

A System to Detect Forged-Origin BGP Hijacks

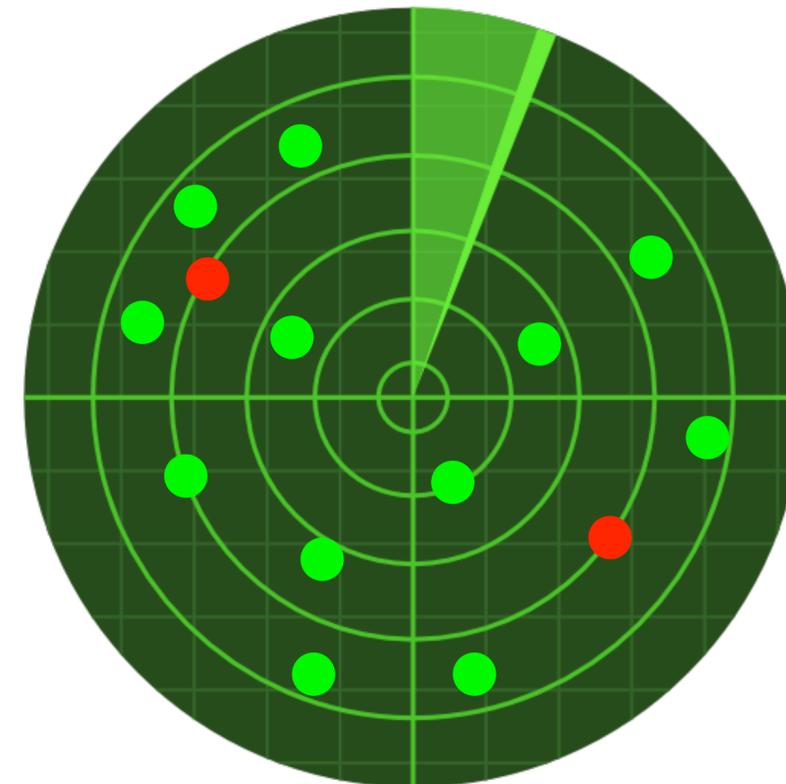
Thomas Holterbach
University of Strasbourg

USENIX NSDI'24
Thursday, April 18

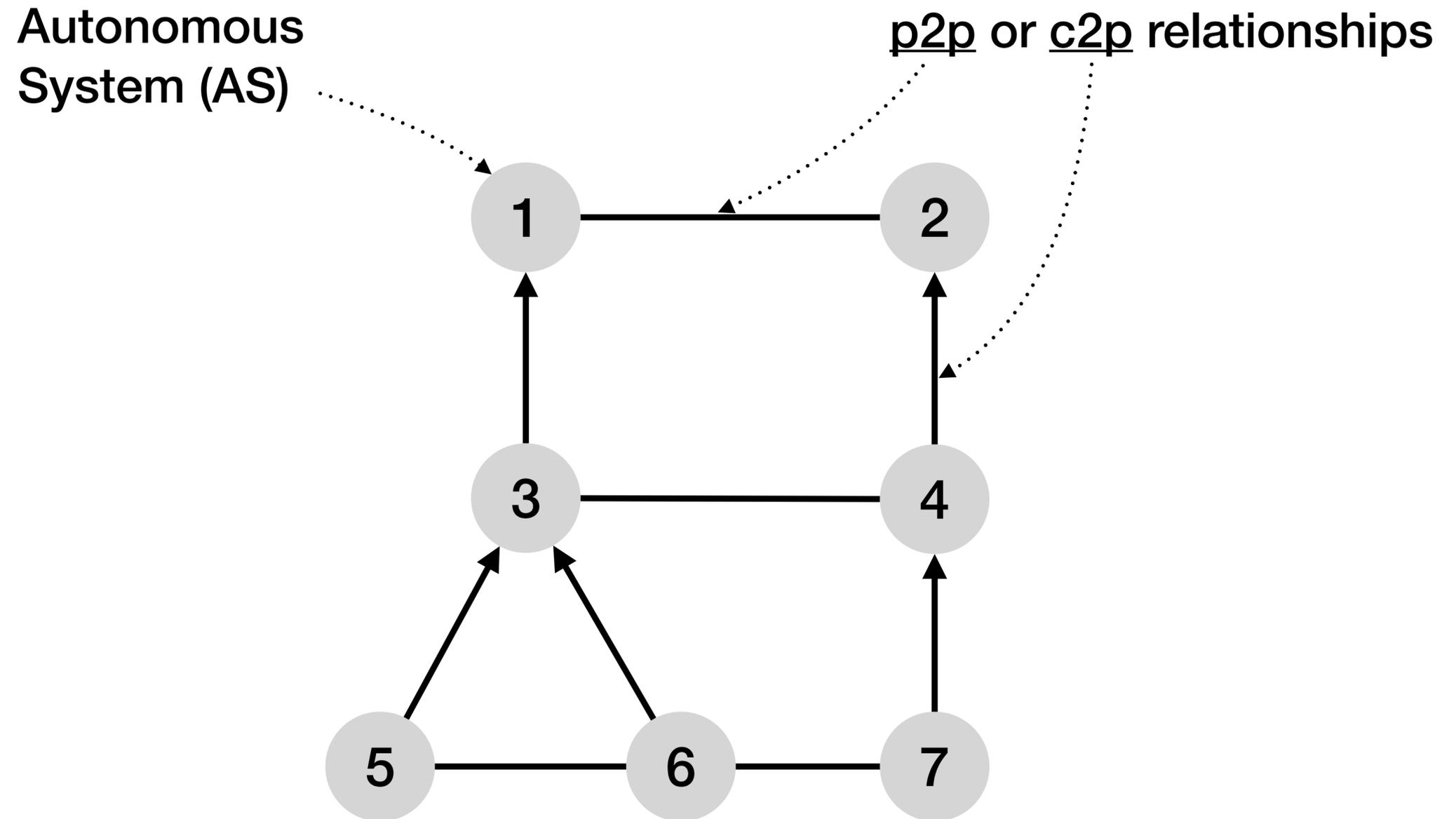
Joint work with:

Thomas Alfroy
Amreesh D. Phokeer

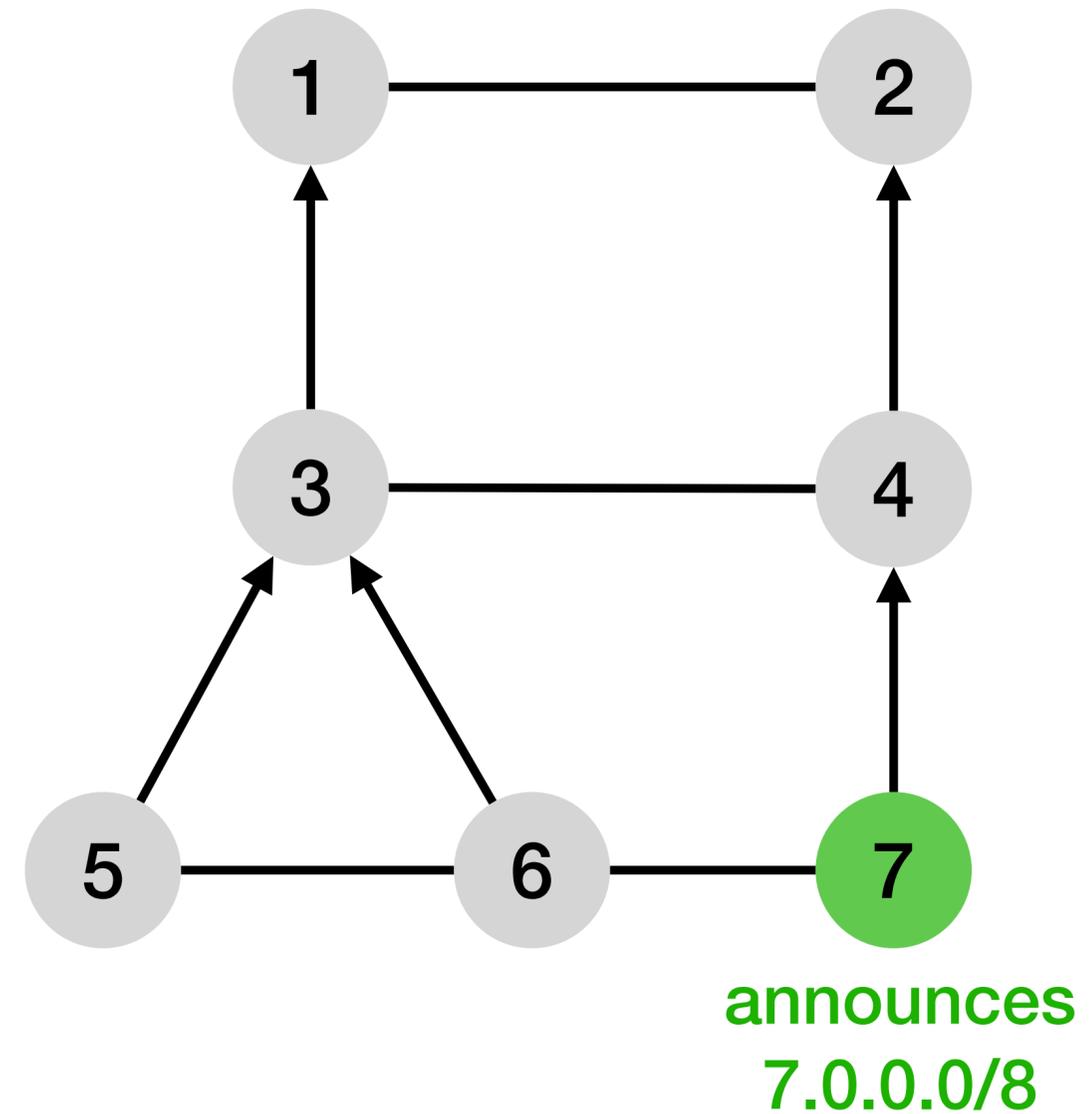
Alberto Dainotti
Cristel Pelsser



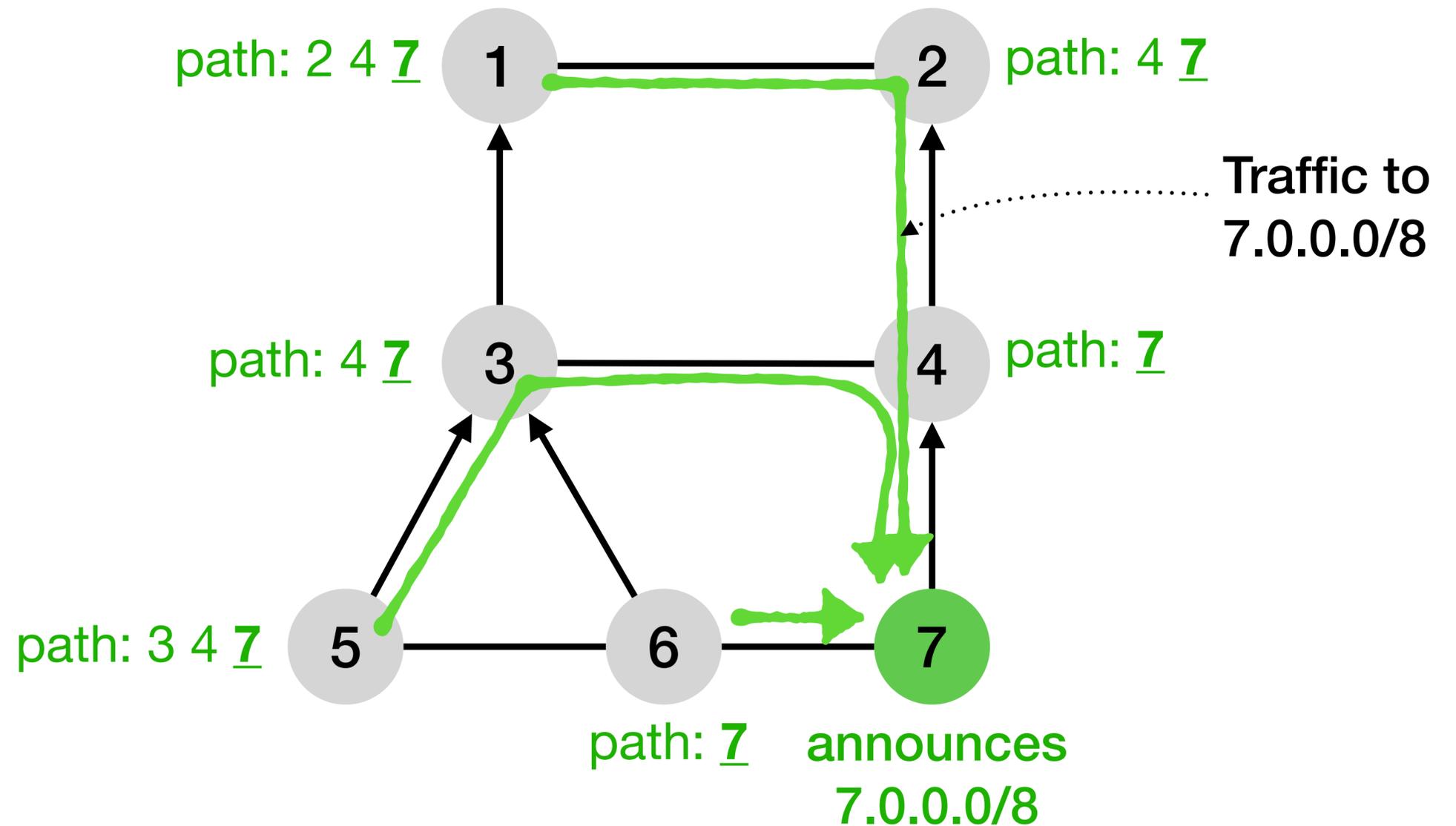
Internet routing (BGP) is vulnerable to traffic hijacking



