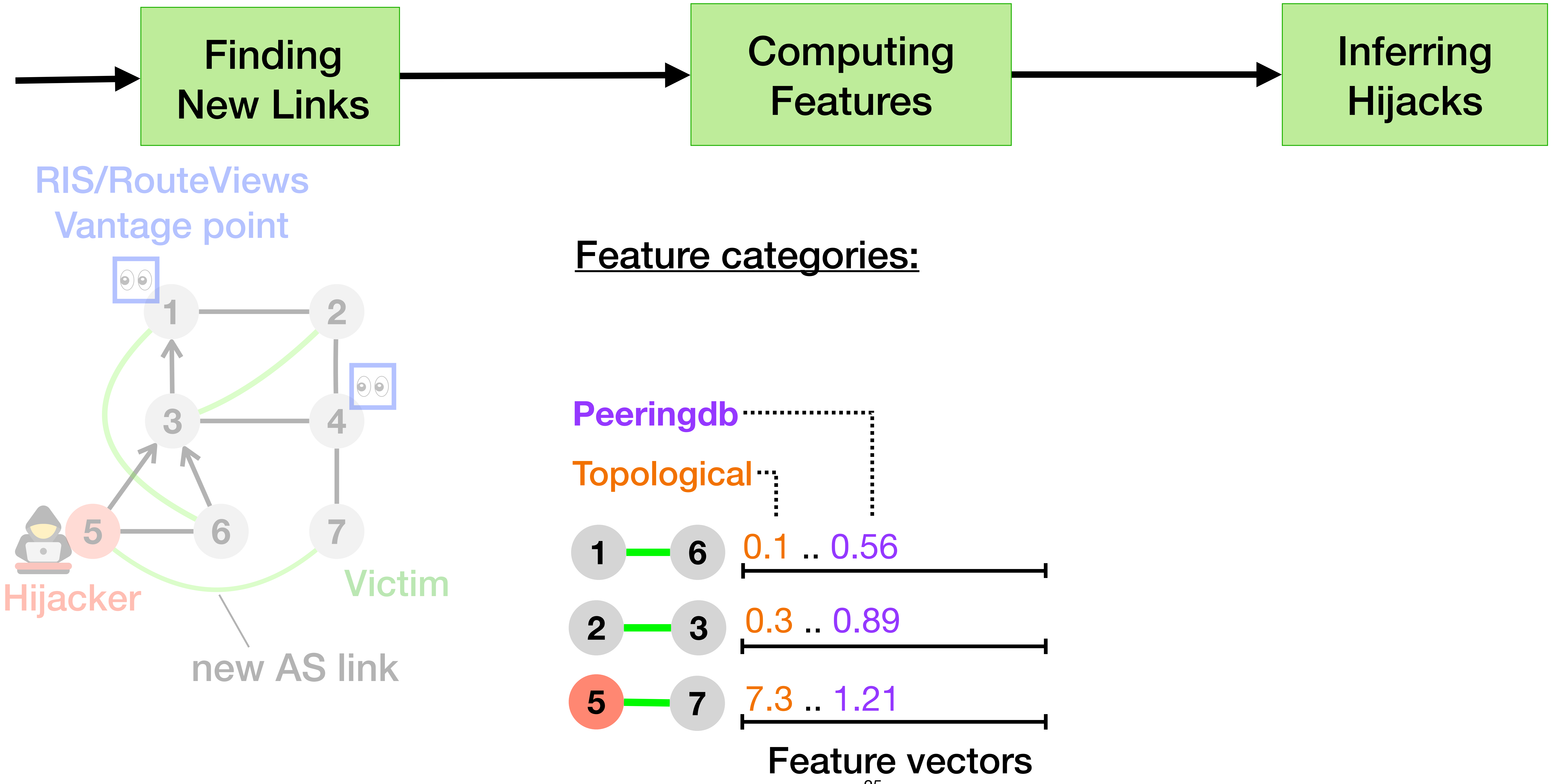
Topological patterns



Closeness



DFOH's fake AS links inference algorithm comprises three steps



DFOH looks at public peering information and identifies when two ASes are unlikely to peer

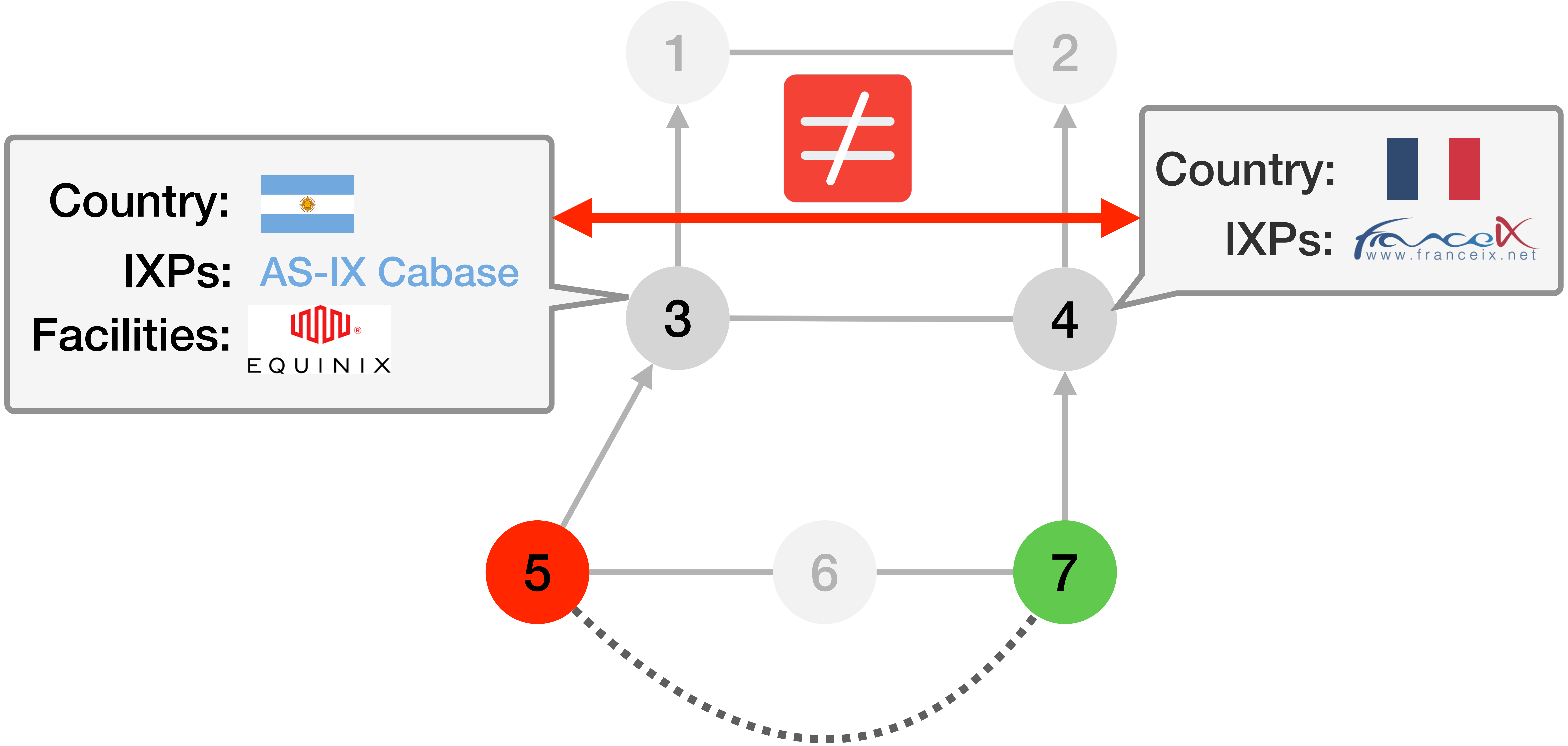
DFOH looks for three types of information in PeeringDB:

1. Country
2. Public peering exchange points
3. Private peering facilities

DFOH compares the peering information of the neighbors of the hypothetical victim and attacker

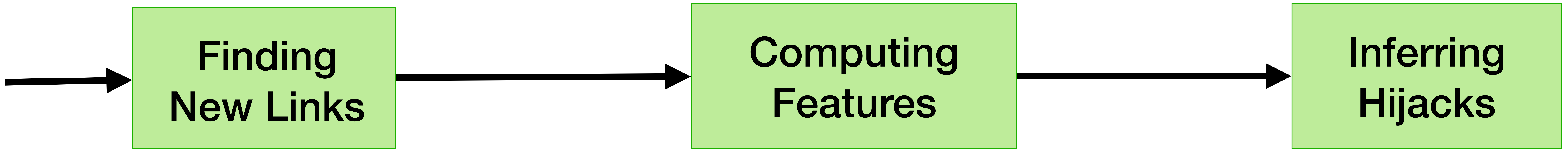
Reason #1:
Protect against adversarial inputs

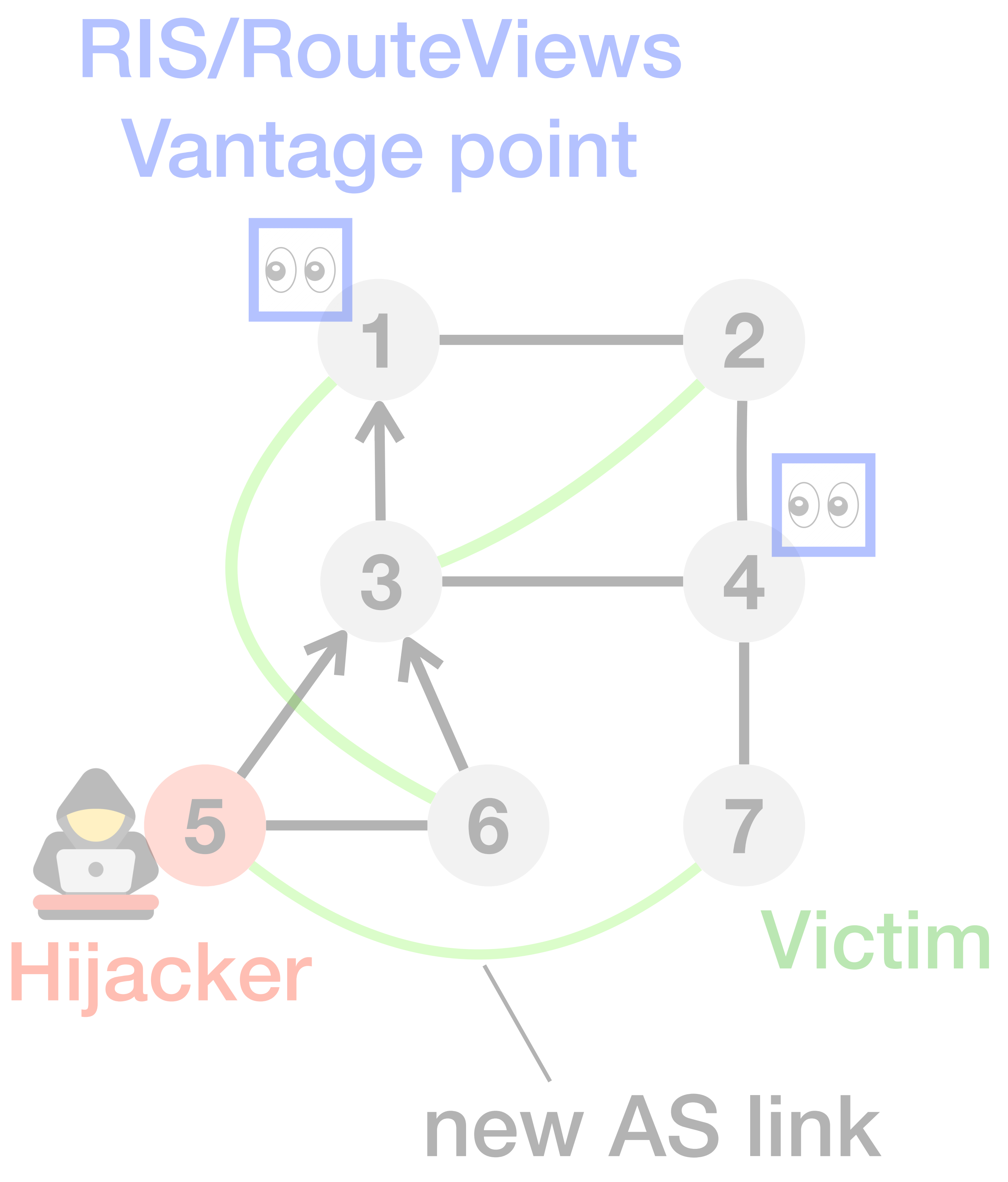
Reason #2:
Mitigate missing peering information