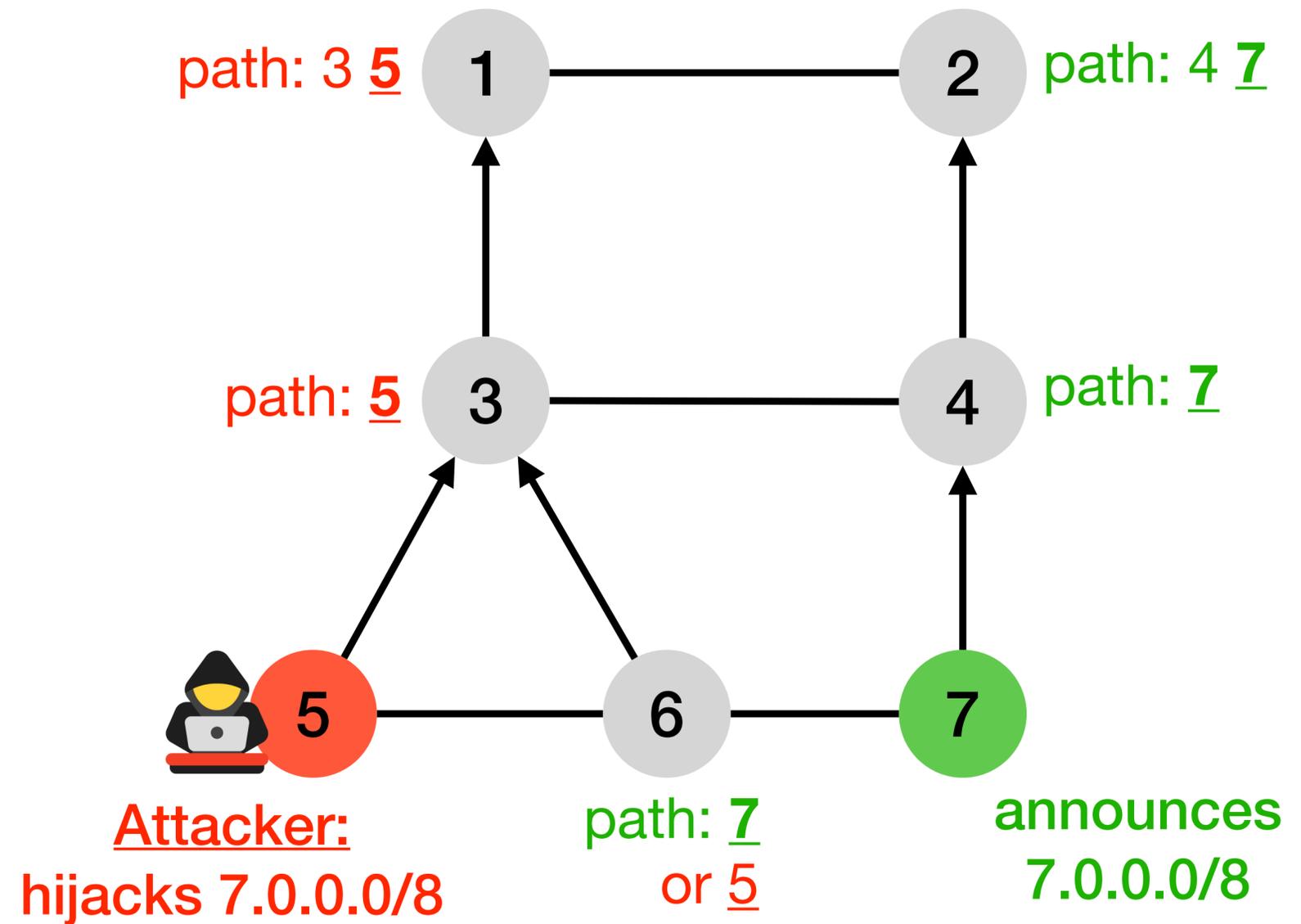
Internet routing (BGP) is vulnerable to traffic hijacking



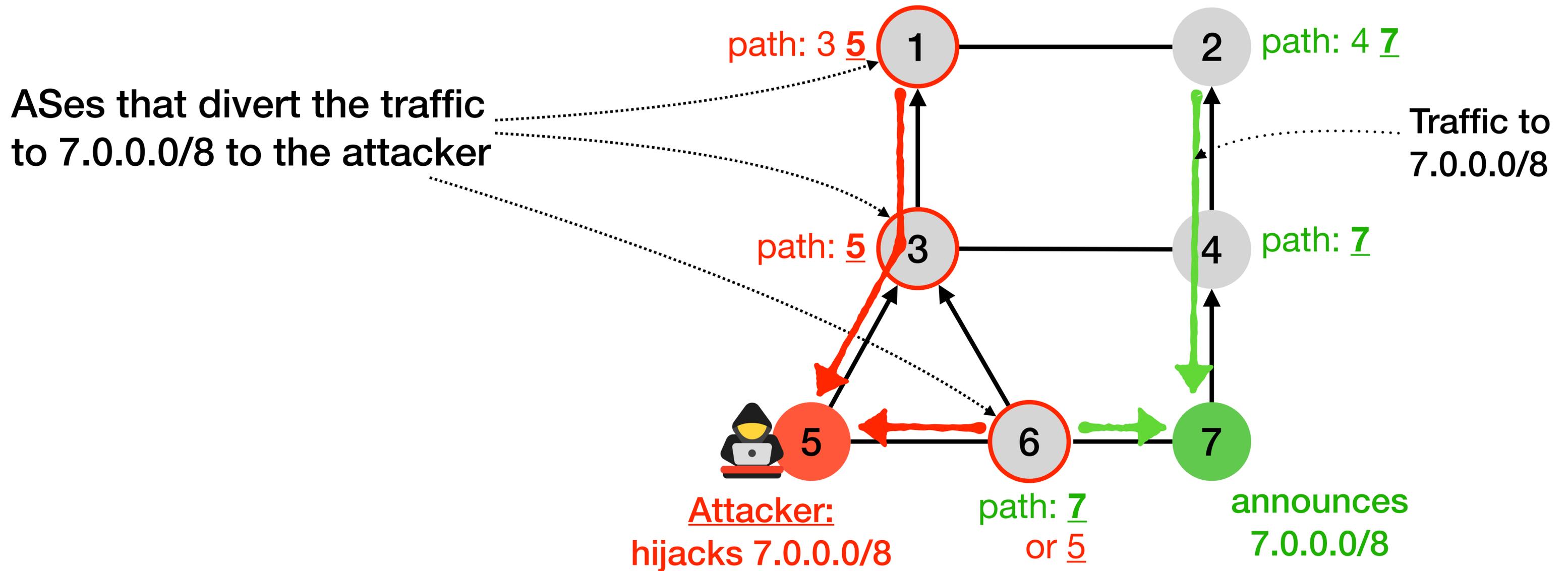
Internet routing (BGP) is vulnerable to traffic hijacking



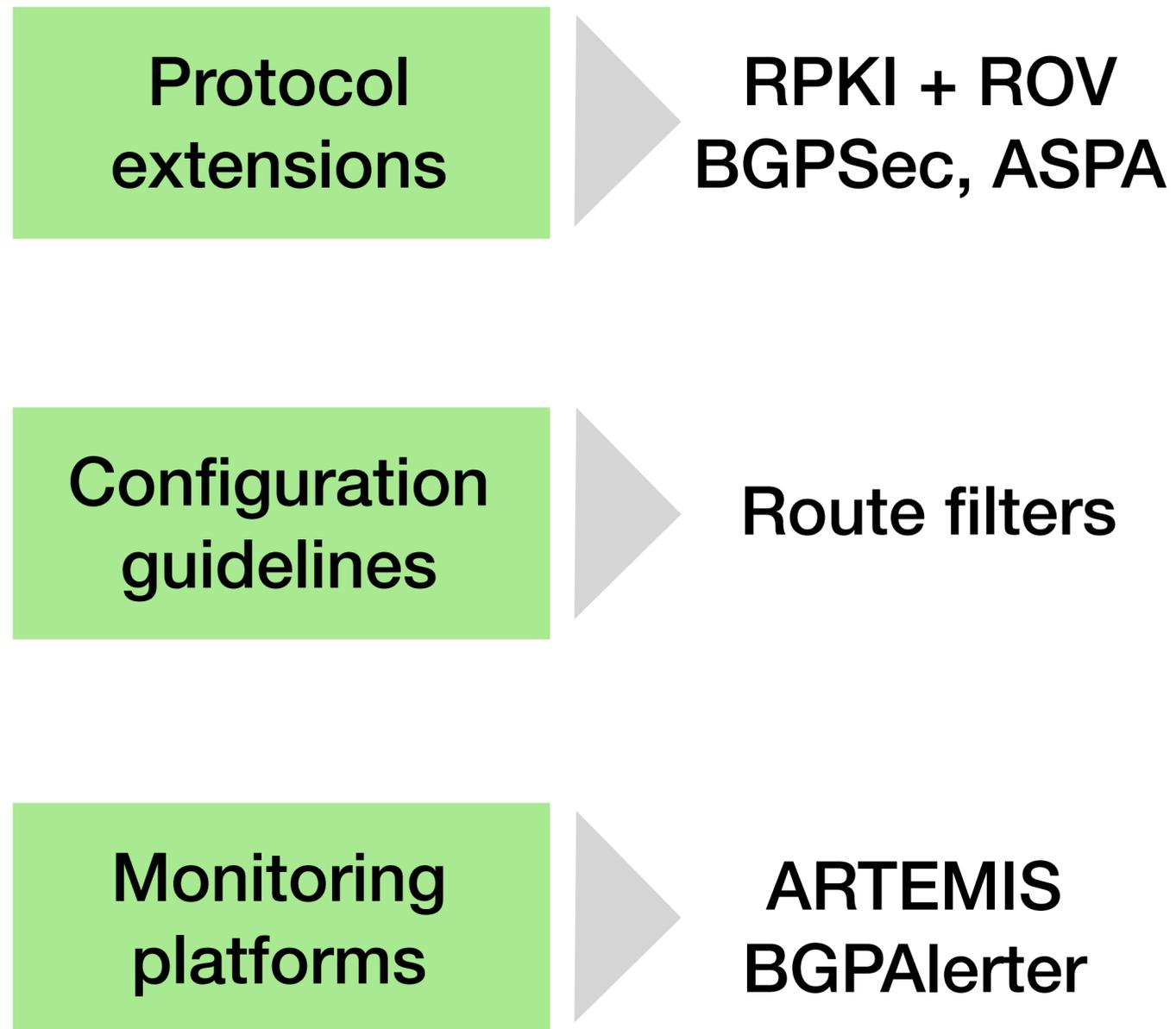
Internet routing (BGP) is vulnerable to traffic hijacking