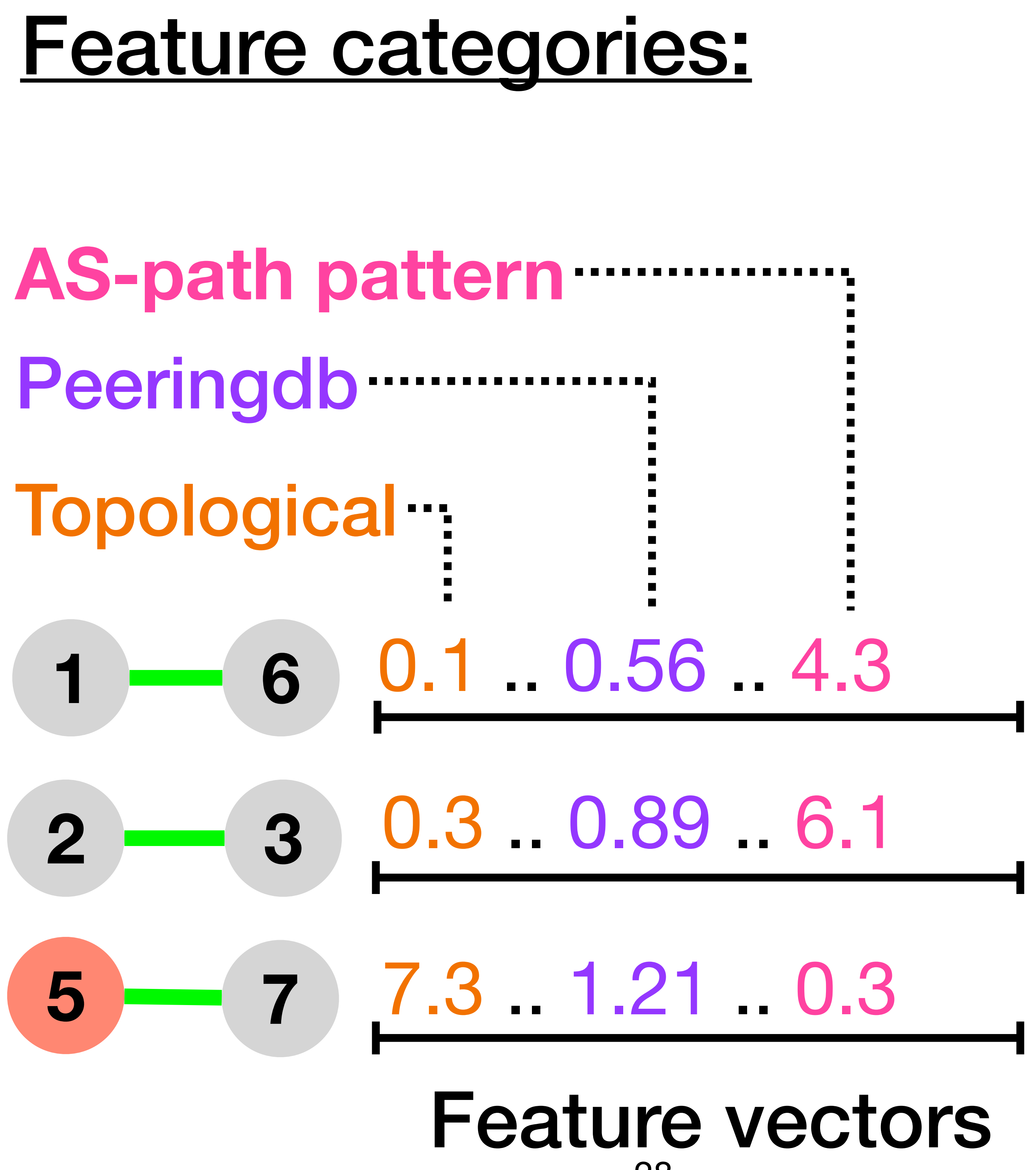


New AS link is fake

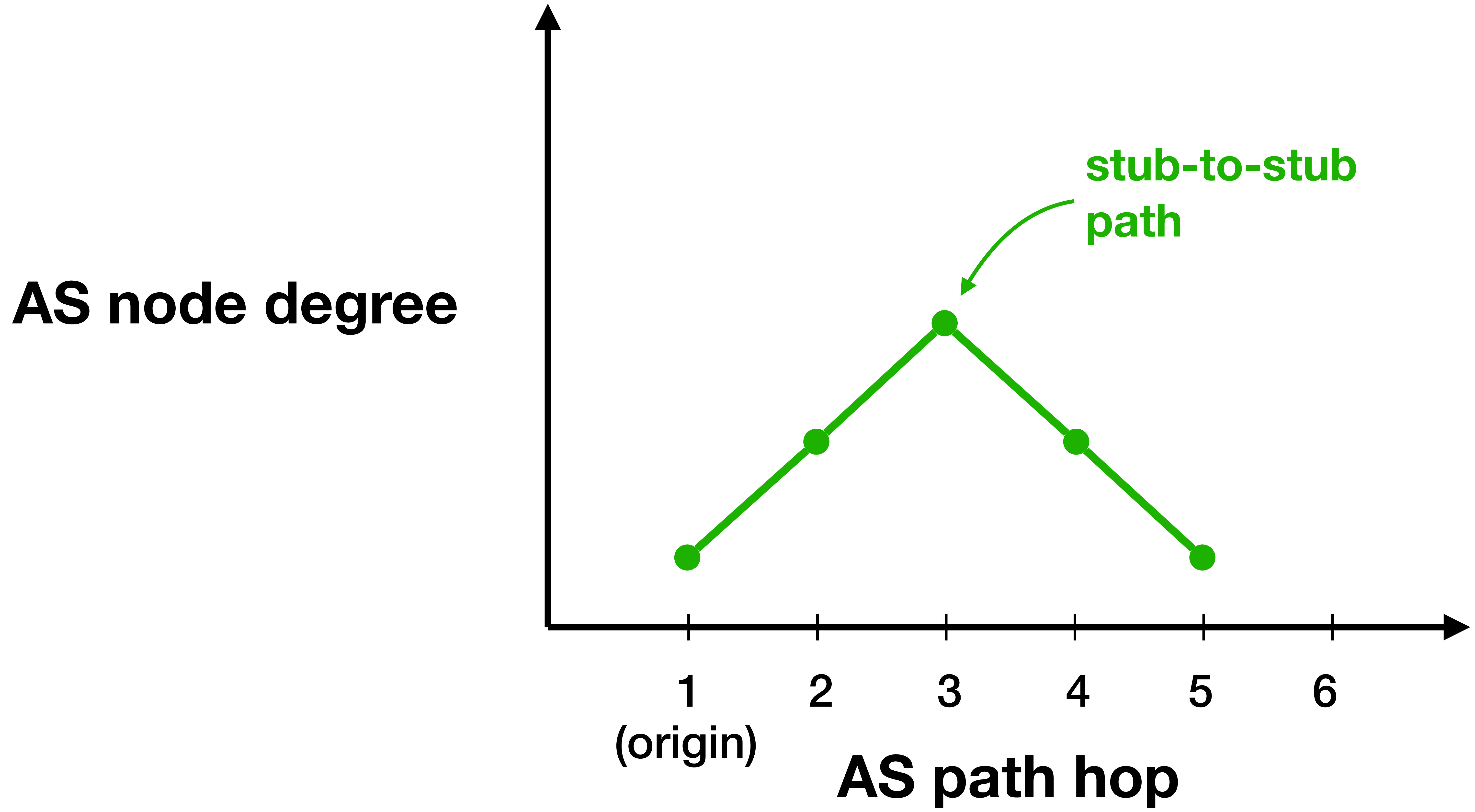
DFOH’s fake AS links inference algorithm comprises three steps



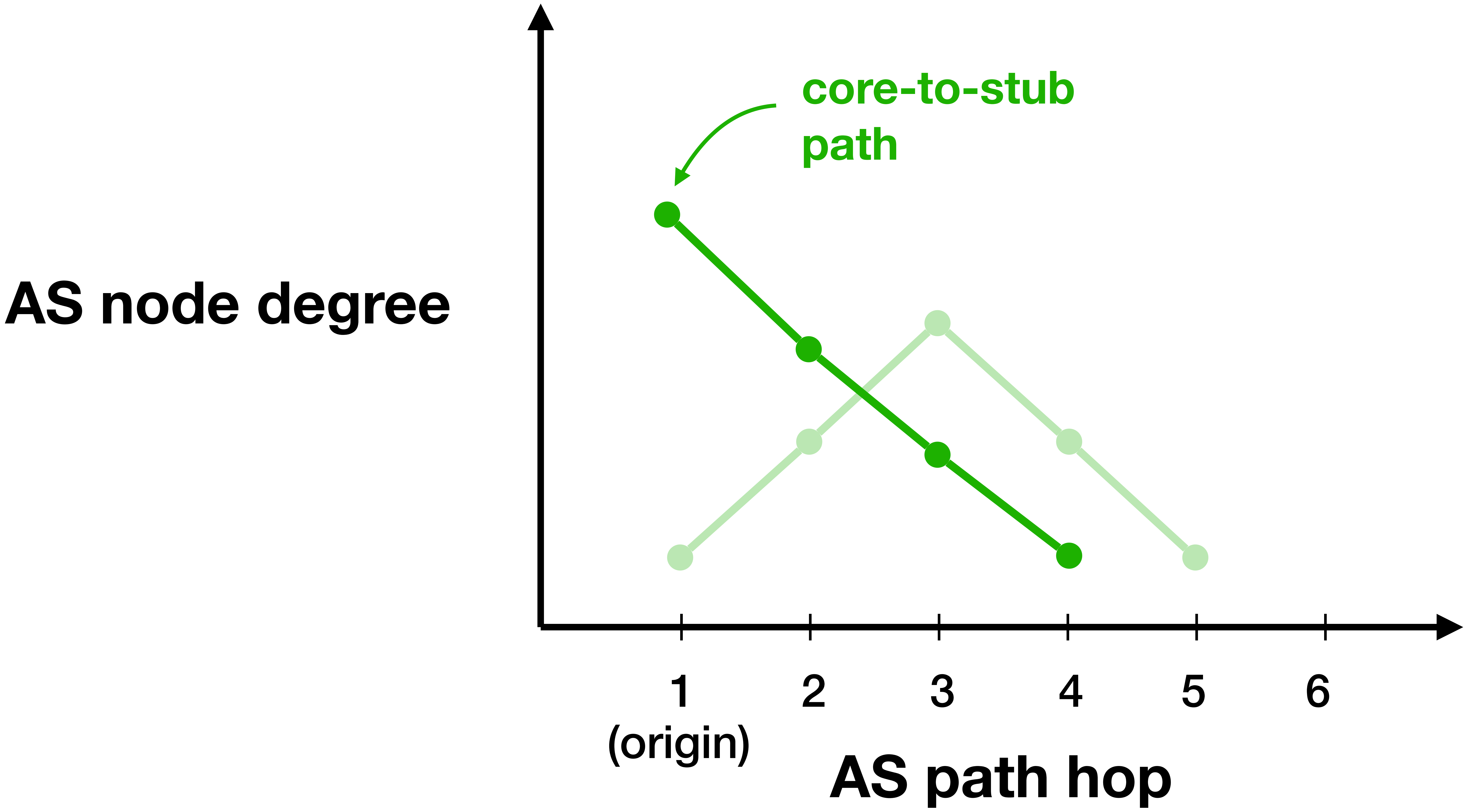




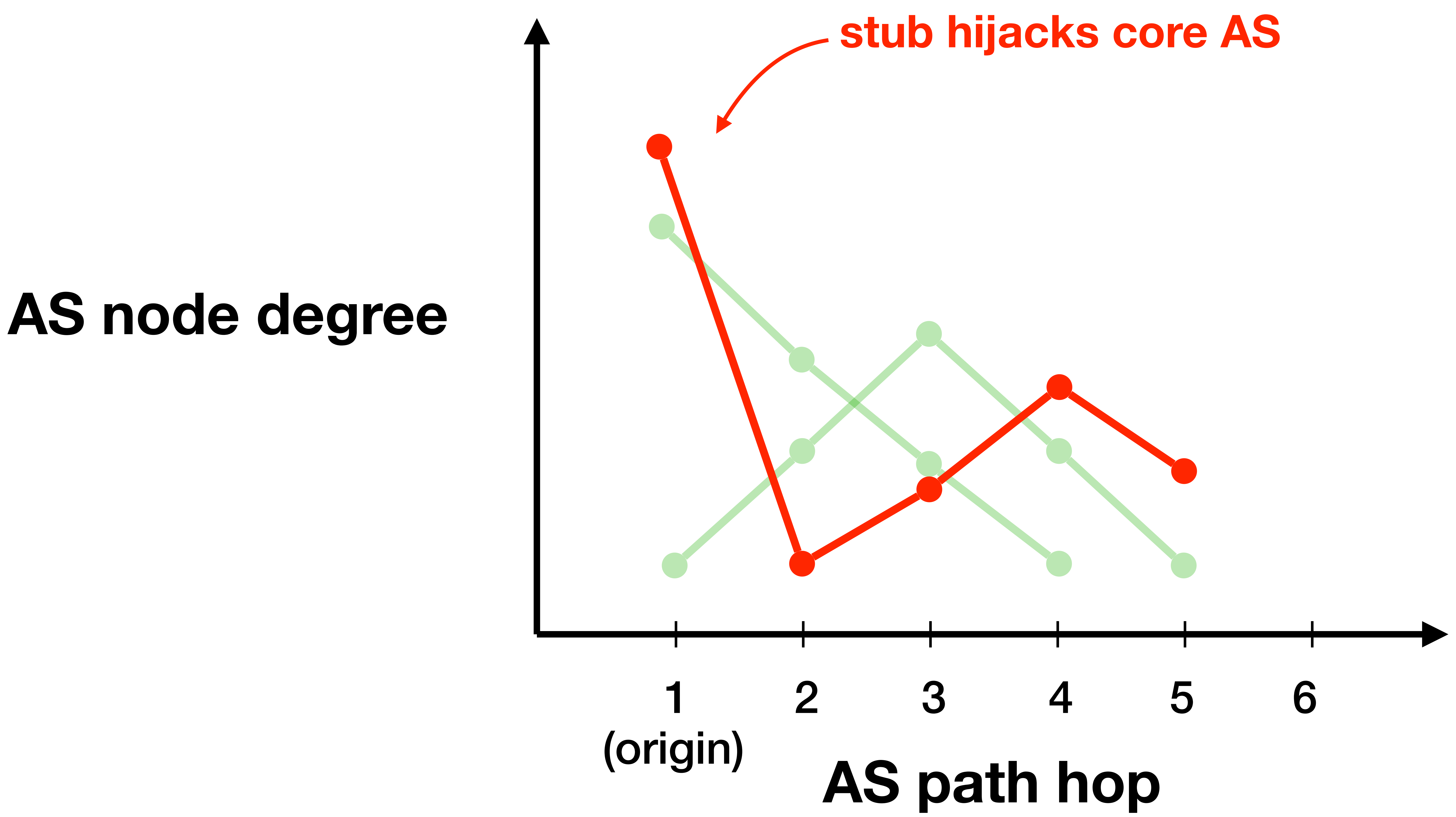
DFOH looks at the AS paths that include the new link and identifies suspicious sequence of ASes