Internet routing (BGP) is vulnerable to traffic hijacking

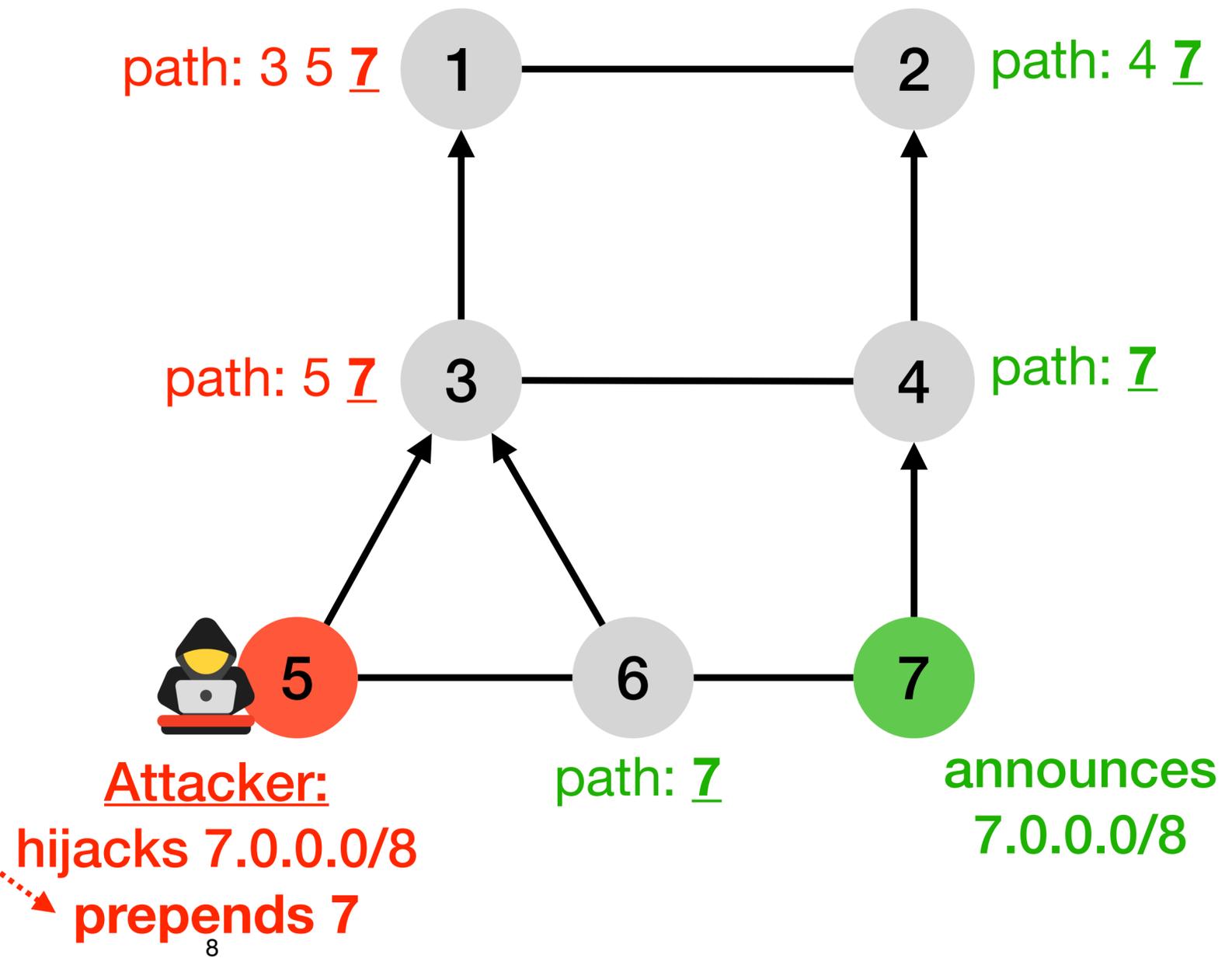


Fortunately, there are defenses against BGP hijacking



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

The attacker prepends the legitimate AS number to the AS path



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

August 17, 2022

The Record.
Recorded Future® News

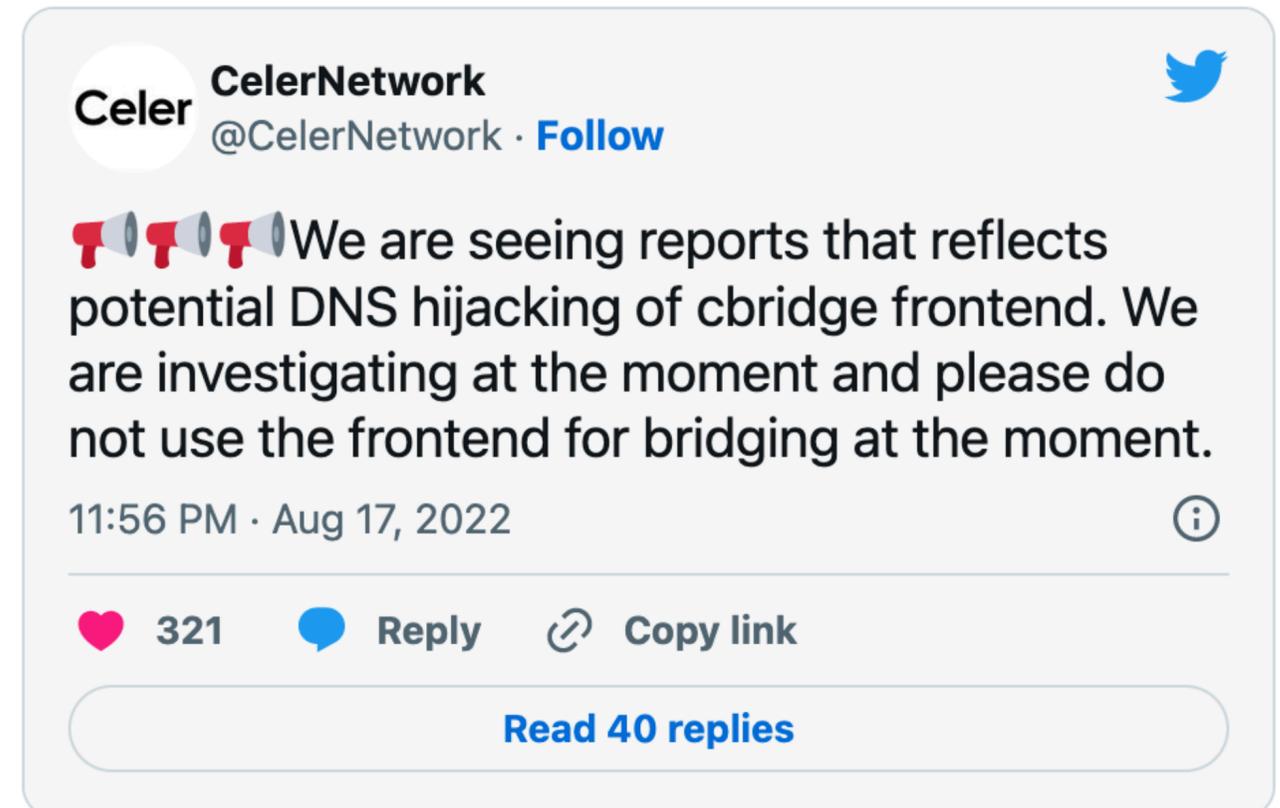
February 3, 2022

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.



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.

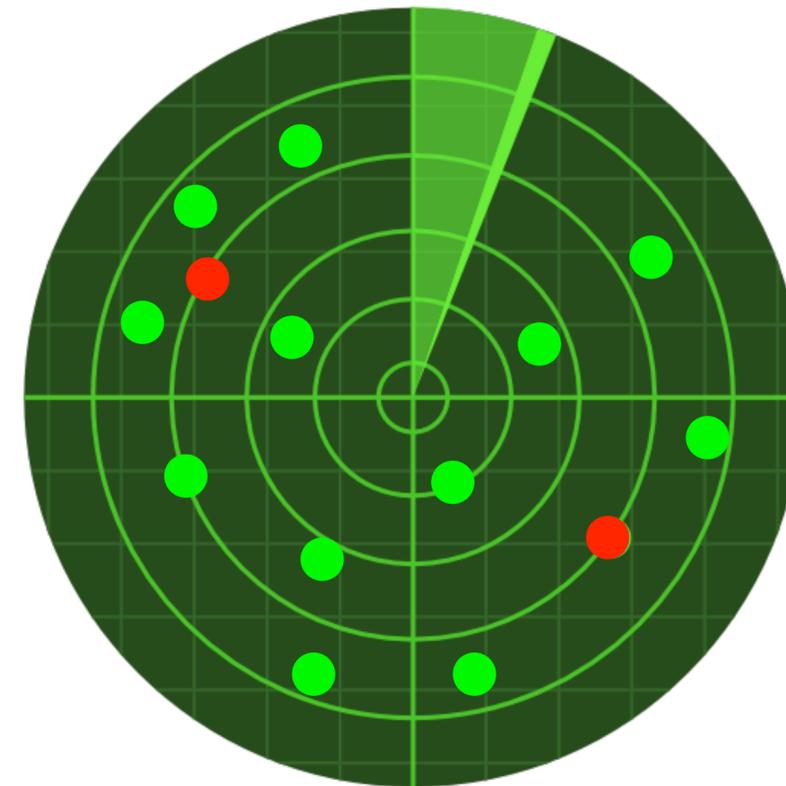
11:56 PM · Aug 17, 2022

 321  Reply  Copy link

[Read 40 replies](#)