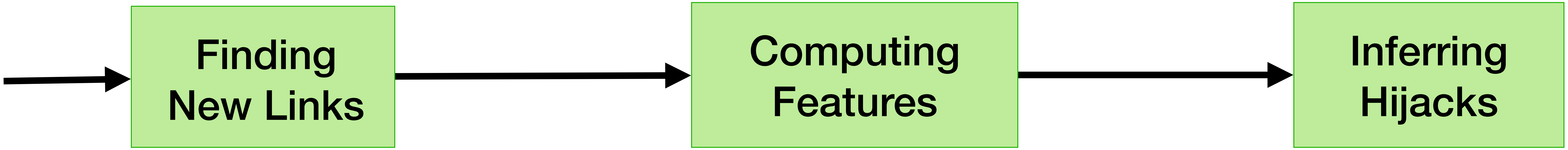
DFOH looks at the AS paths that include the new link and identifies suspicious sequence of ASes



DFOH looks at the AS paths that include the new link and identifies suspicious sequence of ASes



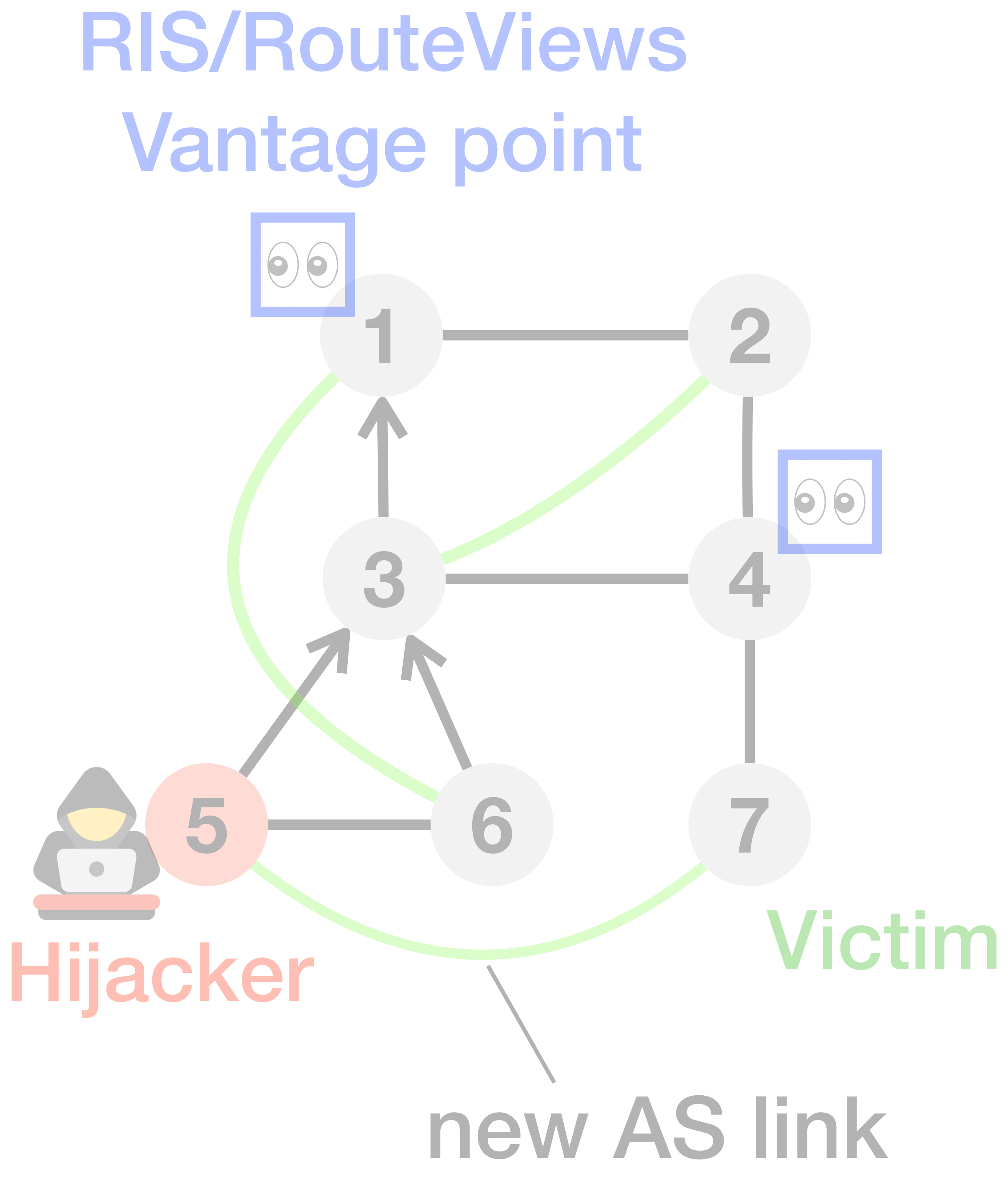
DFOH's fake AS links inference algorithm comprises three steps

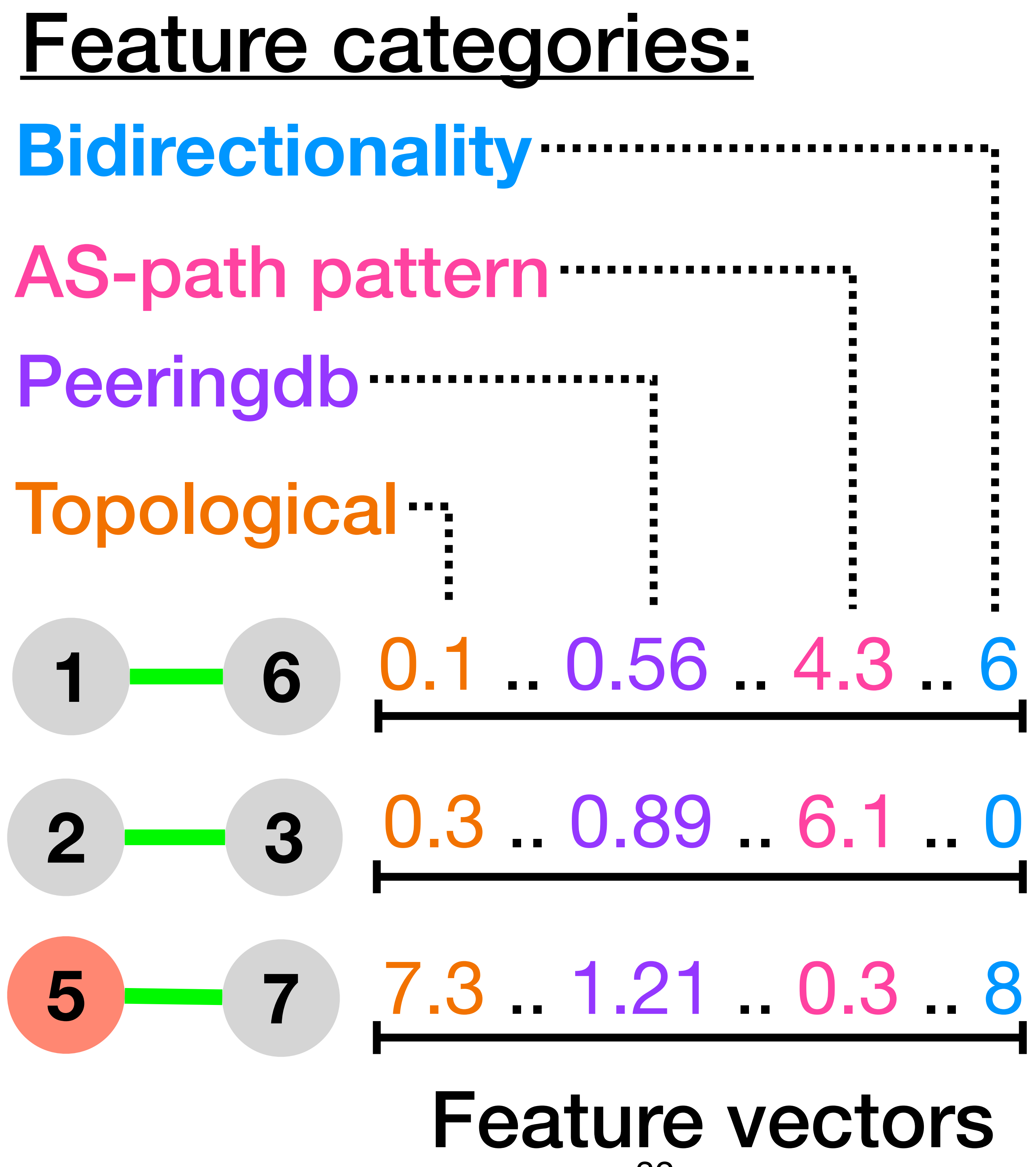


Finding New Links

Computing Features

Inferring Hijacks





Feature categories:

Bidirectionality

AS-path pattern

Peeringdb

Topological

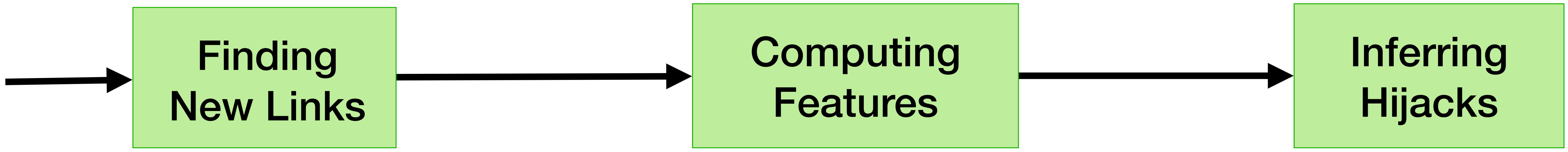
1 6 0.1 .. 0.56 .. 4.3 .. 6

2 3 0.3 .. 0.89 .. 6.1 .. 0

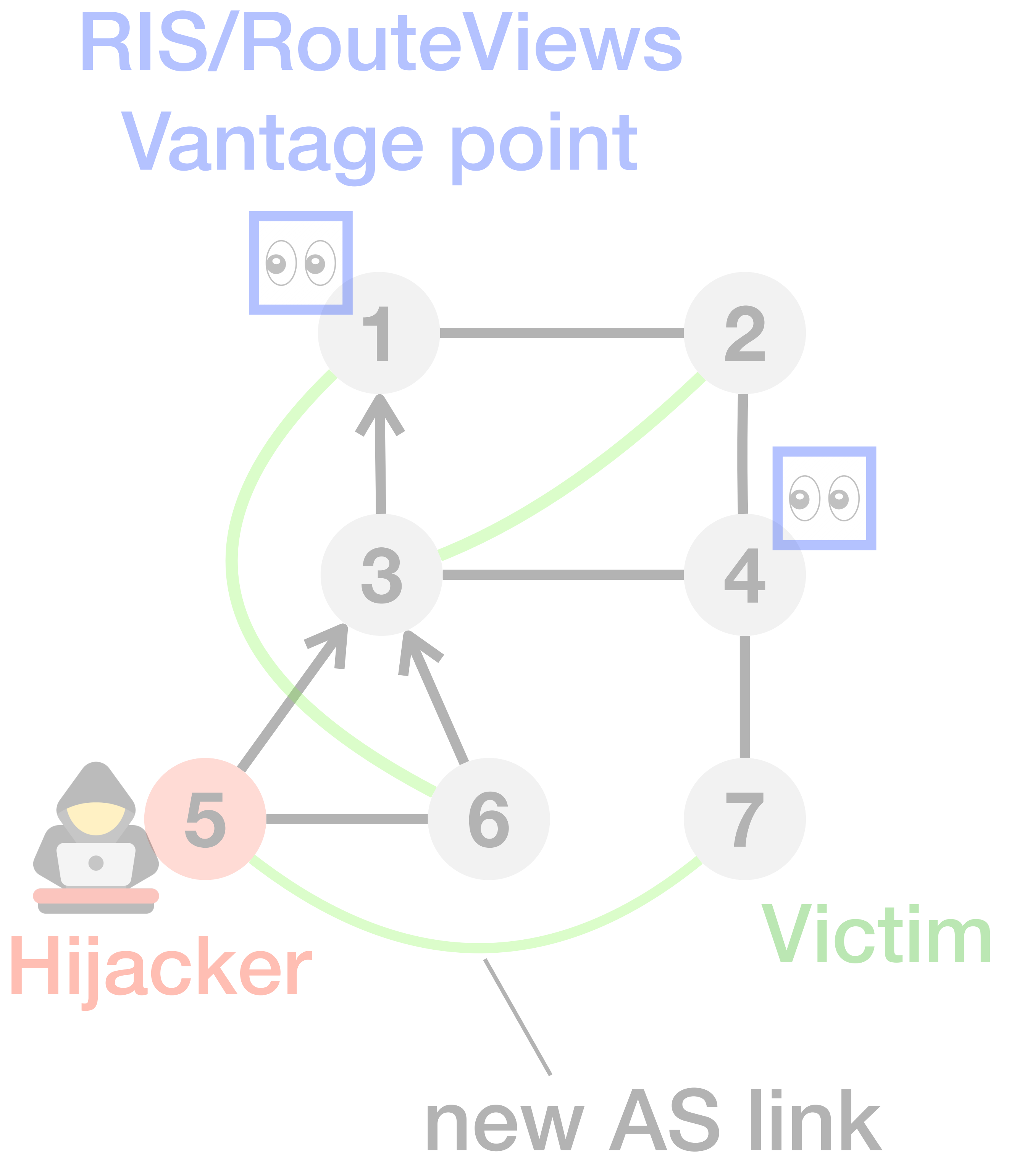
5 7 7.3 .. 1.21 .. 0.3 .. 8

Feature vectors

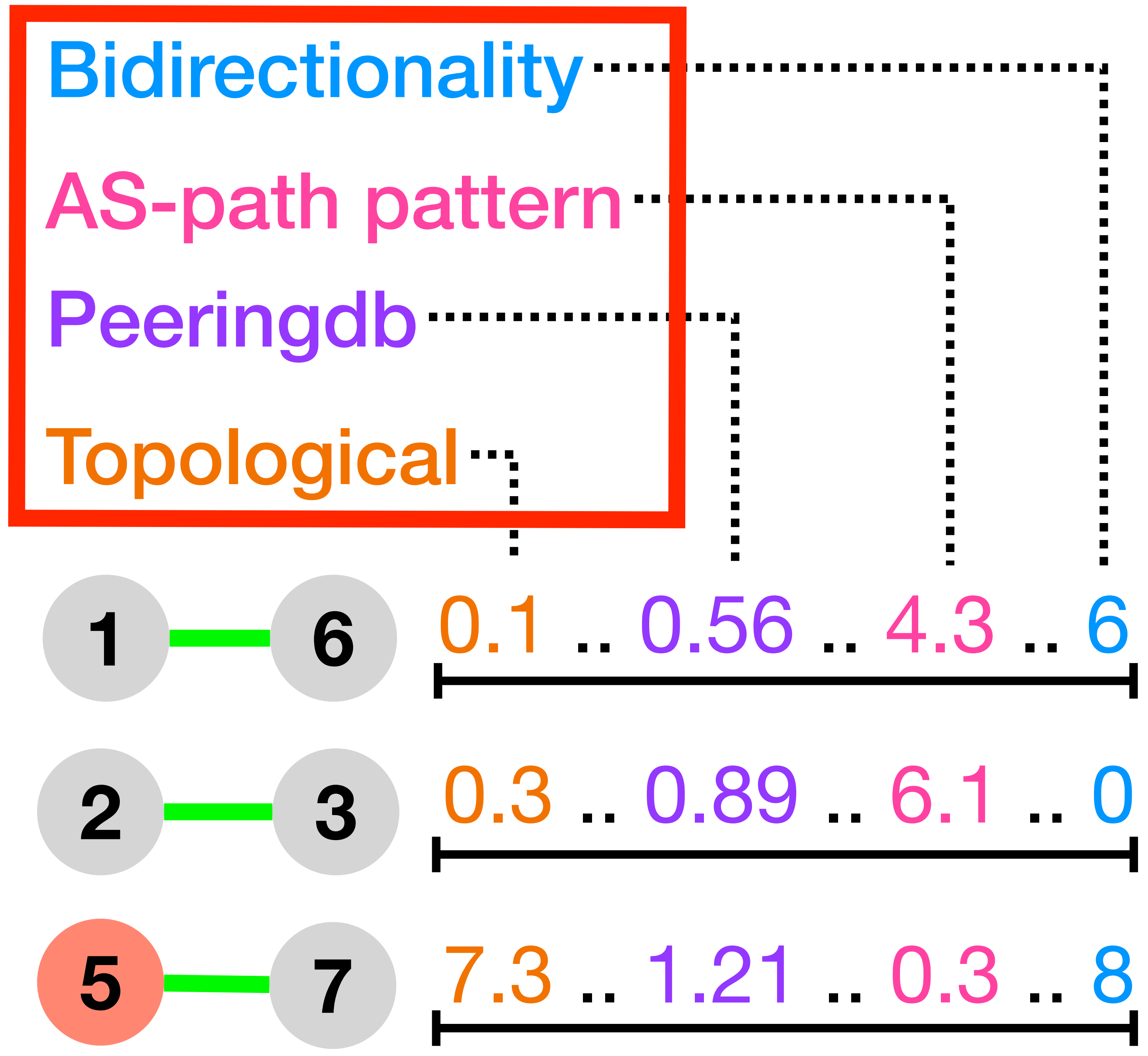
DFOH's fake AS links inference algorithm comprises three steps



Finding New Links → Computing Features → Inferring Hijacks



Domain-specific features that compensate each other



Feature vectors

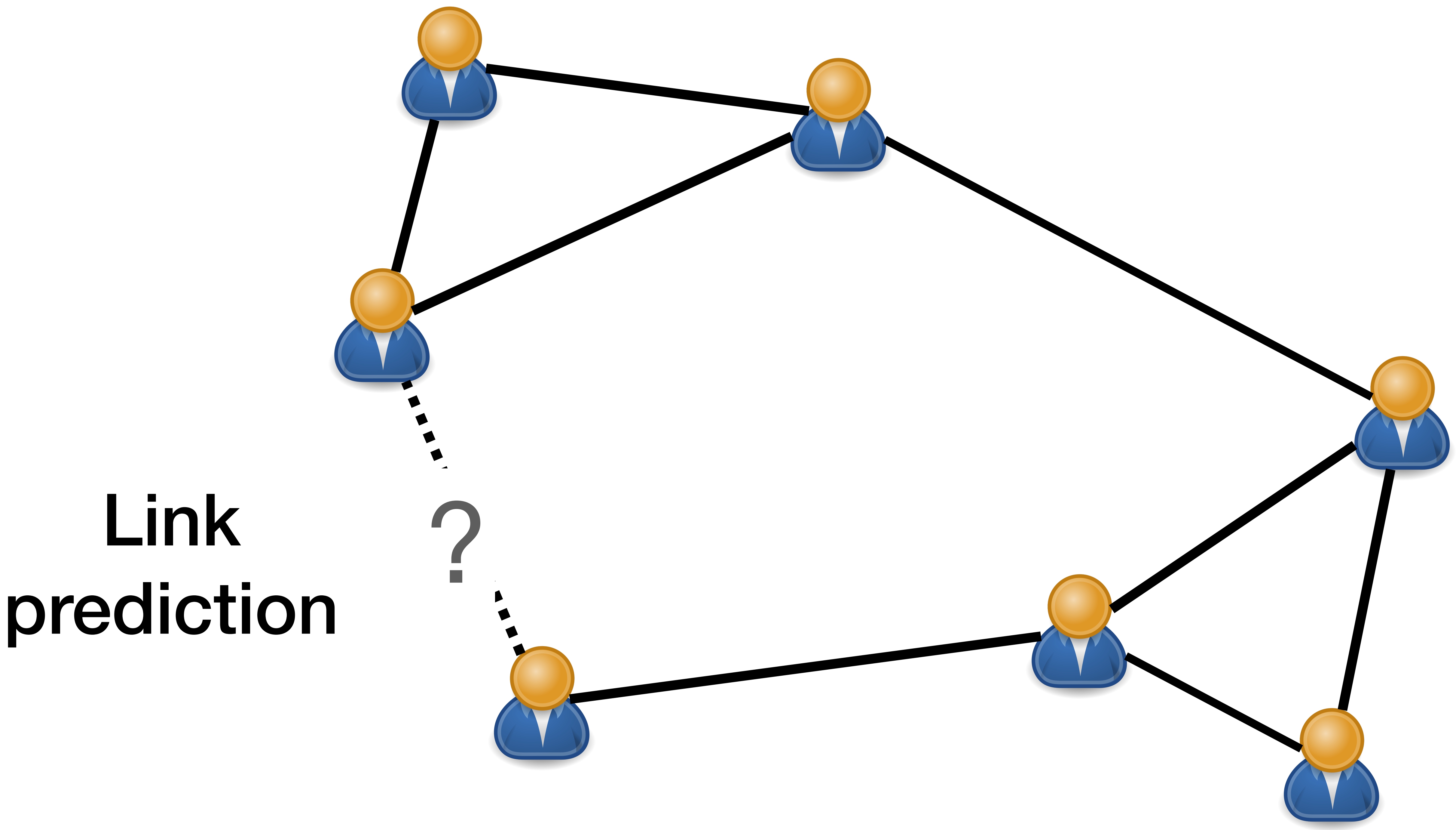
DFOH's fake AS links inference algorithm comprises three steps

Finding New Links → Computing Features → Inferring Hijacks

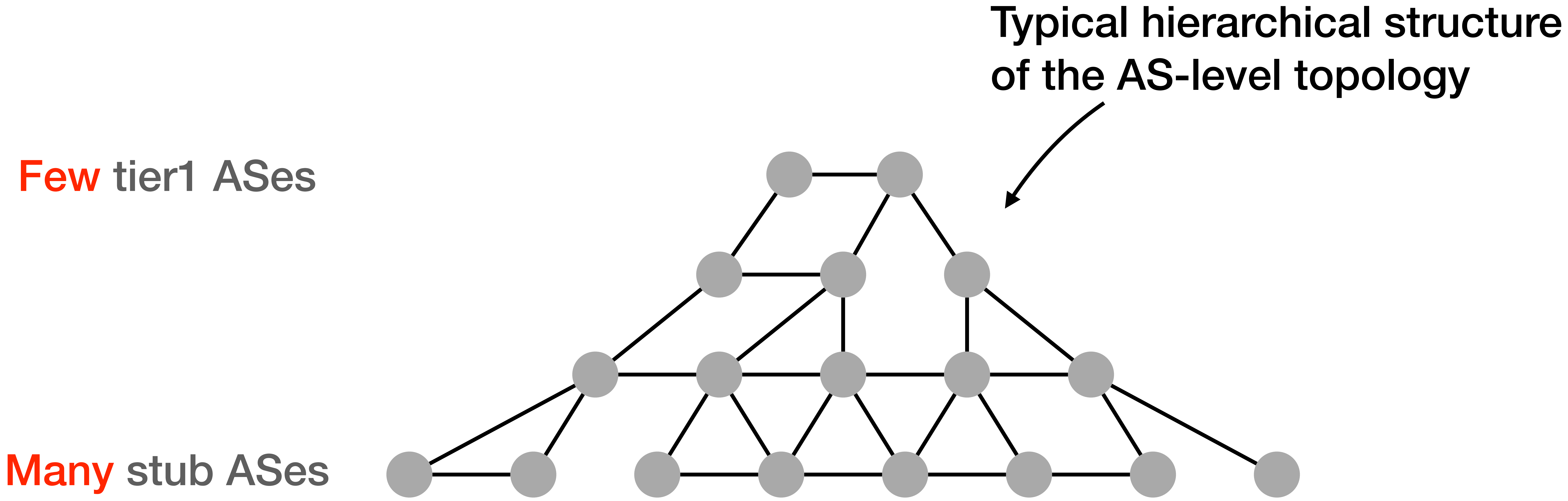
RIS/RouteViews Vantage point
1
2
3
4
5
6
7
Hijacker
Victim
new AS link
Feature categories:
Bidirectionality
AS-path pattern
Peeringdb
Topological
1 — 6 0.1 .. 0.56 .. 4.3 .. 6
2 — 3 0.3 .. 0.89 .. 6.1 .. 0
5 — 7 7.3 .. 1.21 .. 0.3 .. 8
Feature vectors
Random Forest
Inference
1 — 6 ✓
2 — 3 ✓
5 — 7 !
Training
Samples
1 — 2
2 — 4
3 — 6
Existing links
1 — 4
4 — 6
6 — 7
Nonexistent links

There are several link prediction frameworks

SEAL (NIPS'18) is one example

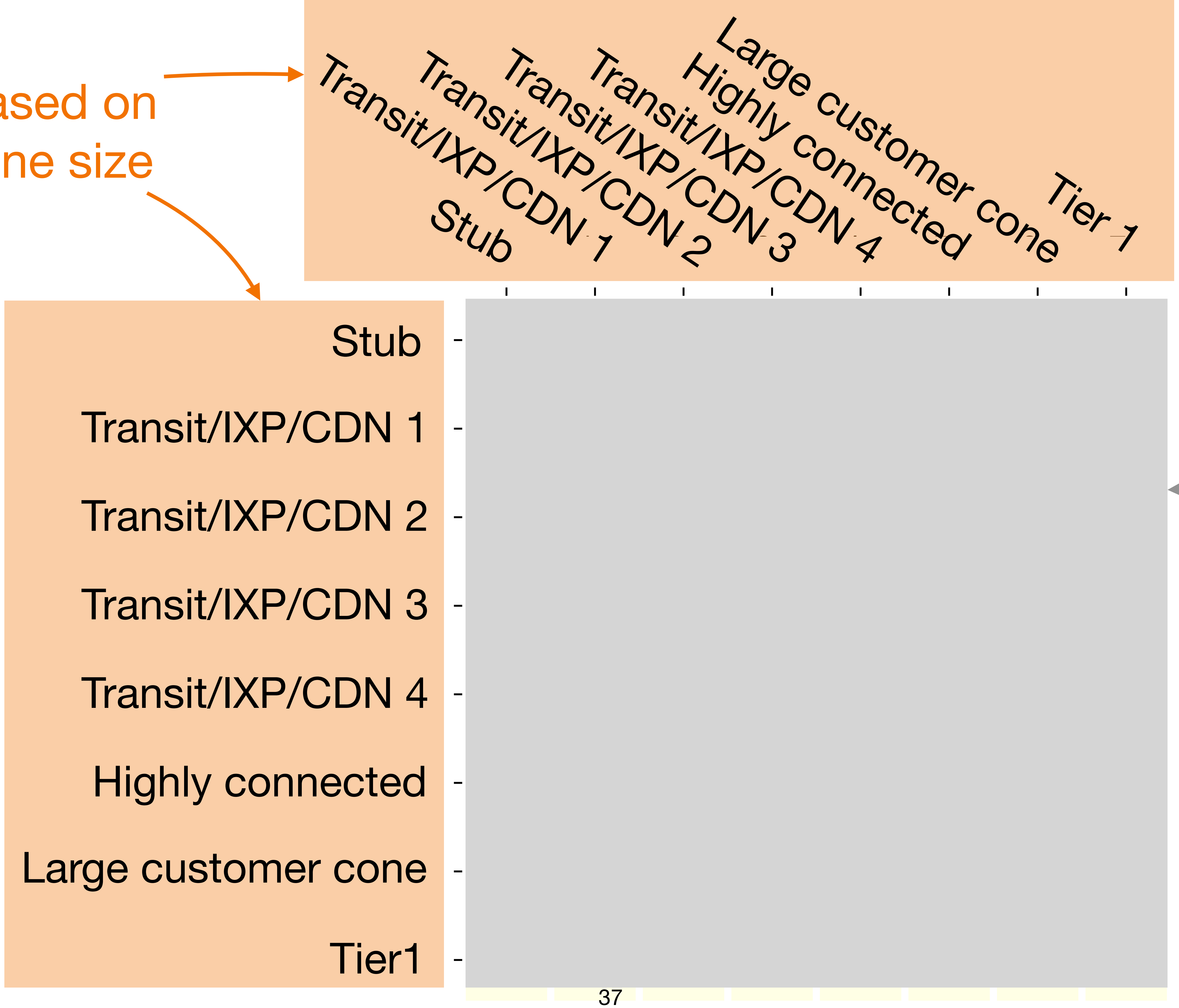


There are several link prediction frameworks
but they do not translate well for detecting fake AS links



Problem: randomly sampling **nonexistent** links makes DFOH skewed towards stub-to-stub links as they are overrepresented

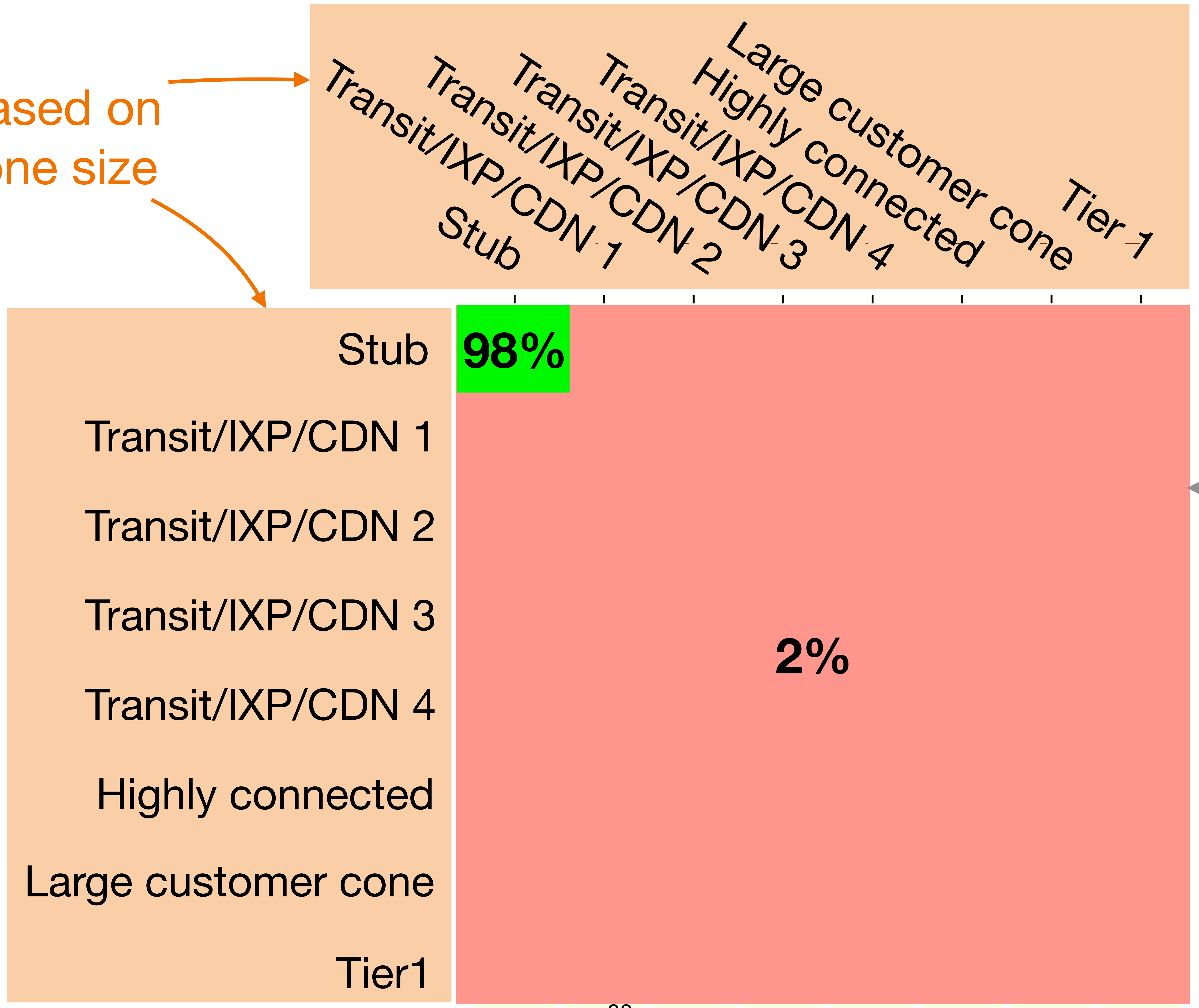
Clusters of ASes based on their degree and cone size



Proportion of sampled **nonexistent** AS links *(random sampling)*

Problem: randomly sampling **nonexistent** links makes DFOH skewed towards stub-to-stub links as they are overrepresented

Clusters of ASes based on their degree and cone size

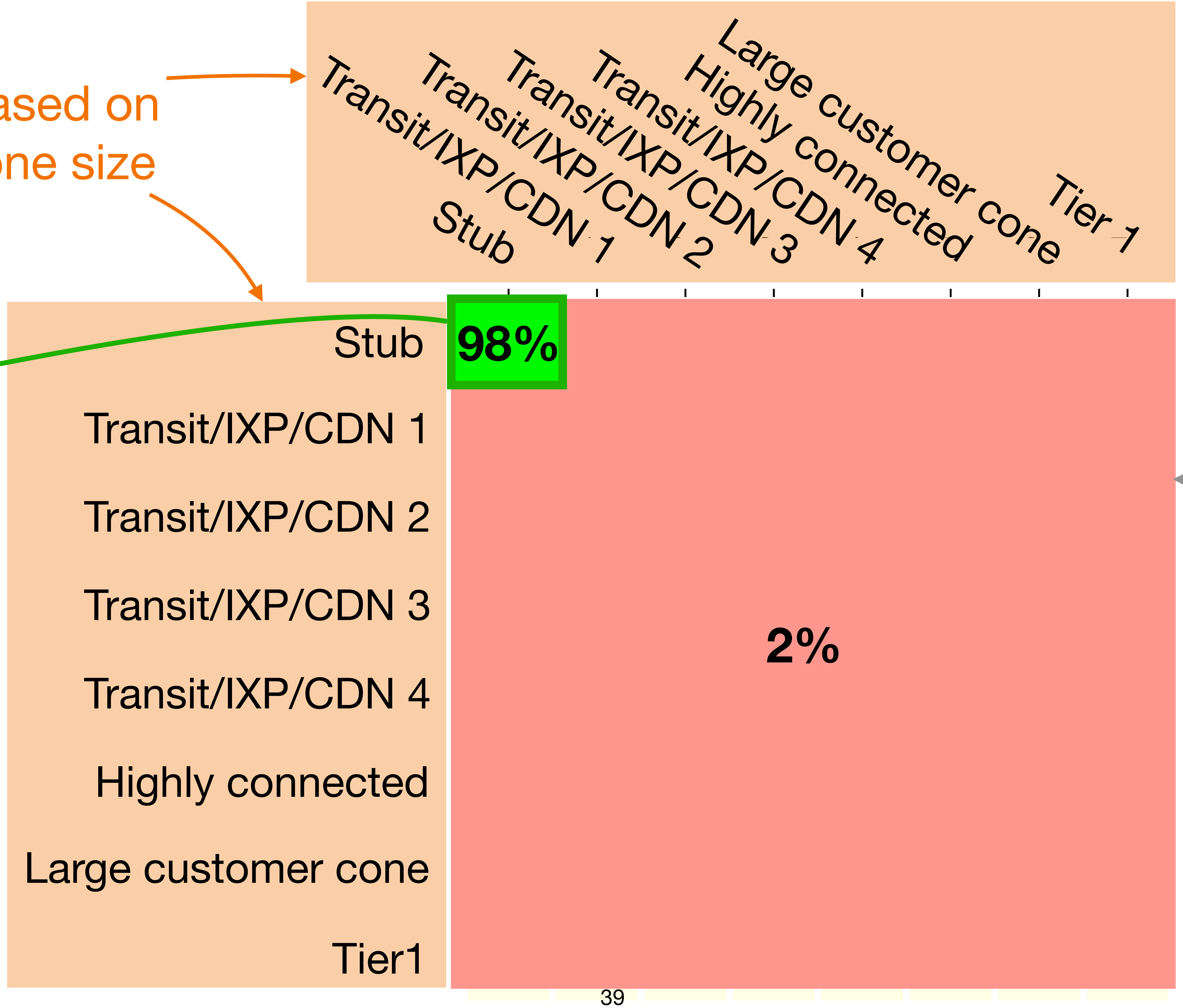


Proportion of sampled **nonexistent** AS links *(random sampling)*

Problem: randomly sampling **nonexistent** links makes DFOH skewed towards stub-to-stub links as they are overrepresented

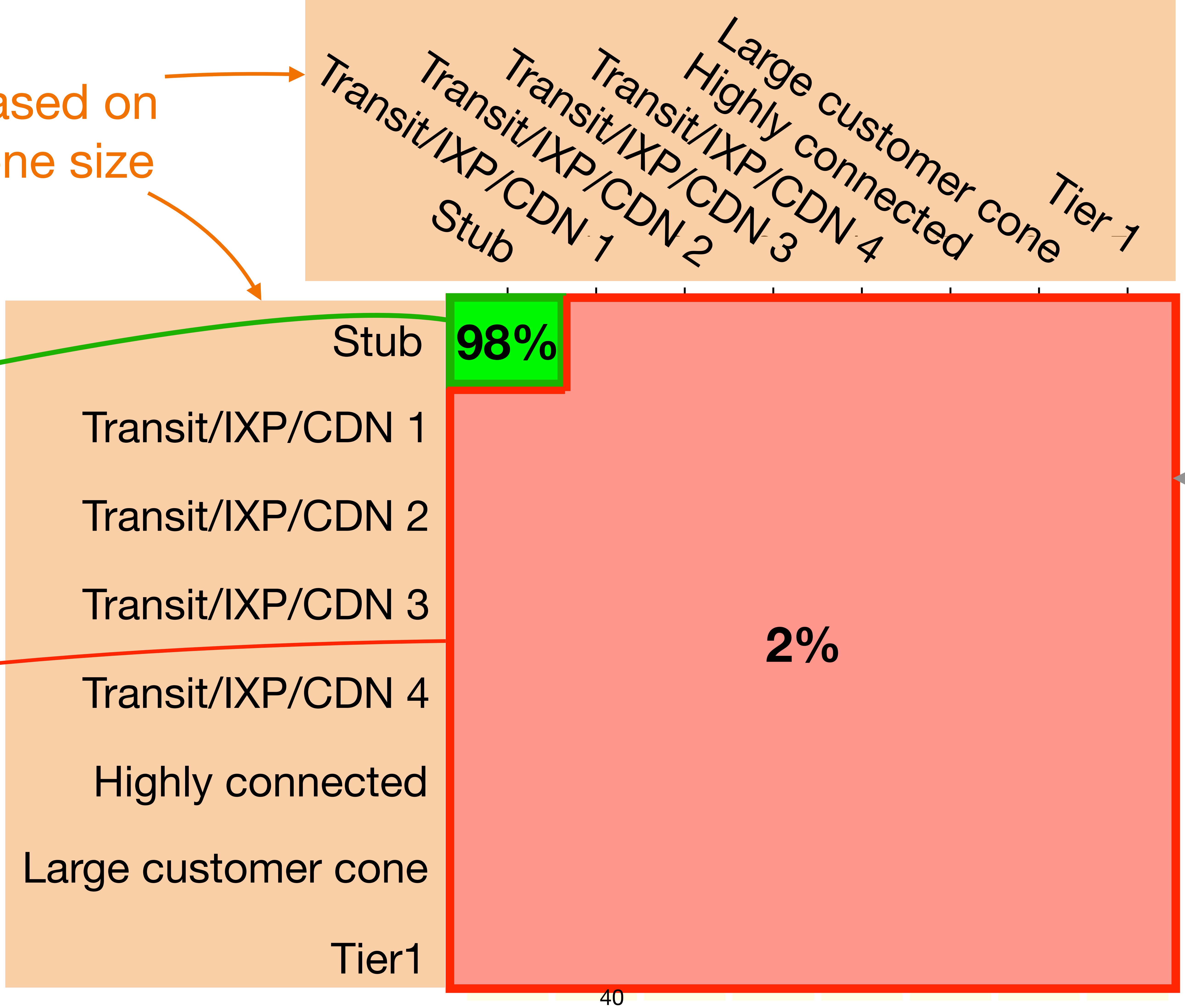
Clusters of ASes based on their degree and cone size

DFOH would perform well on scenarios involving two stubs



Proportion of sampled **nonexistent** AS links *(random sampling)*

Problem: randomly sampling **nonexistent** links makes DFOH skewed towards stub-to-stub links as they are overrepresented



Clusters of ASes based on their degree and cone size

DFOH would perform well on scenarios involving two stubs

But not on the other scenarios

Proportion of sampled **nonexistent** AS links *(random sampling)*

DFOH's fake AS links inference algorithm comprises three steps

Finding New Links

Computing Features

Inferring Hijacks

RIS/RouteViews Vantage point
1
2
3
4
5
6
7
Hijacker
Victim
new AS link
Feature categories:
Bidirectionality
AS-path pattern
Peeringdb
Topological
1 6
0.1 .. 0.56 .. 4.3 .. 6
2 3
0.3 .. 0.89 .. 6.1 .. 0
5 7
7.3 .. 1.21 .. 0.3 .. 8
Feature vectors
Random Forest
Inference
1 6
2 3
5 7
Training
Balanced sampling
Stub-Stub
Tier2-Stub
Tier1-Tier2
⋮
1 2
2 4
3 6
Existing links
1 4
4 6
6 7
Nonexistent links

Outline

DFOH's main challenge is to detect fake AS links

DFOH's inference pipeline relies on domain-specific knowledge and a tailored link prediction framework