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

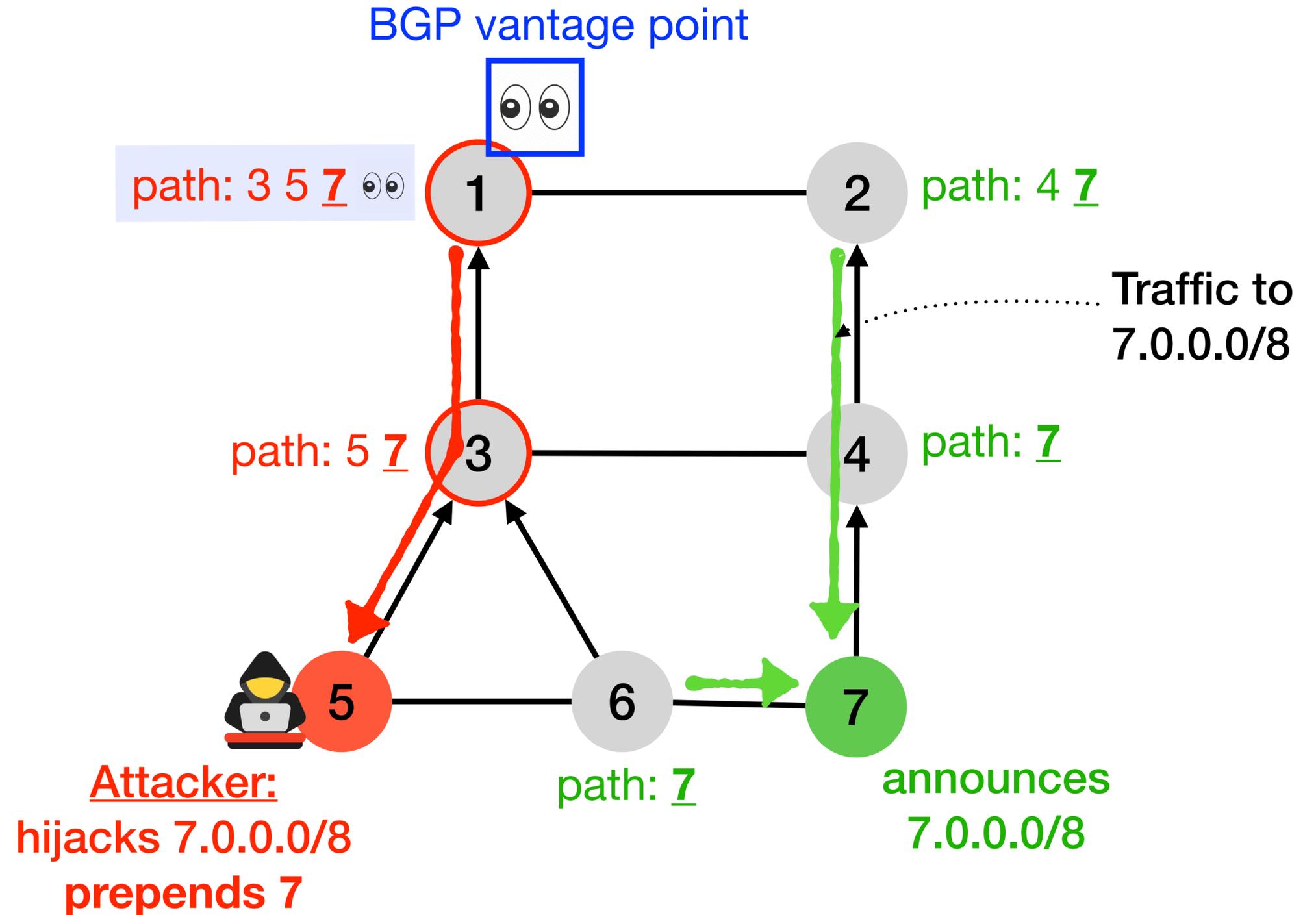
is to detect **fake** AS links

DFOH's inference pipeline

DFOH's inferences are accurate

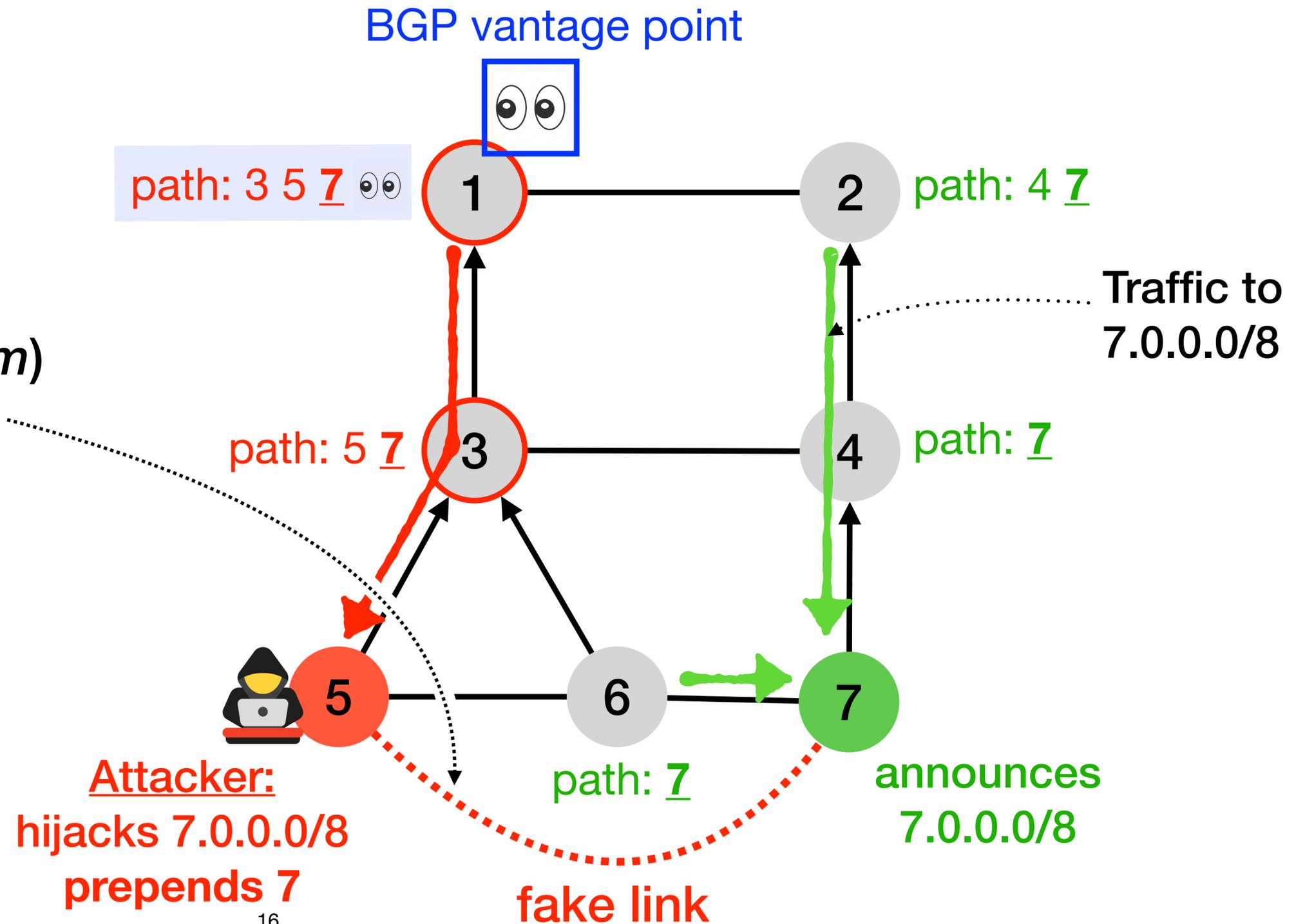
DFOH is up and running

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

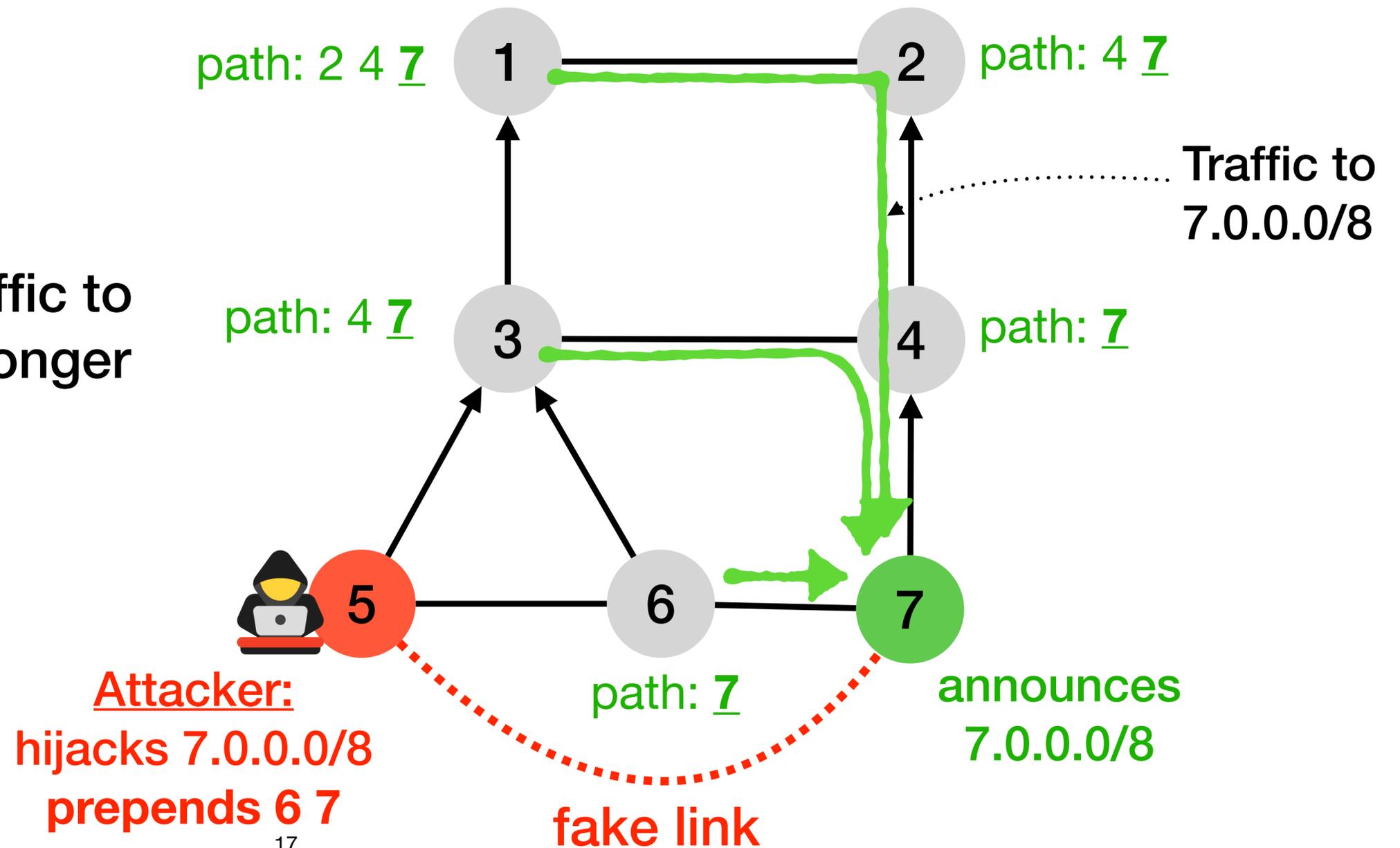
Upon the attack:
AS5 (*attacker*) and AS7 (*victim*) appear directly connected



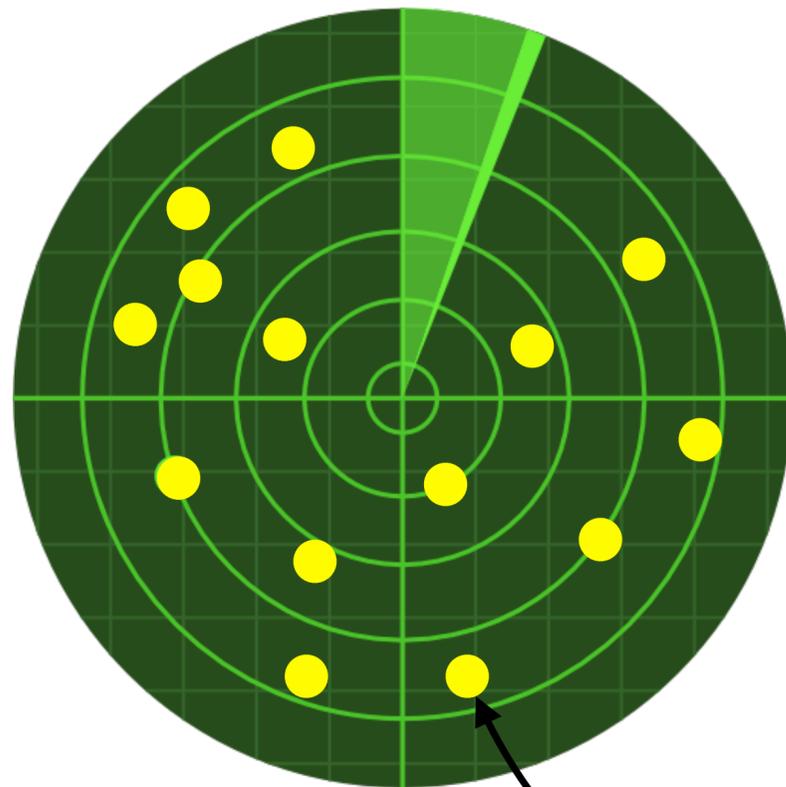
An attacker **cannot escape** from creating a fake AS link without hampering the effectiveness of its attack

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

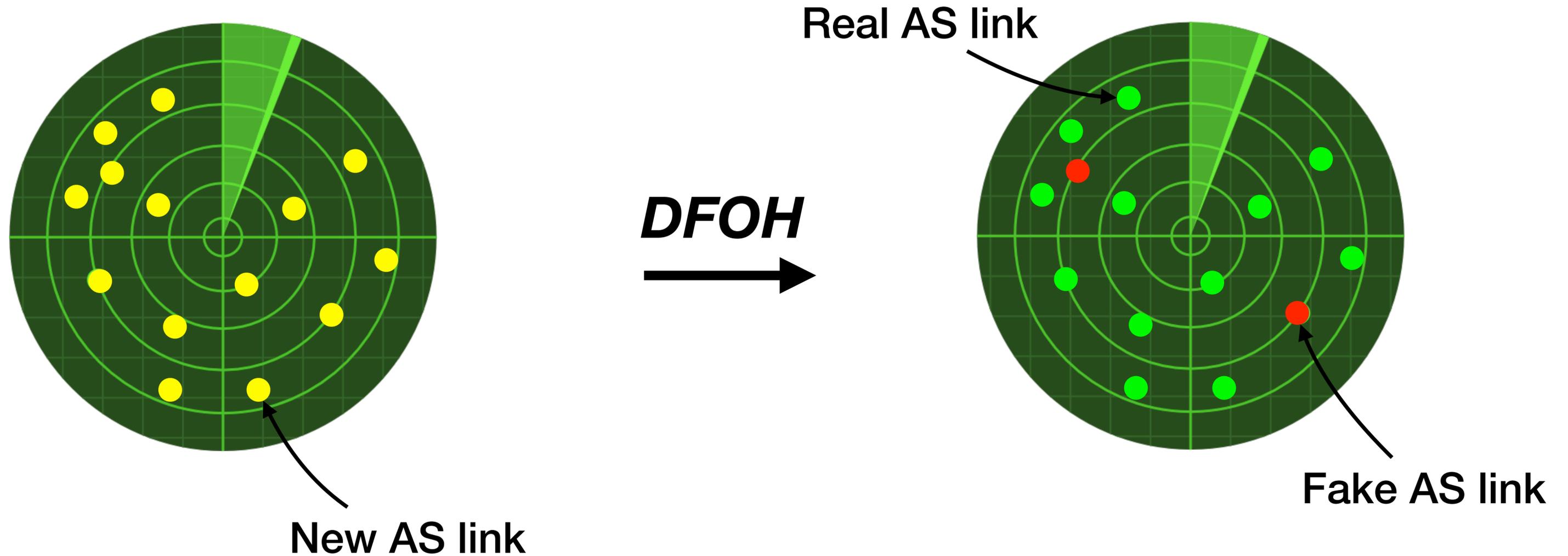


New AS link

**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**

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



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

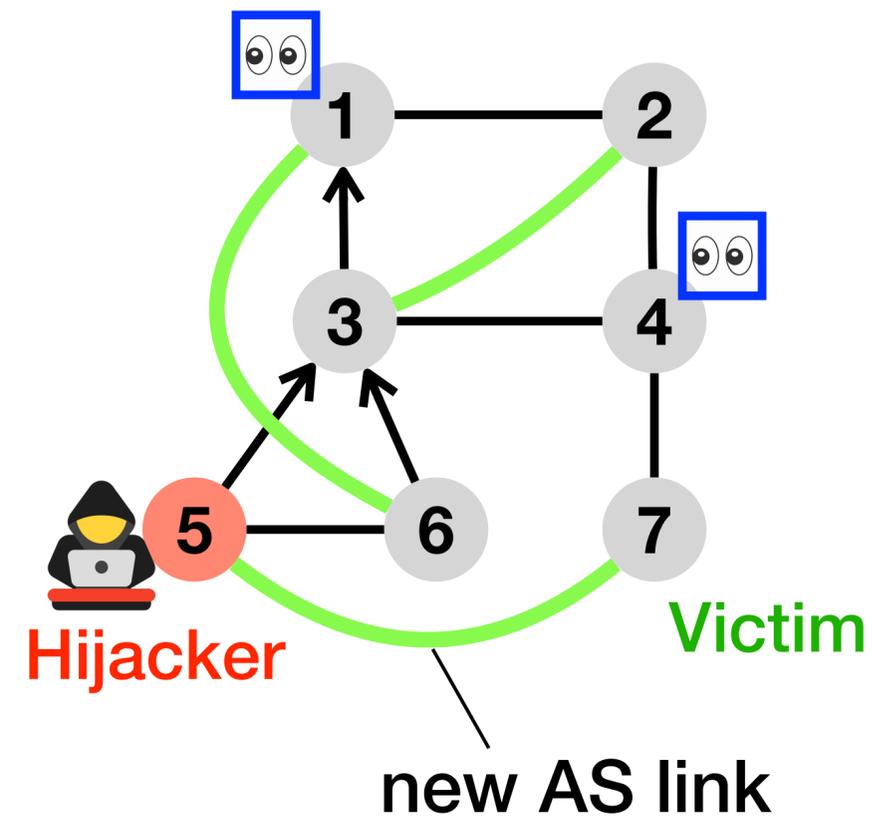
DFOH's inferences are accurate

DFOH is up and running

DFOH's fake AS links inference algorithm comprises three steps



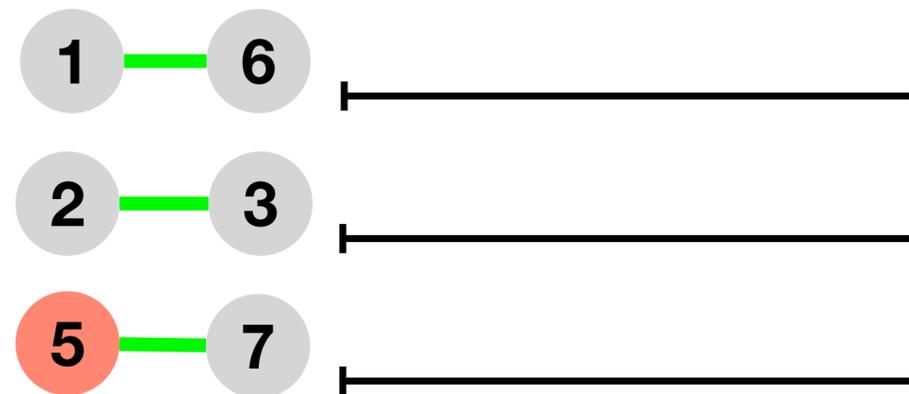
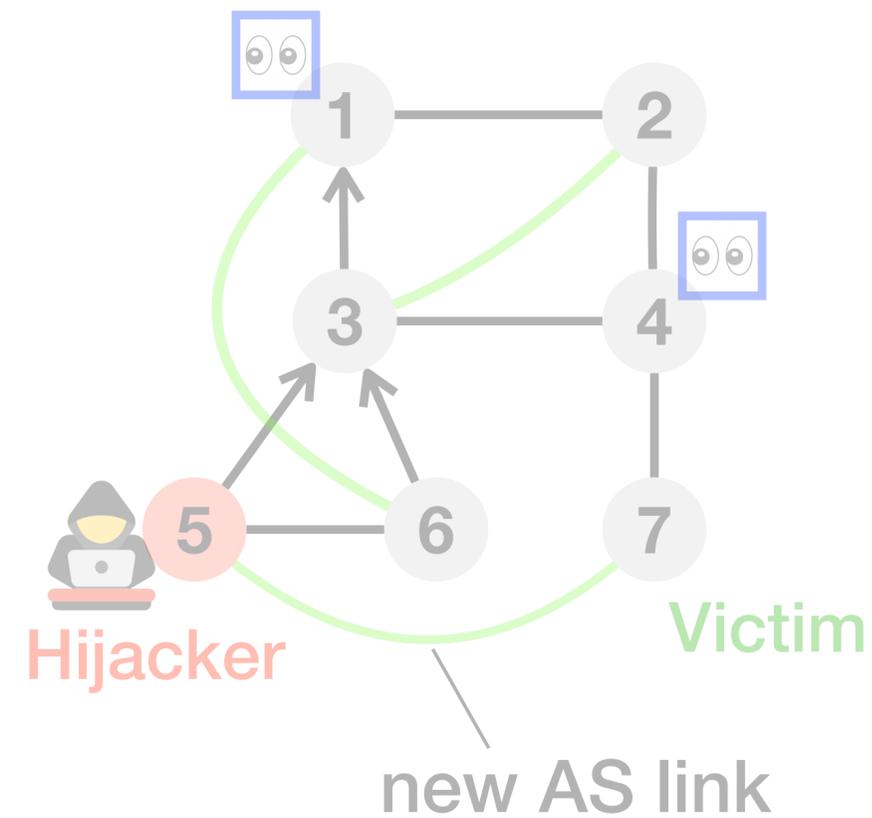
RIS/RouteViews
Vantage point