***DFOH*'s inferences are accurate** **in every attack scenario**

DFOH is up and running

We evaluate ***DFOH*** on artificially created forged-origin hijacks as there is no ground truth at scale

Methodology:

We take existing AS paths
and prepend a new origin to create a new link

We take 9k cases where the new link exists (*legitimate* or “*negative*” *cases*)
and 9k cases where the new link does not exist (*suspicious or “positive” cases*)

We evaluate ***DFOH*** on artificially created forged-origin hijacks as there is no ground truth at scale

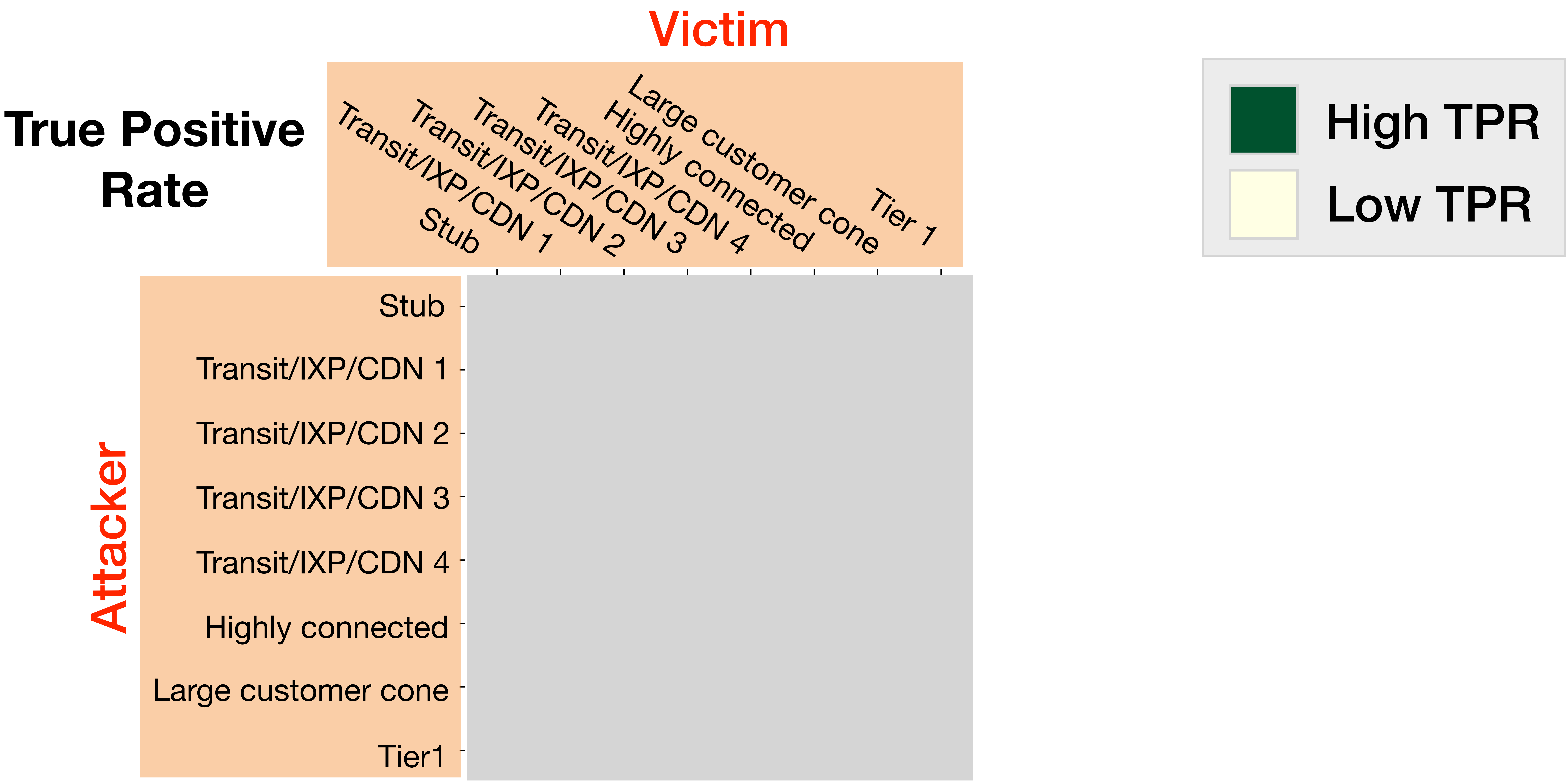
Methodology:

We take existing AS paths
and prepend a new origin to create a new link

We take 9k cases where the new link exists (*legitimate* or “*negative” cases*)
and 9k cases where the new link does not exist (*suspicious or “positive” cases*)

🎯 We focus on the **True Positive Rate** (TPR)
and the **False Positive Rate** (FPR)

DFOH **is accurate upon every attack scenario**

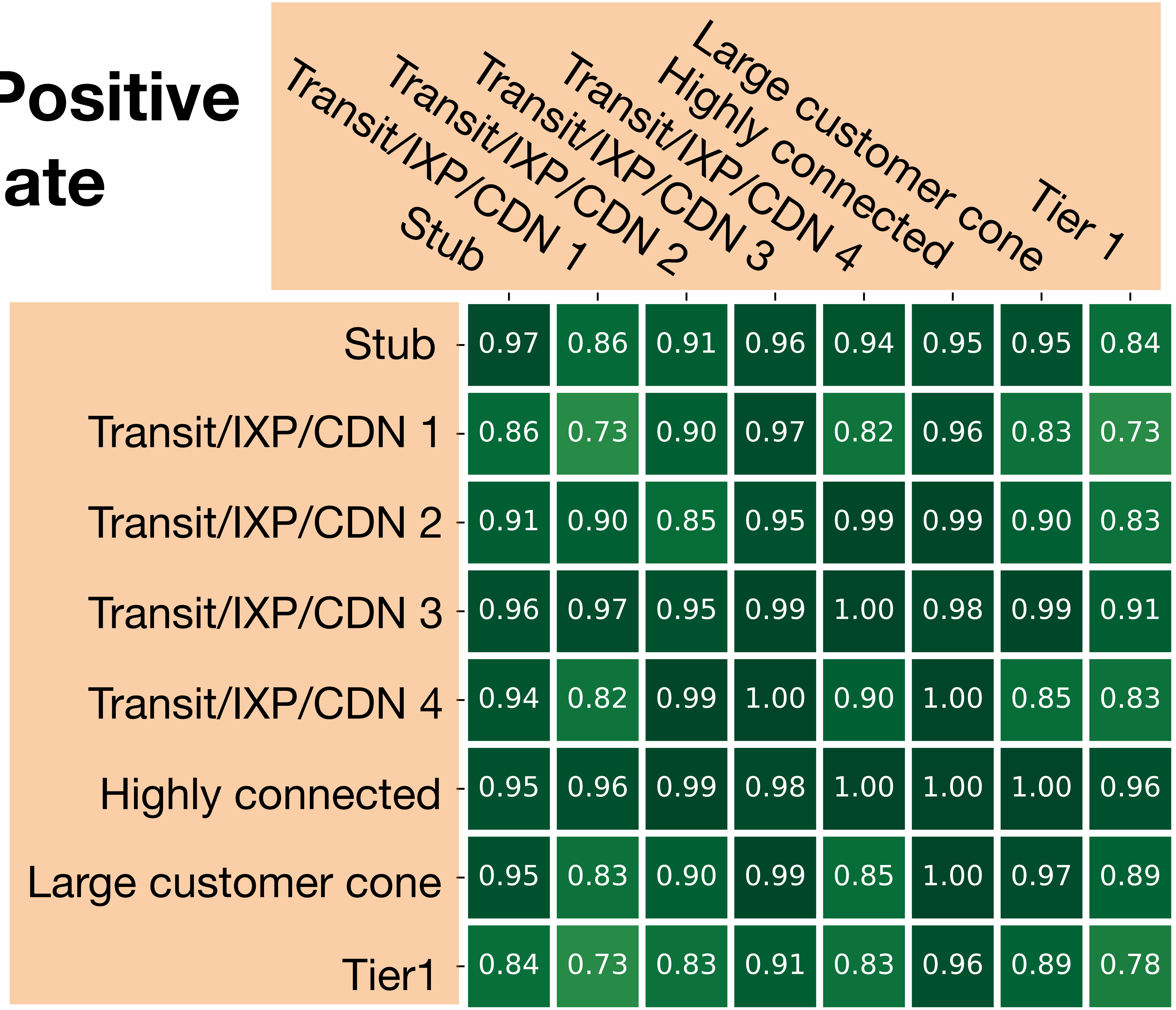


DFOH is accurate upon every attack scenario

True Positive Rate

Victim

Attacker



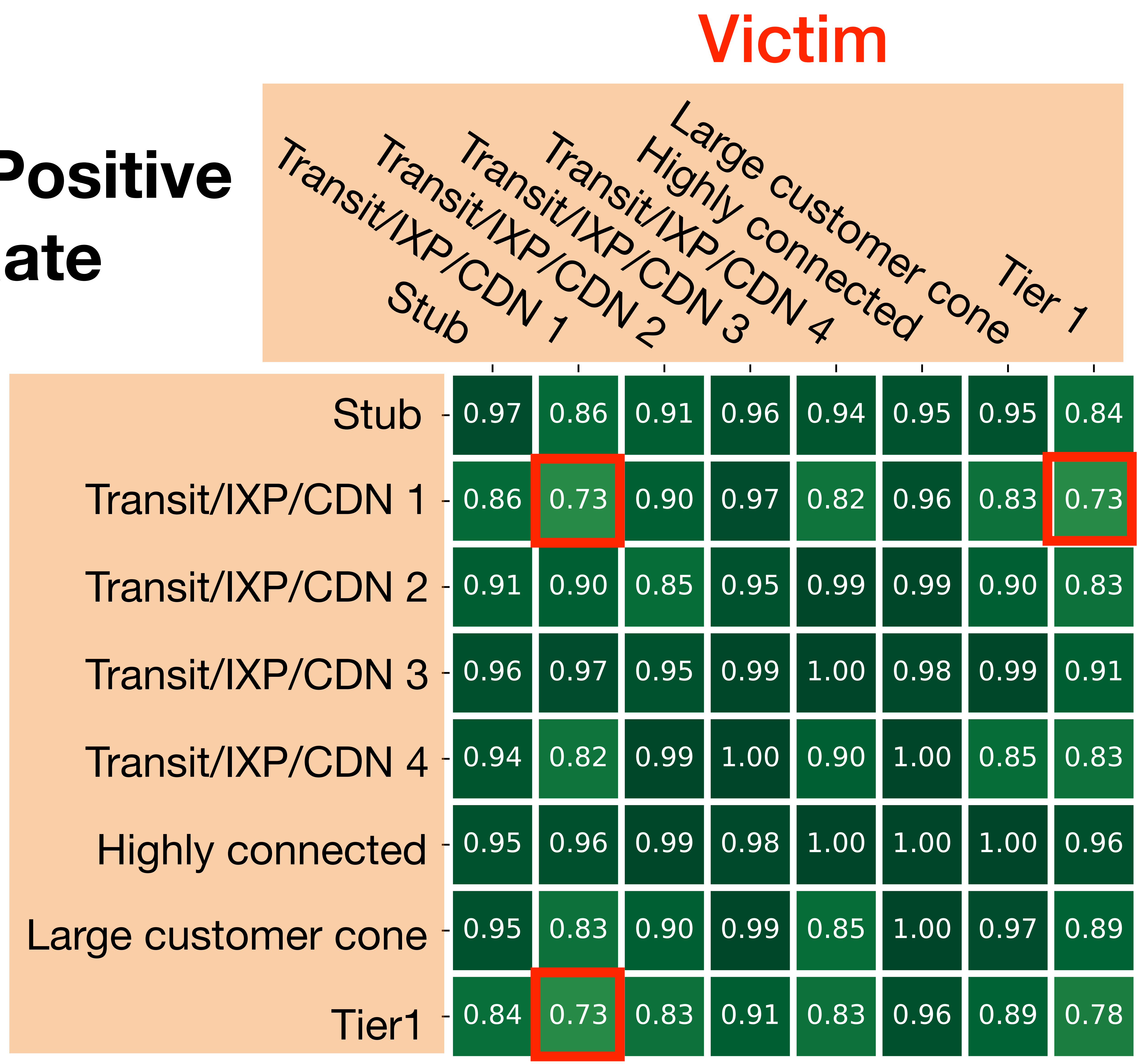
Attacker \ Victim	Stub	Transit/IXP/CDN 1	Transit/IXP/CDN 2	Transit/IXP/CDN 3	Transit/IXP/CDN 4	Highly connected	Large customer cone	Tier 1
Stub	0.97	0.86	0.91	0.96	0.94	0.95	0.95	0.84
Transit/IXP/CDN 1	0.86	0.73	0.90	0.97	0.82	0.96	0.83	0.73
Transit/IXP/CDN 2	0.91	0.90	0.85	0.95	0.99	0.99	0.90	0.83
Transit/IXP/CDN 3	0.96	0.97	0.95	0.99	1.00	0.98	0.99	0.91
Transit/IXP/CDN 4	0.94	0.82	0.99	1.00	0.90	1.00	0.85	0.83
Highly connected	0.95	0.96	0.99	0.98	1.00	1.00	1.00	0.96
Large customer cone	0.95	0.83	0.90	0.99	0.85	1.00	0.97	0.89
Tier1	0.84	0.73	0.83	0.91	0.83	0.96	0.89	0.78