DFOH's fake AS links inference algorithm comprises three steps



RIS/RouteViews
Vantage point

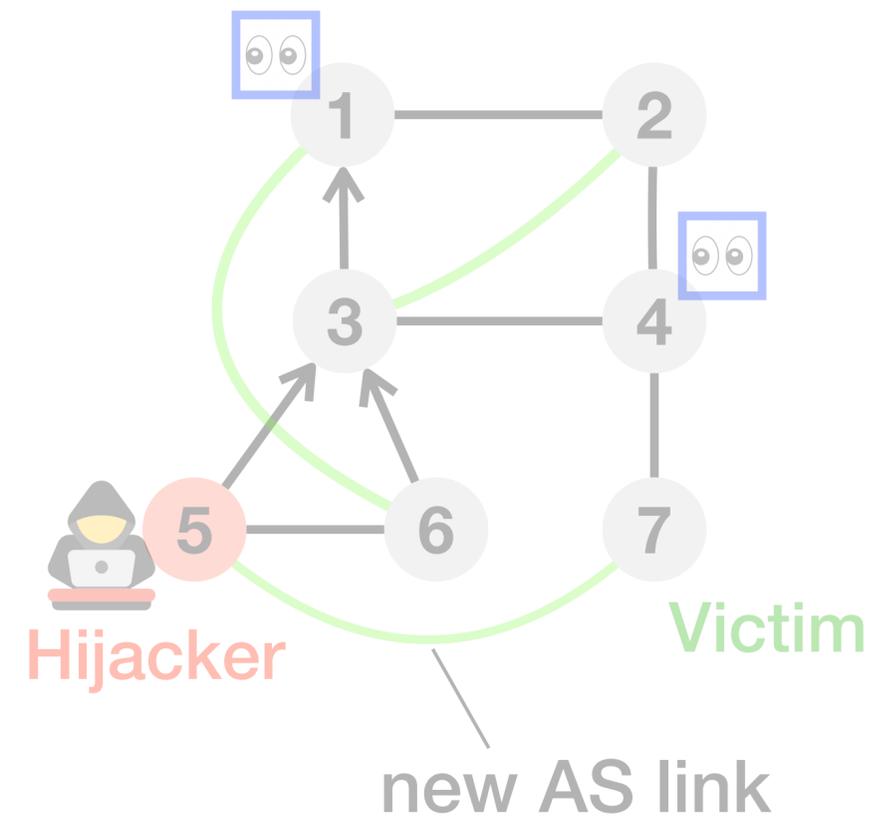


Feature vectors

DFOH's fake AS links inference algorithm comprises three steps

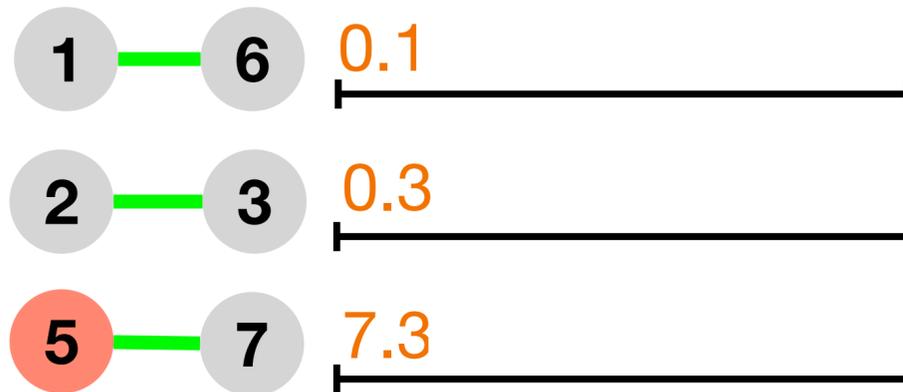


RIS/RouteViews
Vantage point



Feature categories:

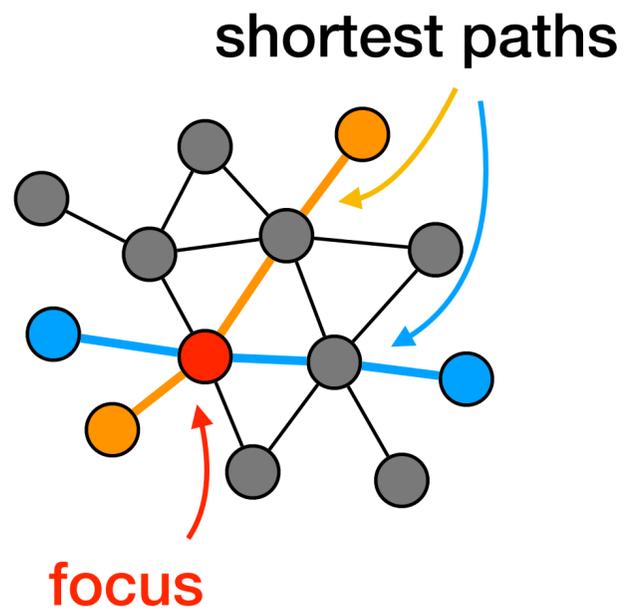
Topological



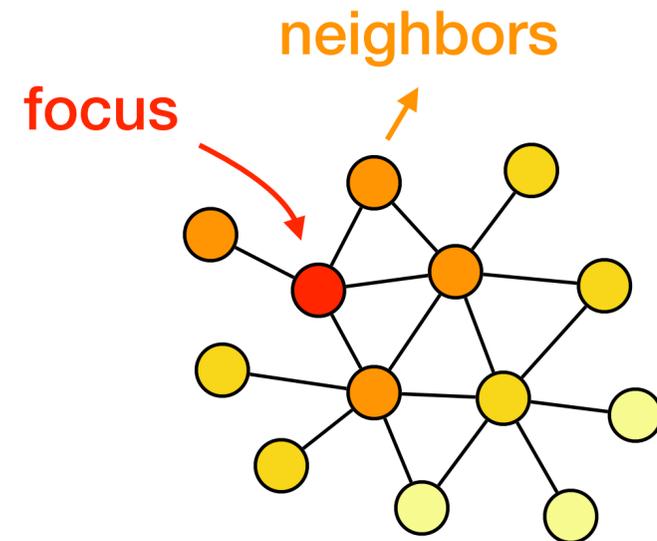
Feature vectors

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

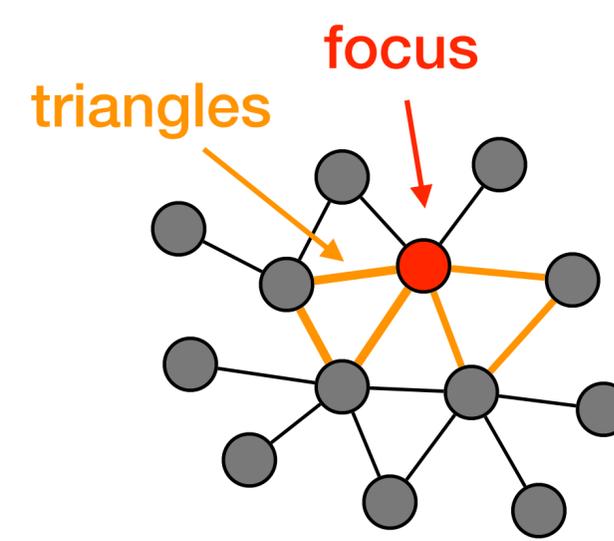
Node centrality



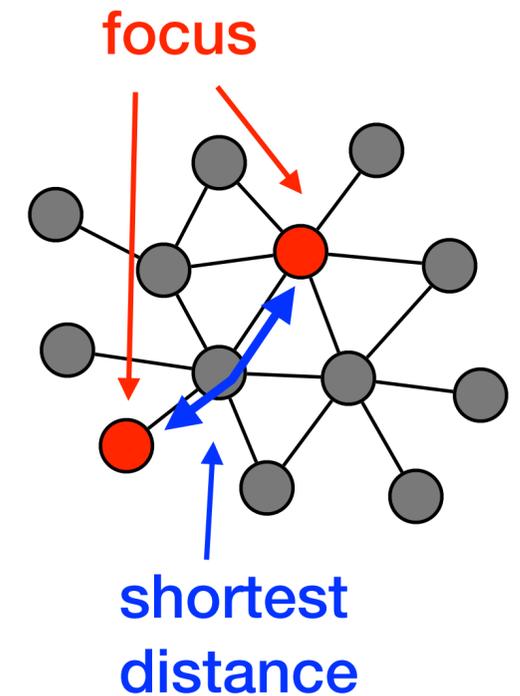
Neighborhood richness



Topological patterns



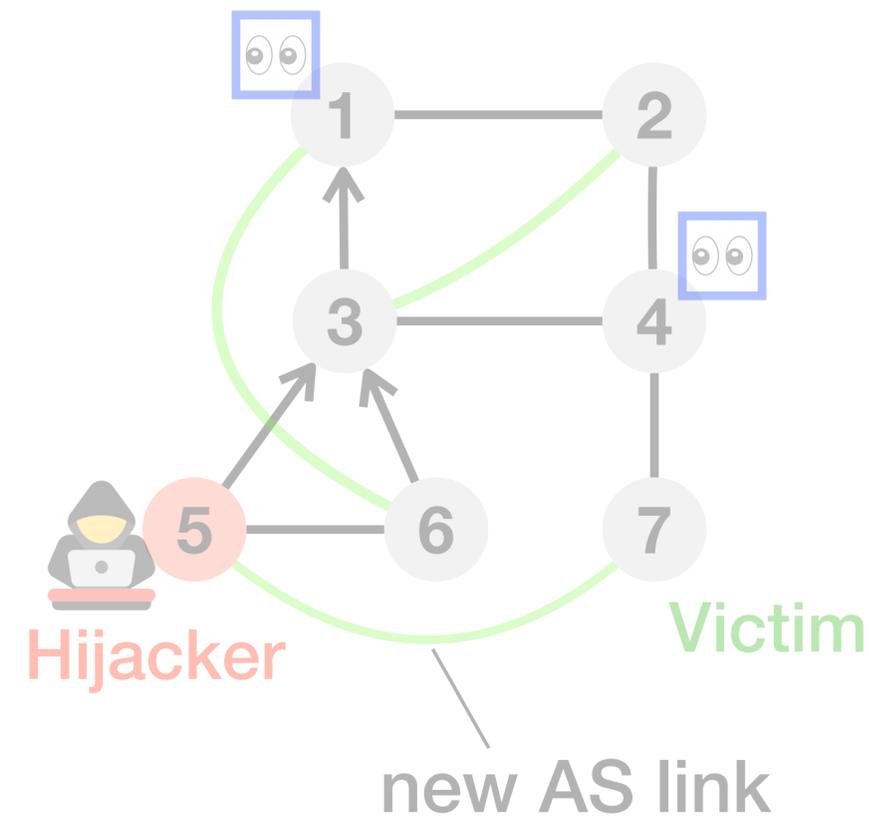
Closeness



DFOH's fake AS links inference algorithm comprises three steps



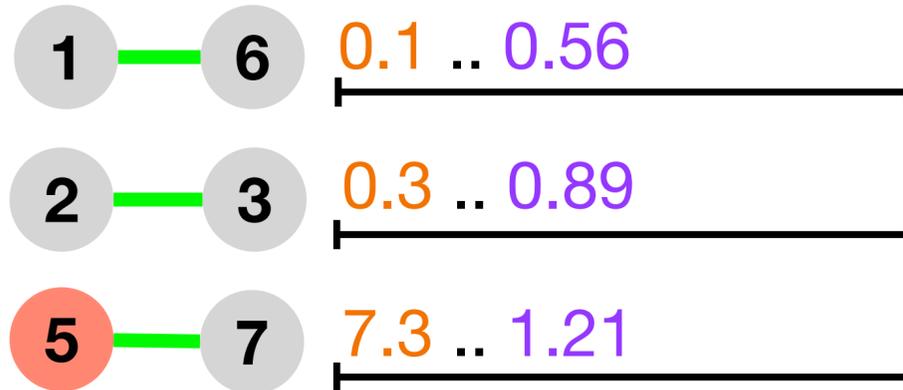
RIS/RouteViews
Vantage point



Feature categories:

Peeringdb

Topological



Feature vectors

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

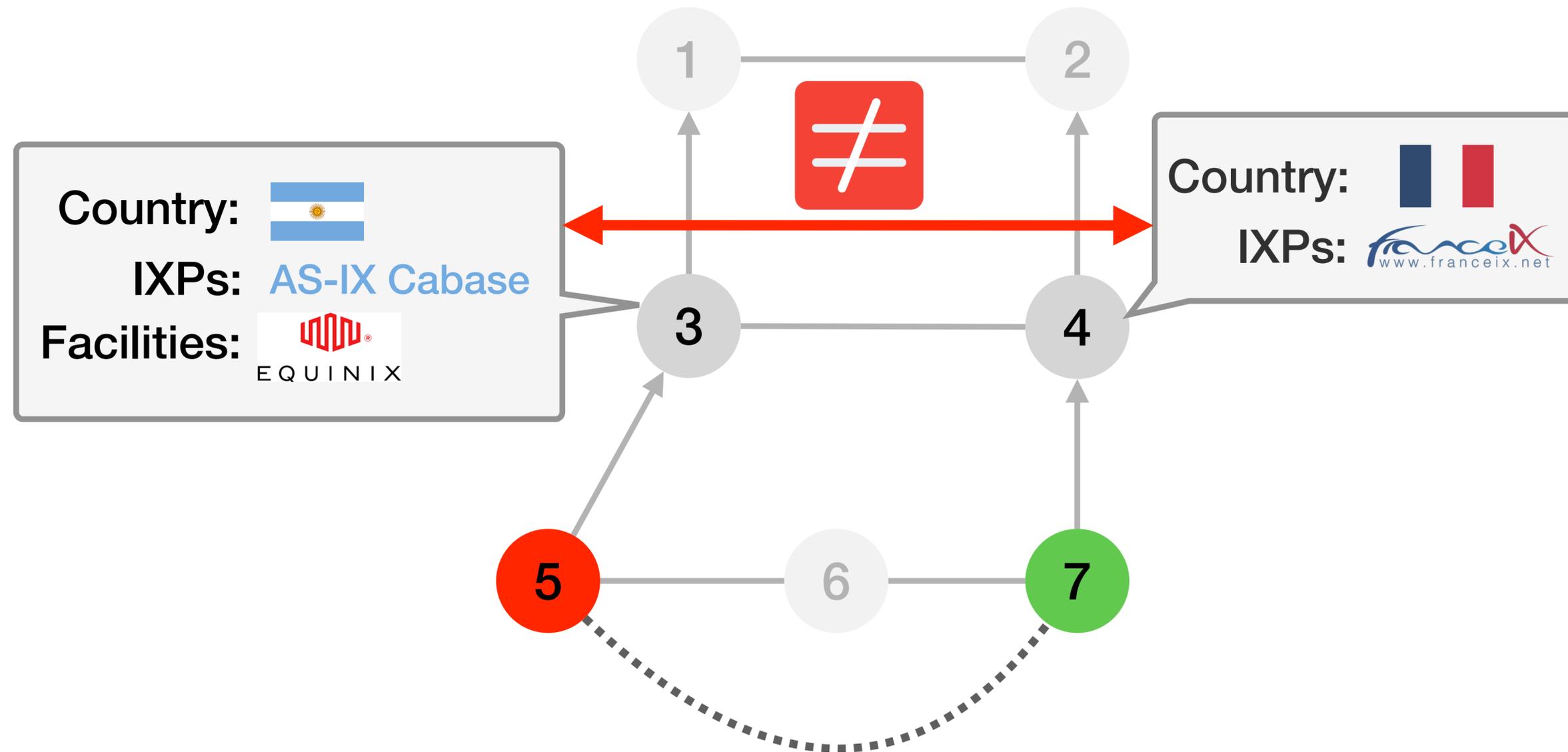
DFOH looks for three types of information in PeeringDB:

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

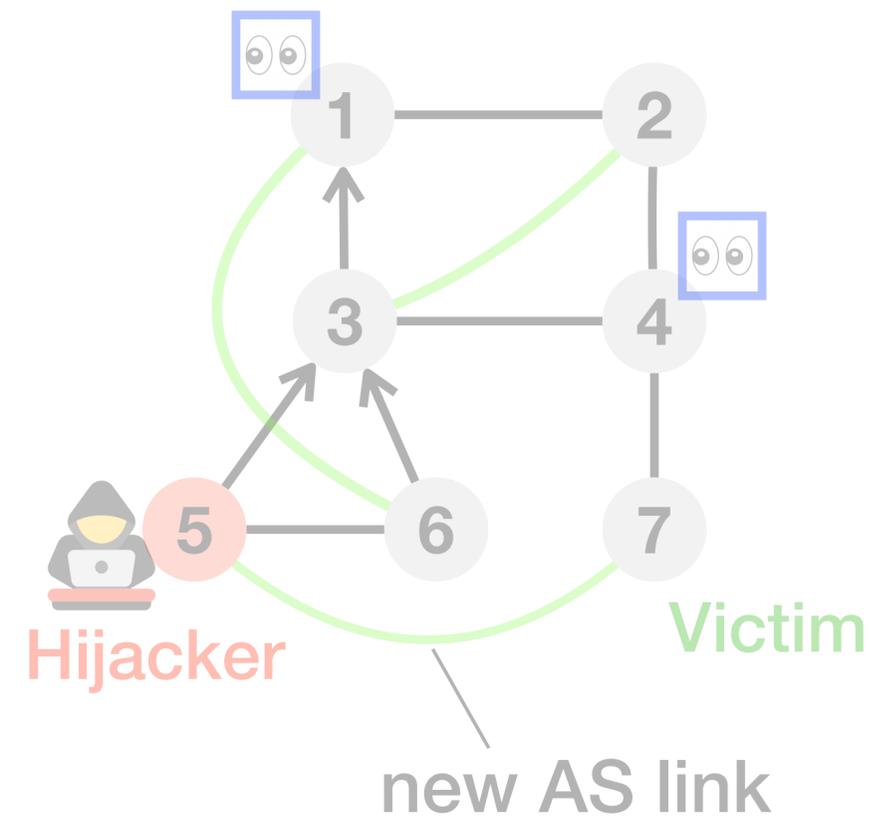
Reason #2:
Mitigate missing
peering information



DFOH's fake AS links inference algorithm comprises three steps



RIS/RouteViews
Vantage point

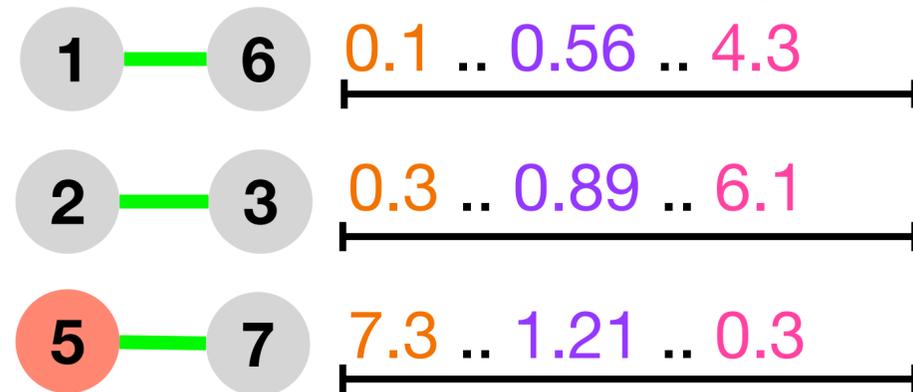


Feature categories:

AS-path pattern

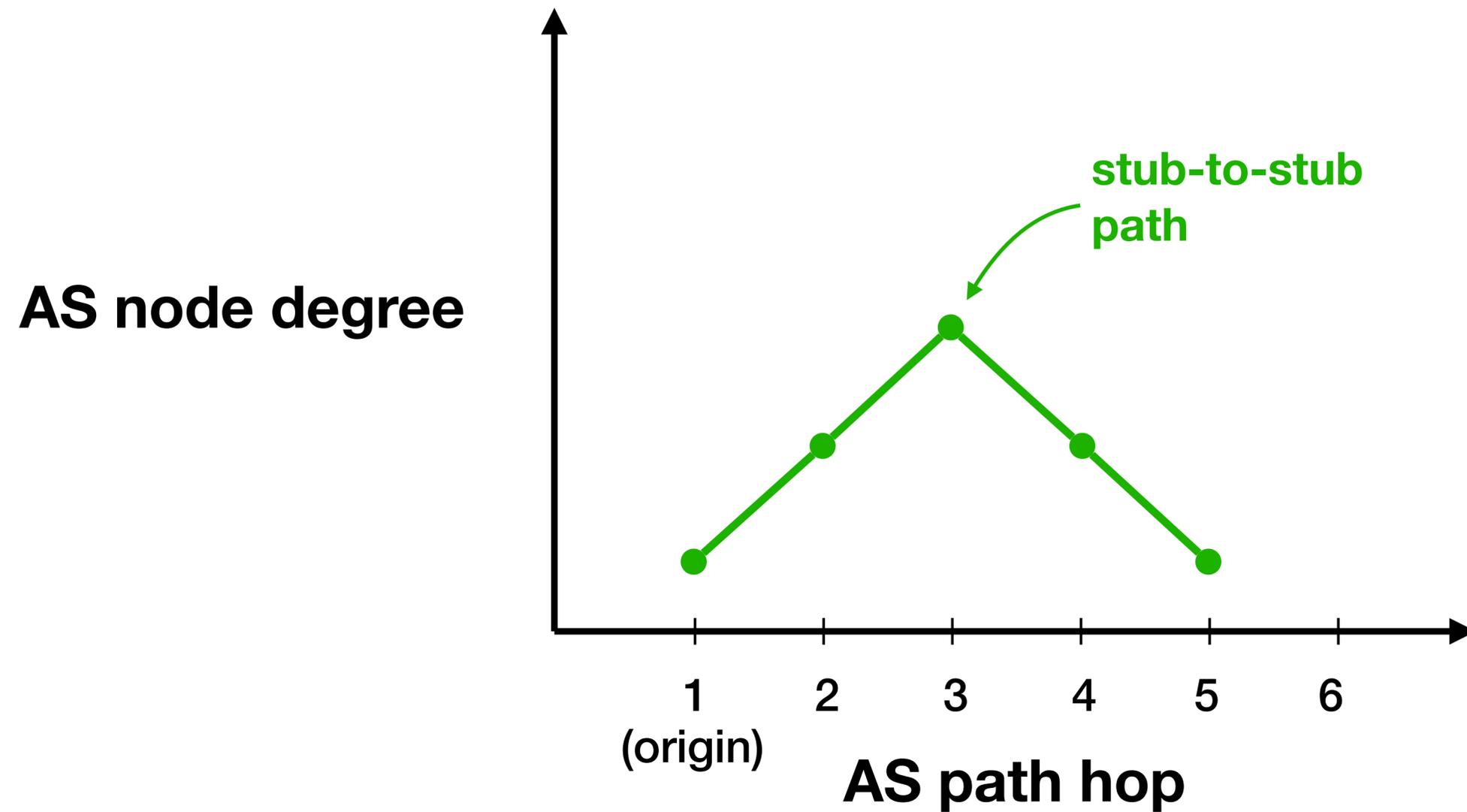
Peeringdb

Topological

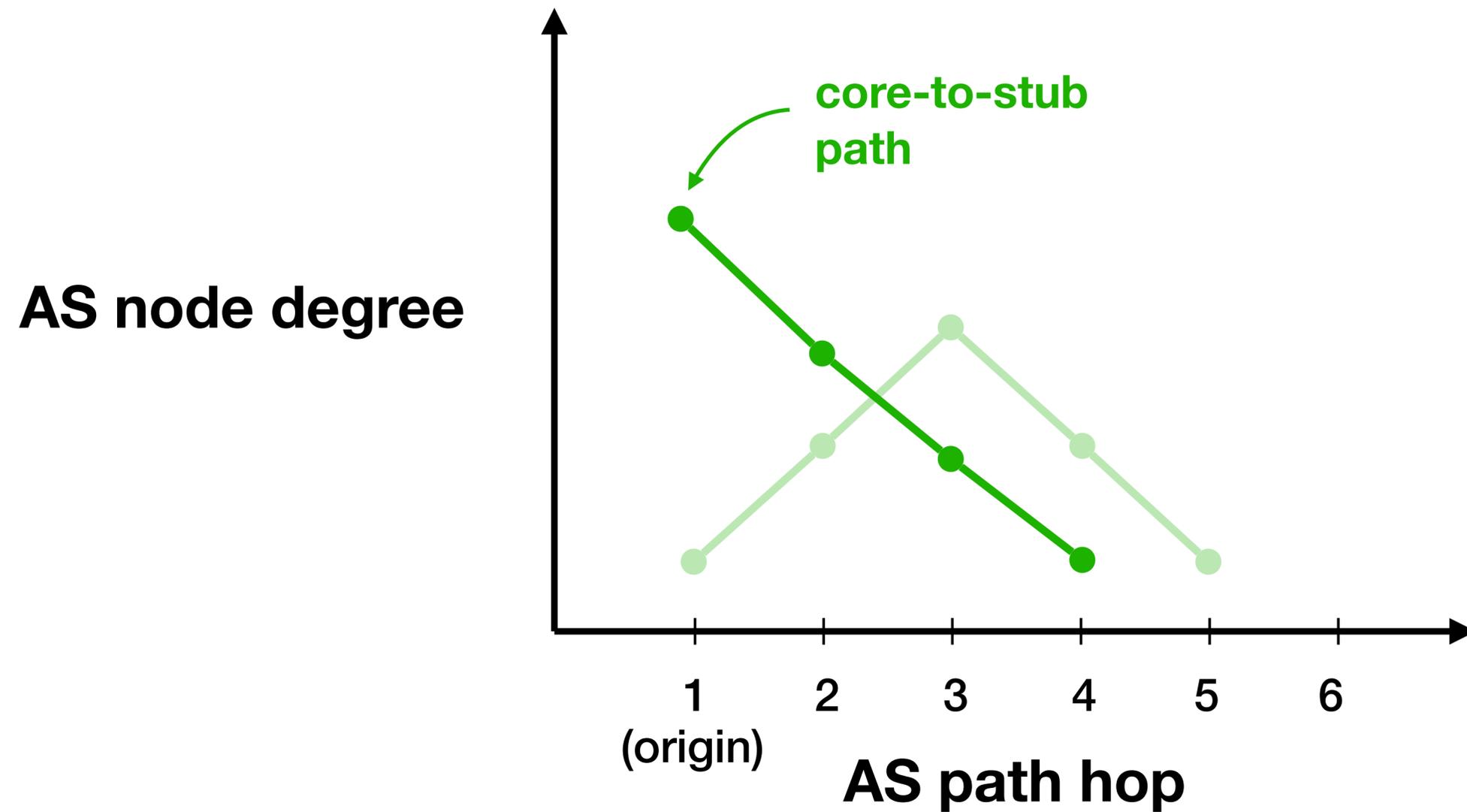


Feature vectors

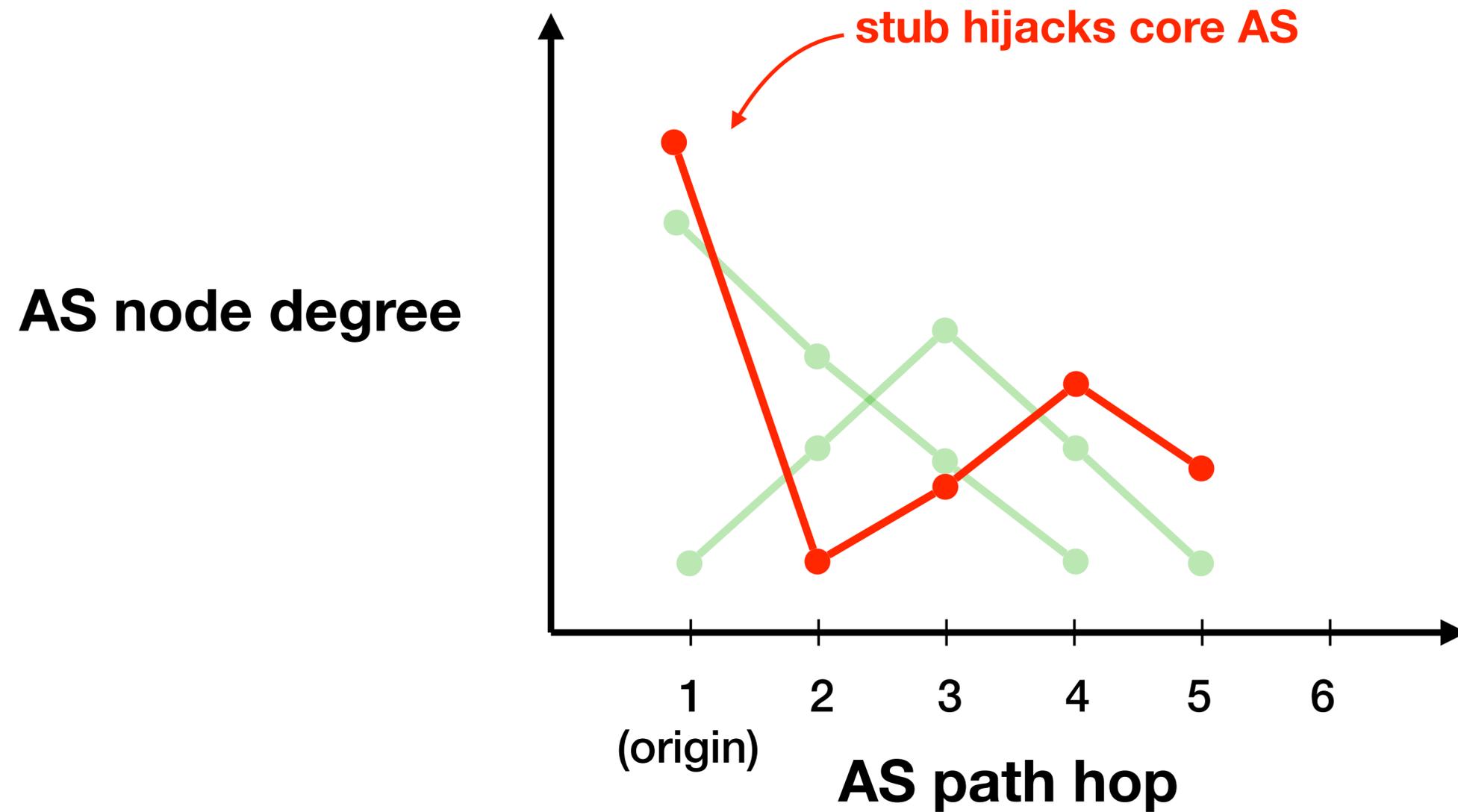
DFOH looks at the AS paths that include the new link and identifies **suspicious sequence of ASes**



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



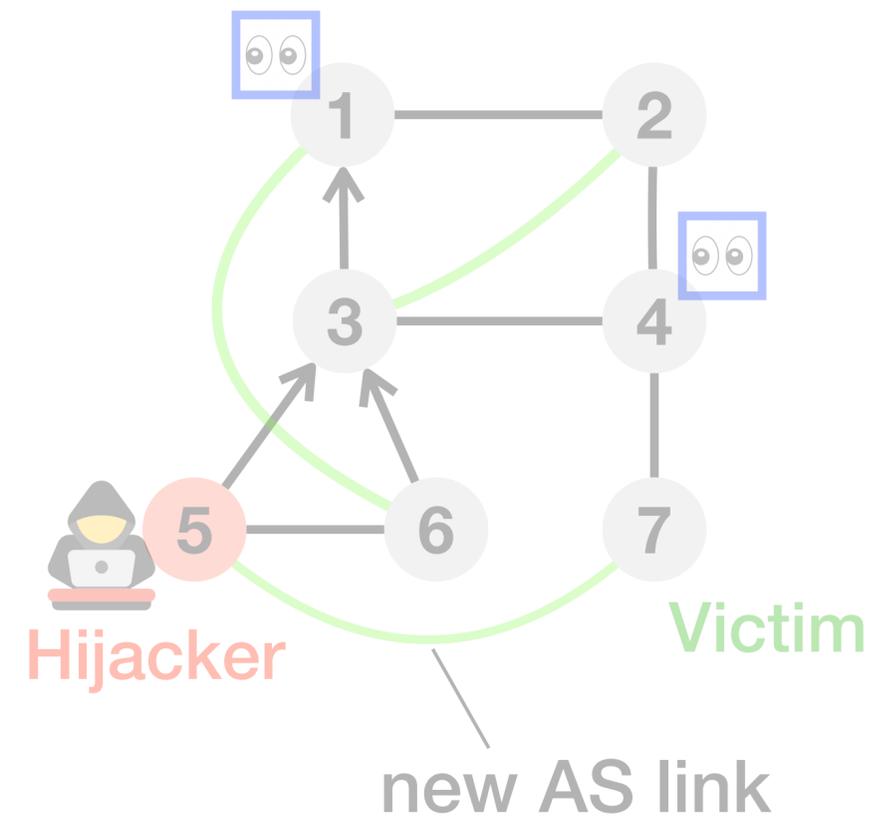
DFOH looks at the AS paths that include the new link and identifies **suspicious sequence of ASes**



DFOH's fake AS links inference algorithm comprises three steps



RIS/RouteViews
Vantage point



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

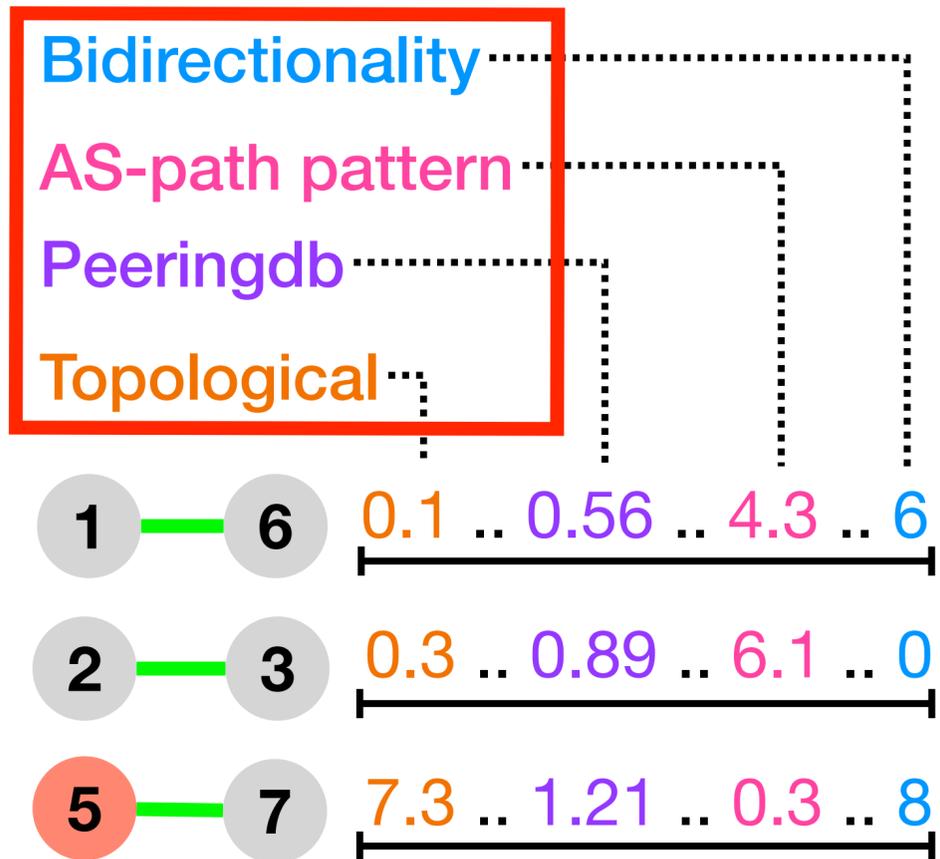
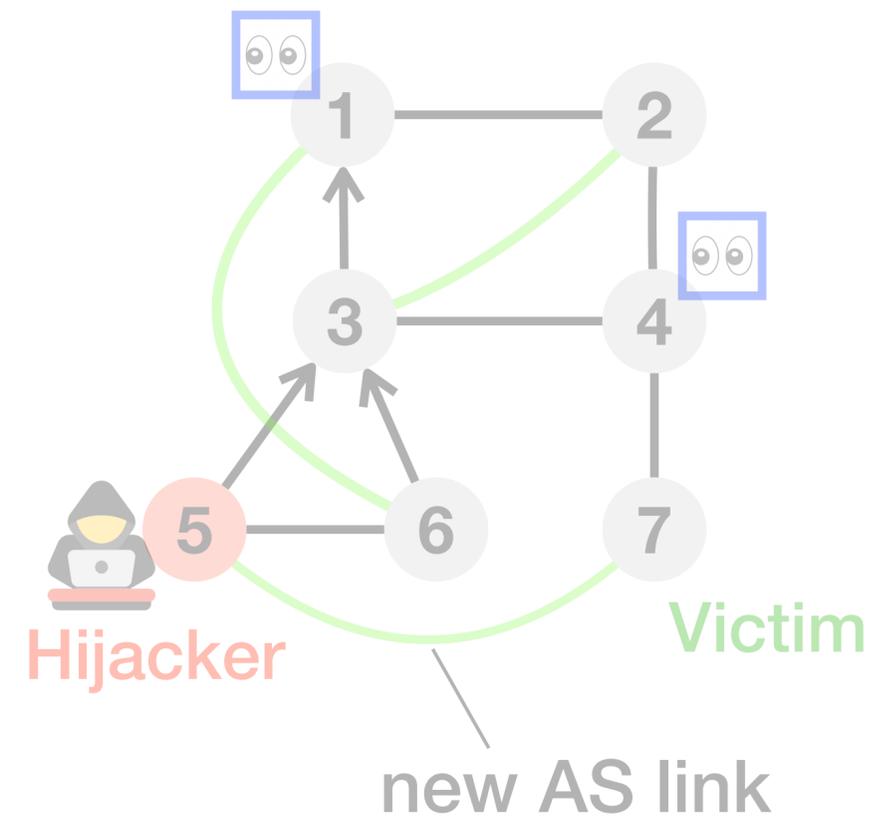
Feature vectors