🟩 High TPR

⬜ Low TPR

DFOH is accurate upon every attack scenario

True Positive Rate



Victim

Attacker

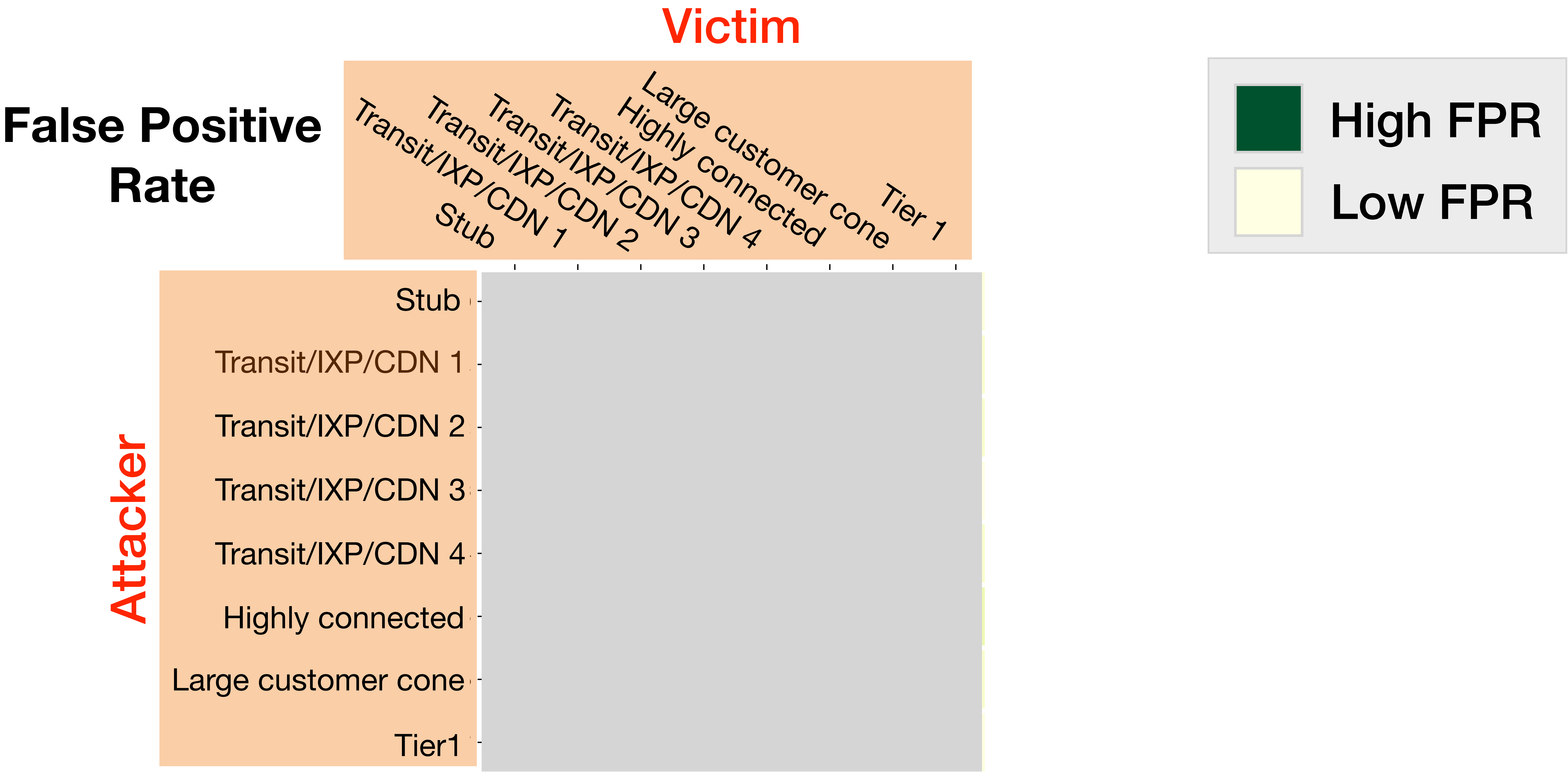
	Stub	Transit/IXP/CDN 1	Transit/IXP/CDN 2	Transit/IXP/CDN 3	Transit/IXP/CDN 4	Highly connected	Large customer cone	Tier 1
Stub	0.97	0.86	0.91	0.96	0.94	0.95	0.95	0.84
Transit/IXP/CDN 1	0.86	0.73	0.90	0.97	0.82	0.96	0.83	0.73
Transit/IXP/CDN 2	0.91	0.90	0.85	0.95	0.99	0.99	0.90	0.83
Transit/IXP/CDN 3	0.96	0.97	0.95	0.99	1.00	0.98	0.99	0.91
Transit/IXP/CDN 4	0.94	0.82	0.99	1.00	0.90	1.00	0.85	0.83
Highly connected	0.95	0.96	0.99	0.98	1.00	1.00	1.00	0.96
Large customer cone	0.95	0.83	0.90	0.99	0.85	1.00	0.97	0.89
Tier1	0.84	0.73	0.83	0.91	0.83	0.96	0.89	0.78

🟩 High TPR

⬜ Low TPR

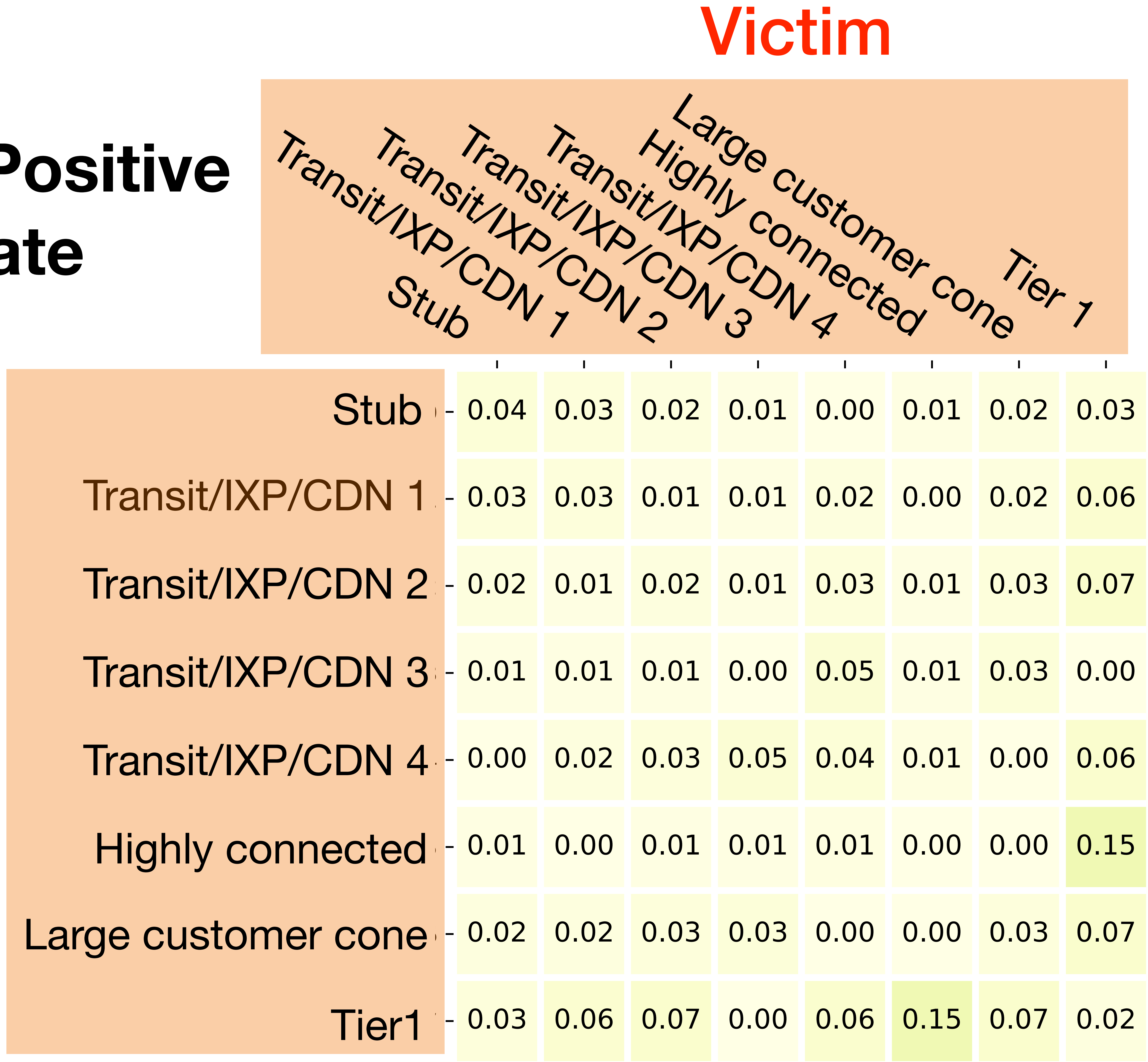
The minimum TPR is 0.73

DFOH is accurate upon every attack scenario



DFOH is accurate upon every attack scenario

False Positive Rate



Attacker \ Victim	Stub	Transit/IXP/CDN 1	Transit/IXP/CDN 2	Transit/IXP/CDN 3	Transit/IXP/CDN 4	Highly connected	Large customer cone	Tier 1
Stub	0.04	0.03	0.02	0.01	0.00	0.01	0.02	0.03
Transit/IXP/CDN 1	0.03	0.03	0.01	0.01	0.02	0.00	0.02	0.06
Transit/IXP/CDN 2	0.02	0.01	0.02	0.01	0.03	0.01	0.03	0.07
Transit/IXP/CDN 3	0.01	0.01	0.01	0.00	0.05	0.01	0.03	0.00
Transit/IXP/CDN 4	0.00	0.02	0.03	0.05	0.04	0.01	0.00	0.06
Highly connected	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.15
Large customer cone	0.02	0.02	0.03	0.03	0.00	0.00	0.03	0.07
Tier1	0.03	0.06	0.07	0.00	0.06	0.15	0.07	0.02

🟩 High FPR

□ Low FPR

DFOH is accurate upon every attack scenario

False Positive Rate

Victim

Attacker

Attacker \ Victim	Stub	Transit/IXP/CDN 1	Transit/IXP/CDN 2	Transit/IXP/CDN 3	Transit/IXP/CDN 4	Highly connected	Large customer cone	Tier 1
Stub	0.04	0.03	0.02	0.01	0.00	0.01	0.02	0.03
Transit/IXP/CDN 1	0.03	0.03	0.01	0.01	0.02	0.00	0.02	0.06
Transit/IXP/CDN 2	0.02	0.01	0.02	0.01	0.03	0.01	0.03	0.07
Transit/IXP/CDN 3	0.01	0.01	0.01	0.00	0.05	0.01	0.03	0.00
Transit/IXP/CDN 4	0.00	0.02	0.03	0.05	0.04	0.01	0.00	0.06
Highly connected	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.15
Large customer cone	0.02	0.02	0.03	0.03	0.00	0.00	0.03	0.07
Tier1	0.03	0.06	0.07	0.00	0.06	0.15	0.07	0.02

🟩 High FPR

⬜ Low FPR

The maximum FPR is 0.15

Outline

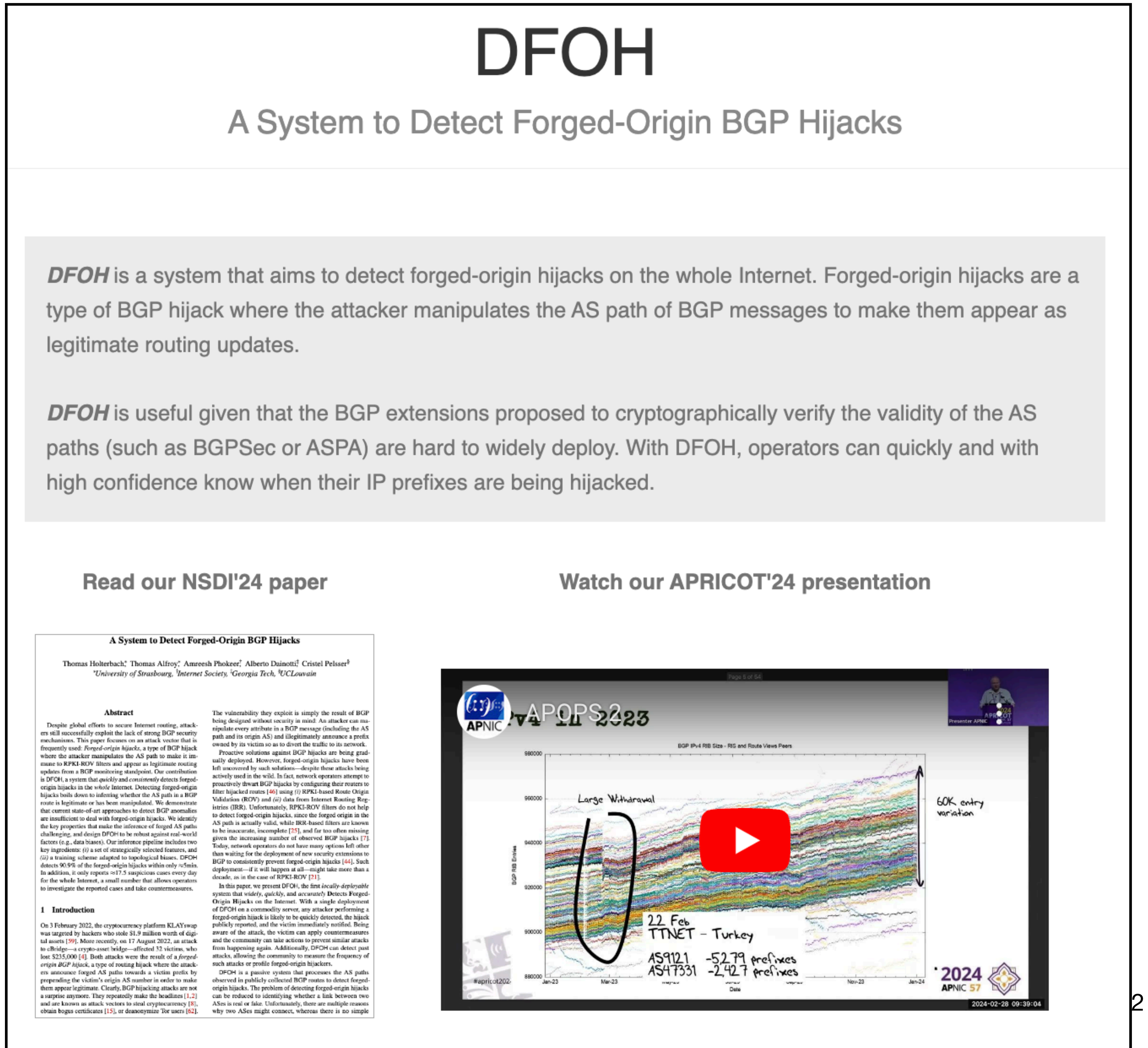
DFOH’s main challenge is to detect fake AS links

DFOH’s inference pipeline discriminates fake AS links from the real ones

DFOH’s inferences are accurate in every attack scenario

DFOH is **up and running** and useful for operators

DFOH runs at https://dfoh.uclouvain.be



DFOH

A System to Detect Forged-Origin BGP Hijacks

DFOH is a system that aims to detect forged-origin hijacks on the whole Internet. Forged-origin hijacks are a type of BGP hijack where the attacker manipulates the AS path of BGP messages to make them appear as legitimate routing updates.

DFOH is useful given that the BGP extensions proposed to cryptographically verify the validity of the AS paths (such as BGPSec or ASPA) are hard to widely deploy. With DFOH, operators can quickly and with high confidence know when their IP prefixes are being hijacked.

Read our NSDI'24 paper

Watch our APRICOT'24 presentation

A System to Detect Forged-Origin BGP Hijacks

Thomas Holterbach, Thomas Alfroy, Amreesh Phokeer, Alberto Dainotti, Cristel Pelsser

DFOH provides past and real-time forged-origin BGP hijacks detection

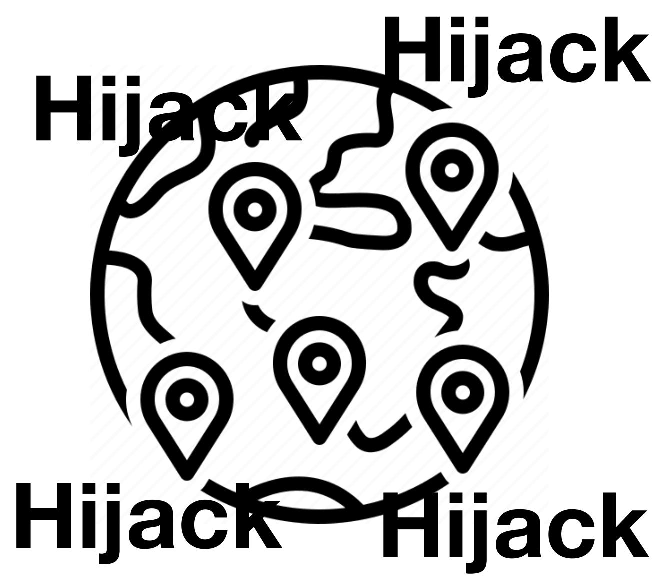
DFOH is useful and practical for network operators

Useful: DFOH detects the two known forged-origin BGP hijacks
(the klayswap and cbridge attacks)

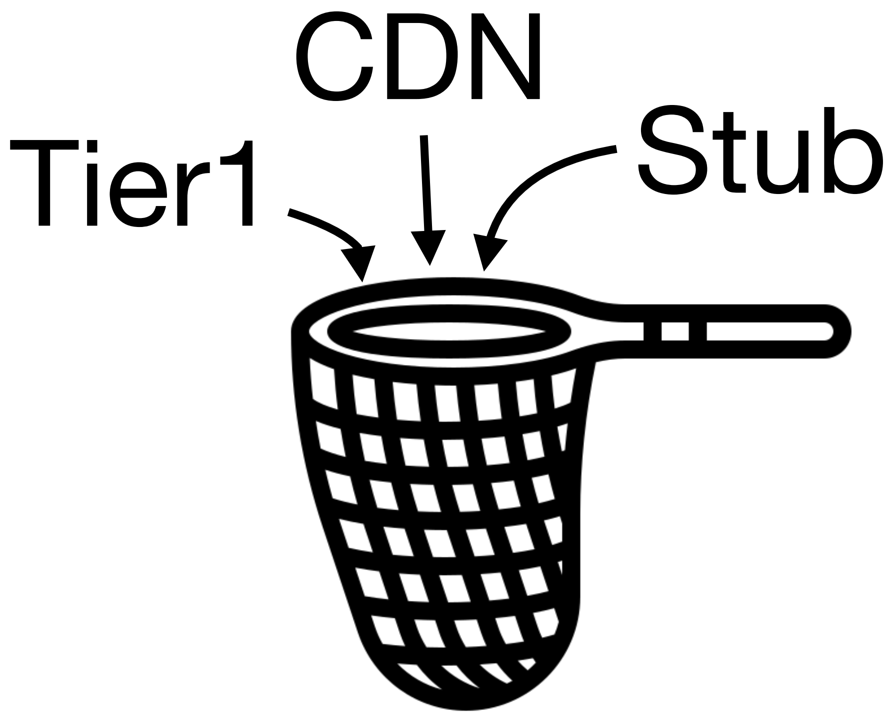
Practical: DFOH only reports zero or one case every month for 99.8% of the ASes
(worse case is 15 cases)

DFOH: A System to Detect Forged-Origin Hijacks

DFOH runs in a commodity server



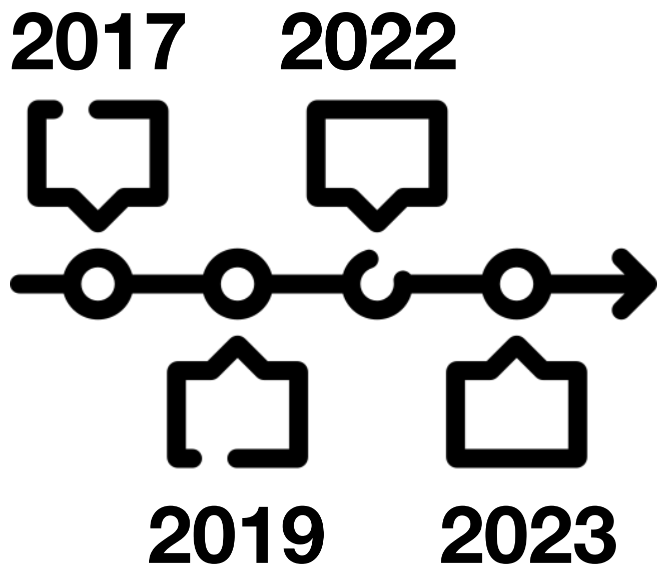
DFOH detects hijacks on the whole Internet



DFOH is accurate in every attack scenario

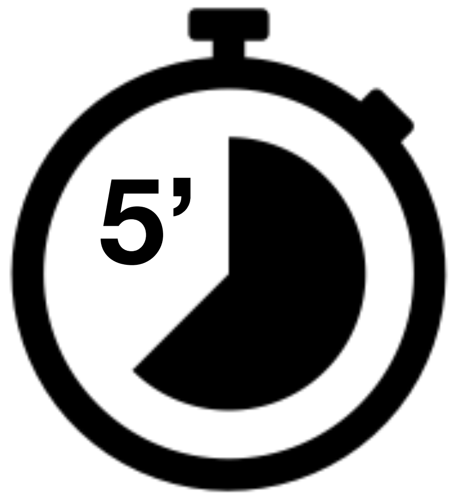
DFOH: A System to Detect Forged-Origin BGP Hijacks

DFOH **runs in a commodity server**

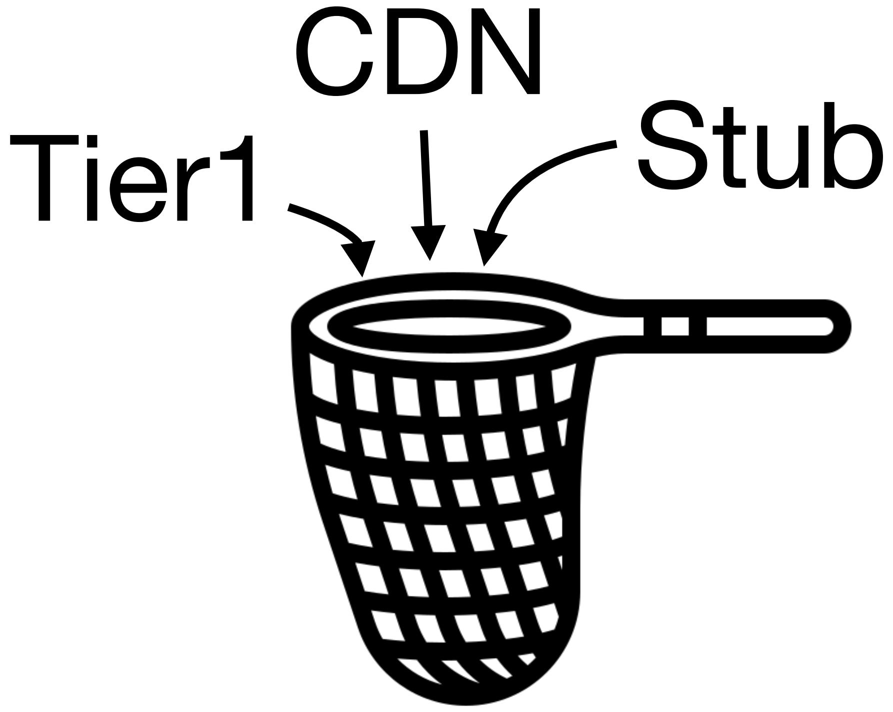


DFOH **detects past hijacks**

DFOH **detects hijacks on the whole Internet**



DFOH **provides near-real-time detection**



DFOH **is accurate in every attack scenario**

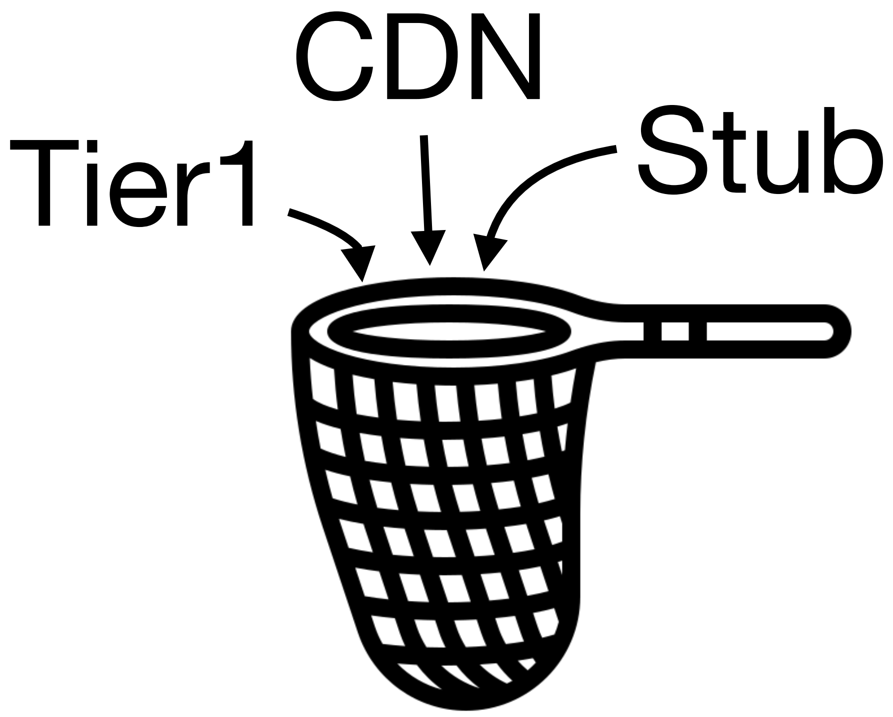
DFOH **is robust against adversarial inputs**

DFOH: A System to Detect Forged-Origin Hijacks

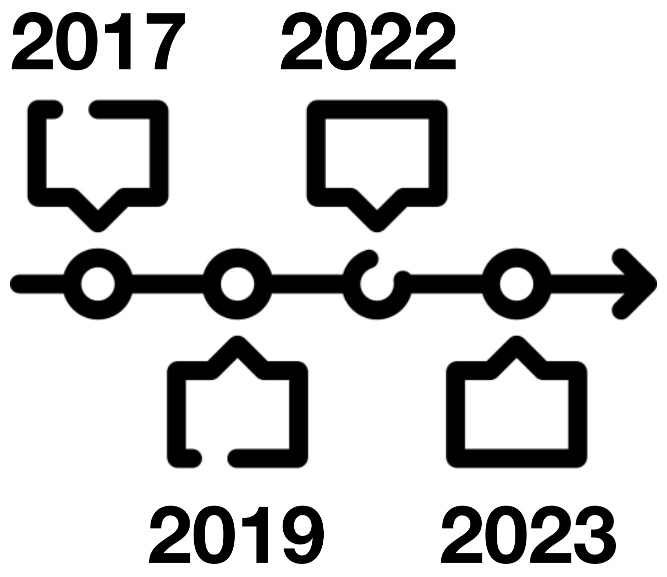
https://dfoh.uclouvain.be

DFOH runs in a commodity server

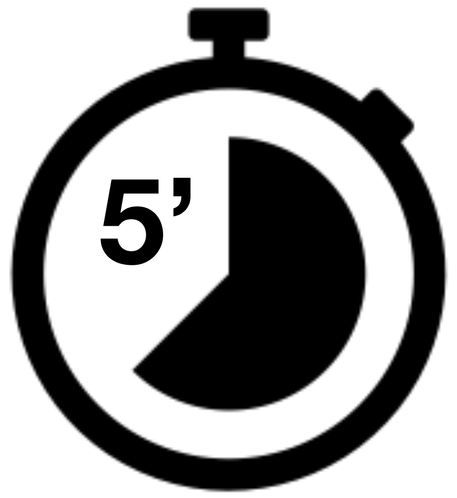
DFOH detects hijacks on the whole Internet



DFOH is accurate in every attack scenario



DFOH detects past hijacks



DFOH provides near-real-time detection

DFOH is robust against adversarial inputs