DFOH's fake AS links inference algorithm comprises three steps



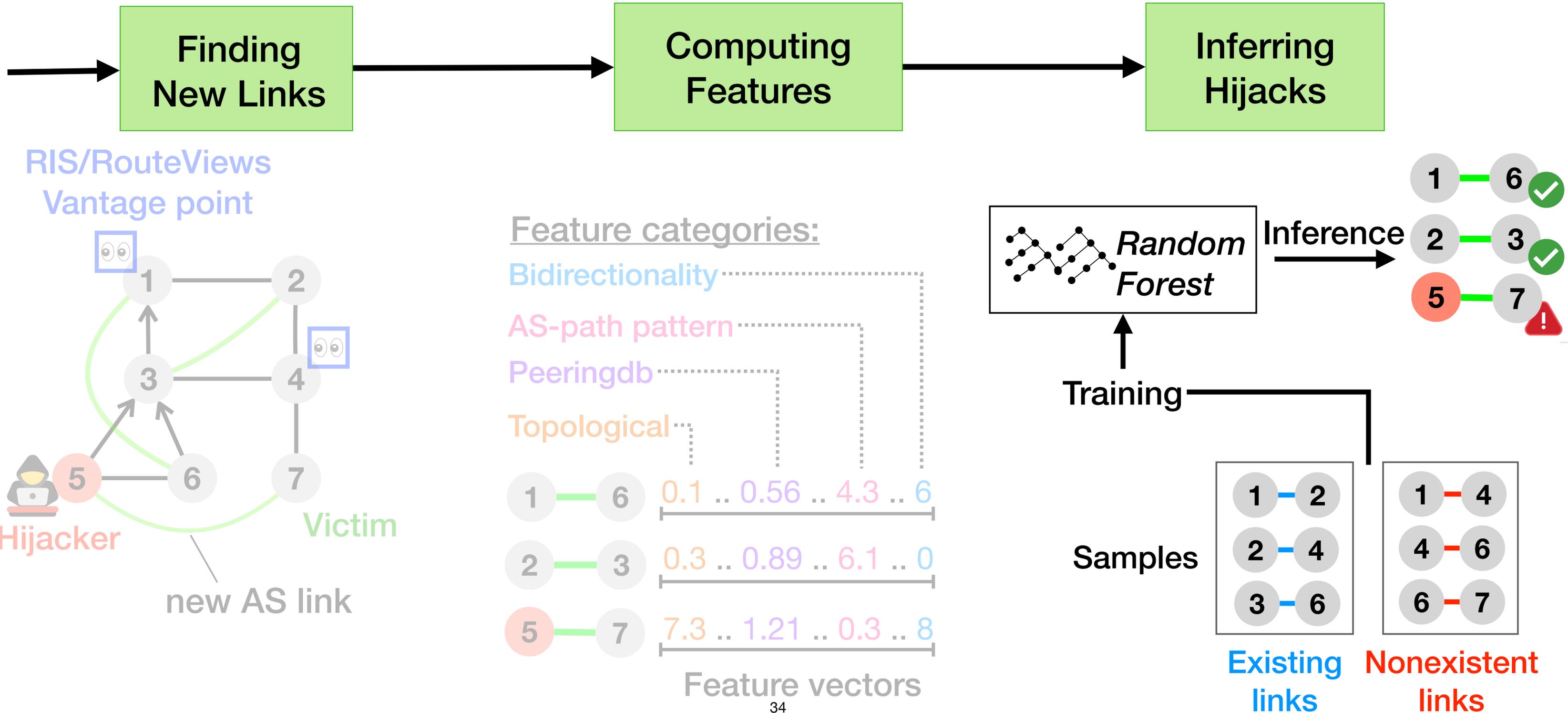
RIS/RouteViews
Vantage point

Domain-specific features
that compensate each other

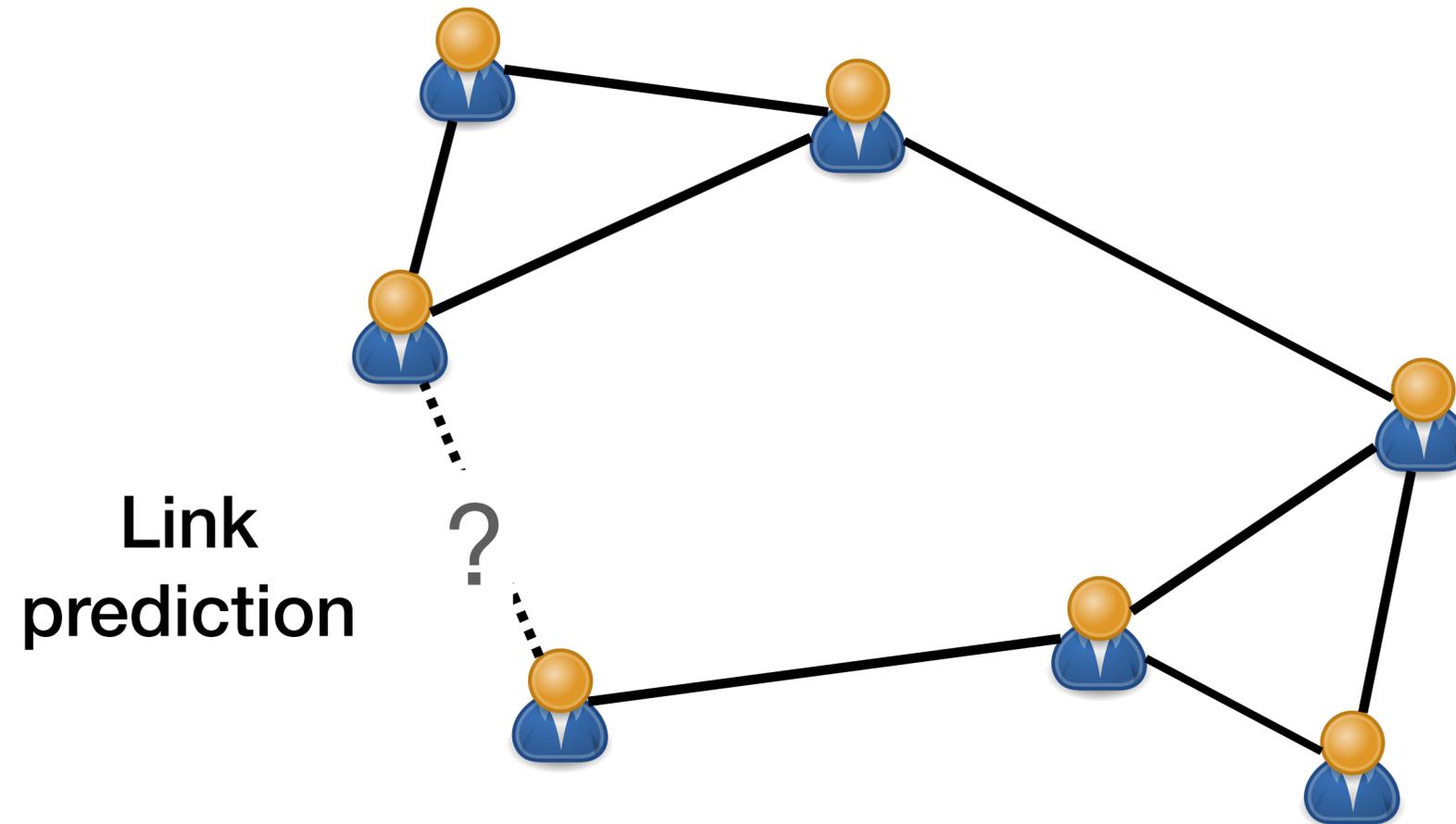


Feature vectors

DFOH's fake AS links inference algorithm comprises three steps



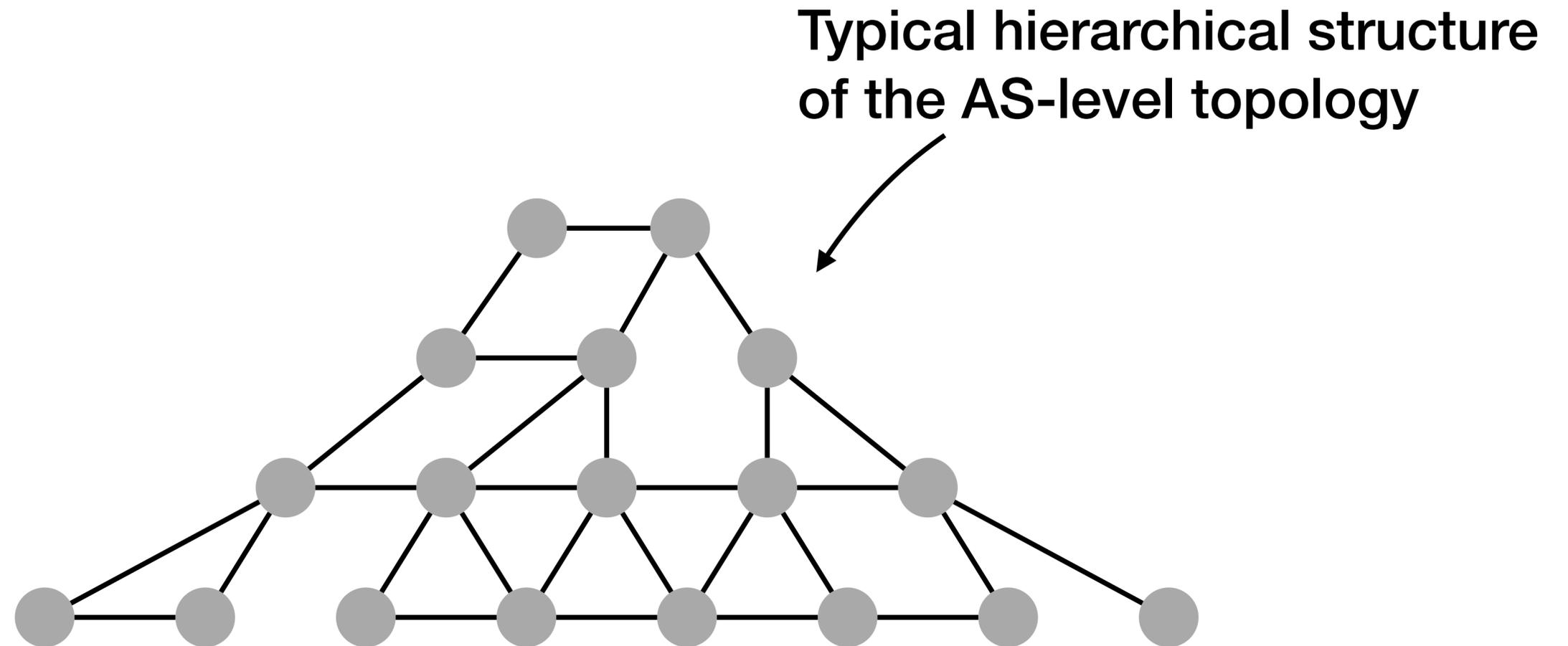
There are several link prediction frameworks
SEAL (NIPS'18) is one example



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

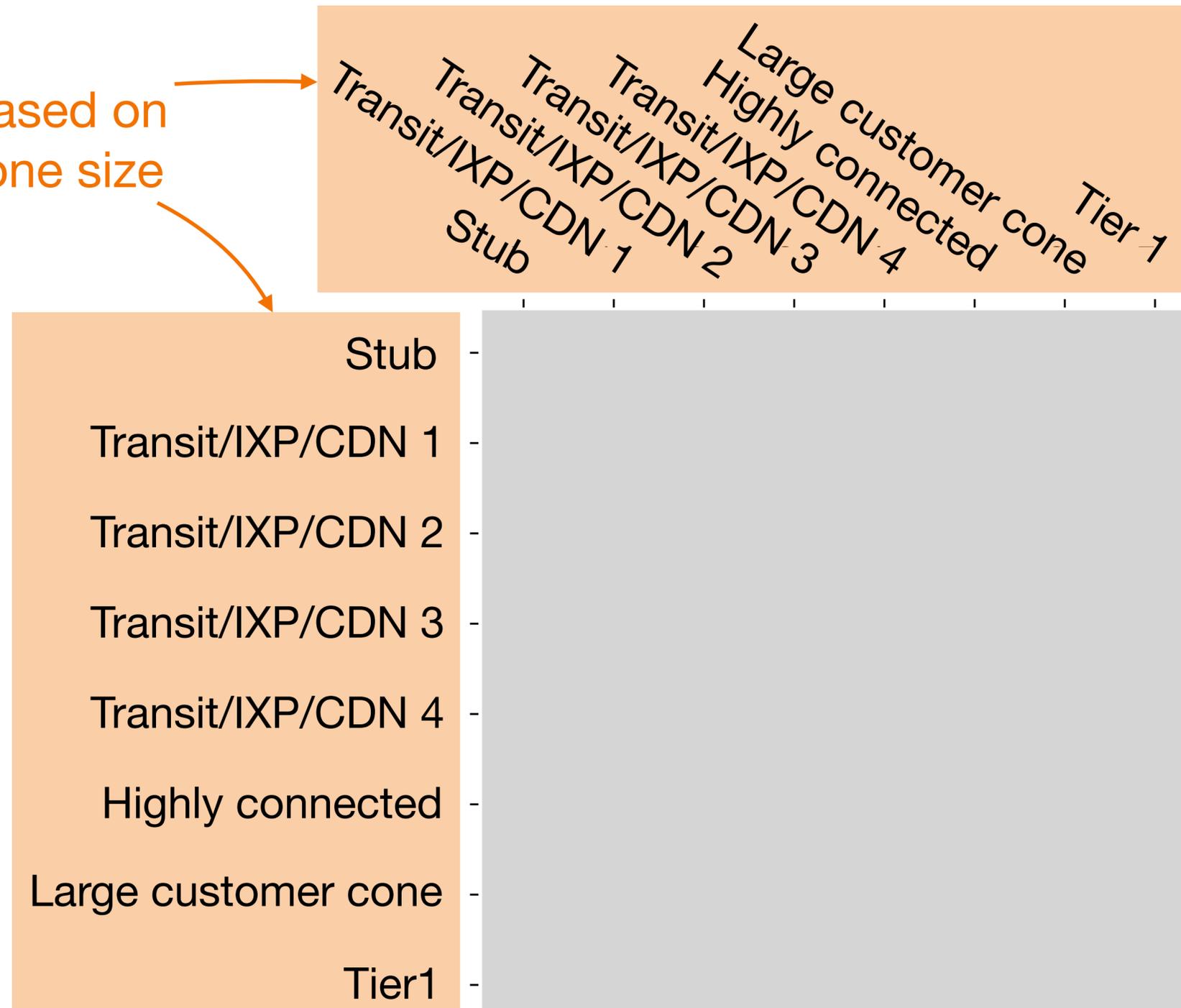
Few tier1 ASes

Many stub ASes



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

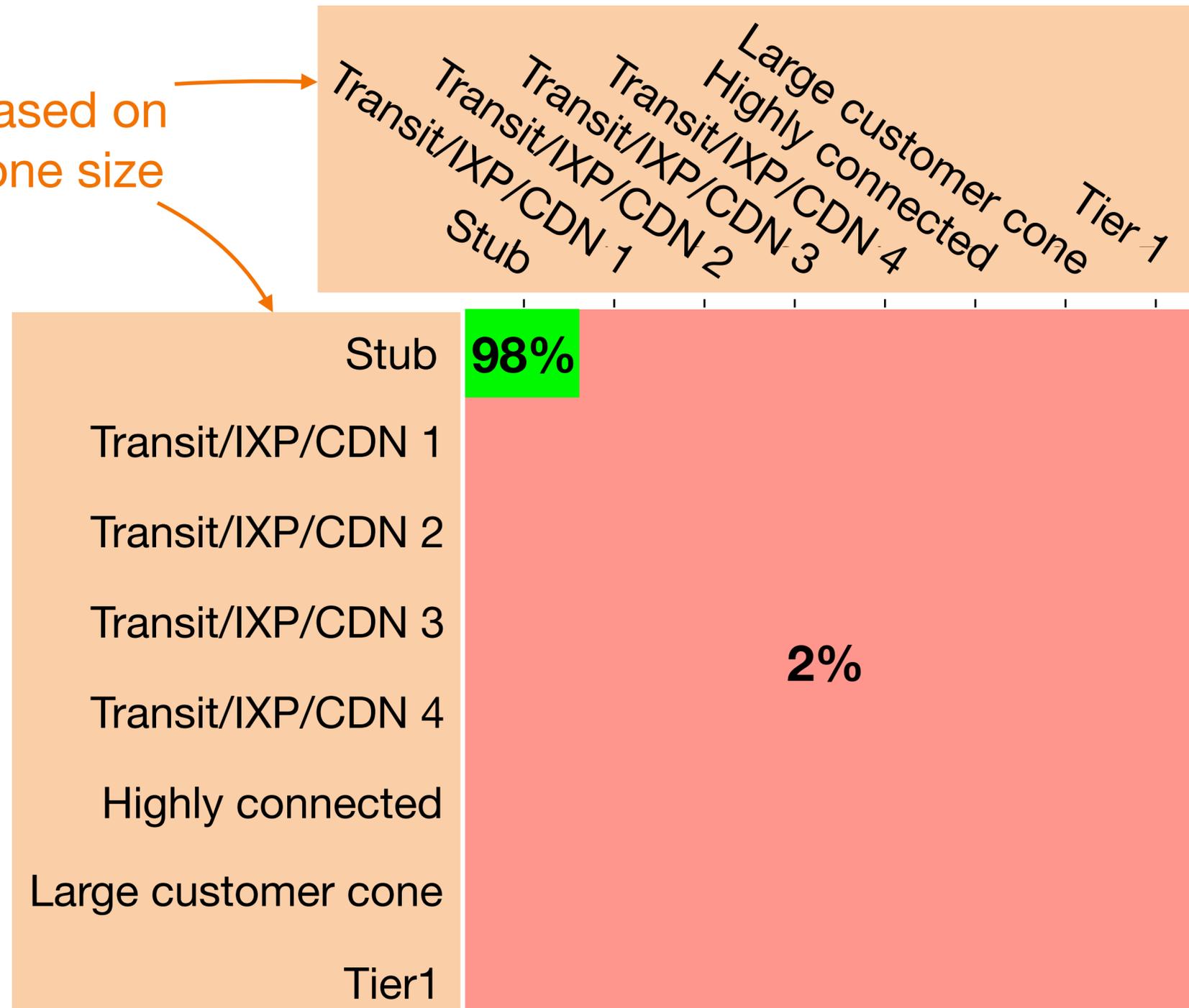
Clusters of ASes based on their degree and cone size



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**

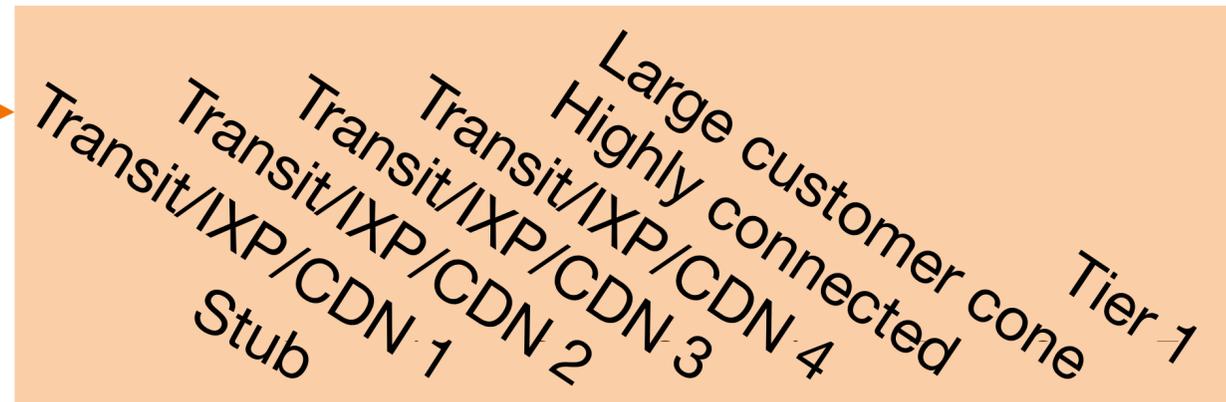
Clusters of ASes based on their degree and cone size



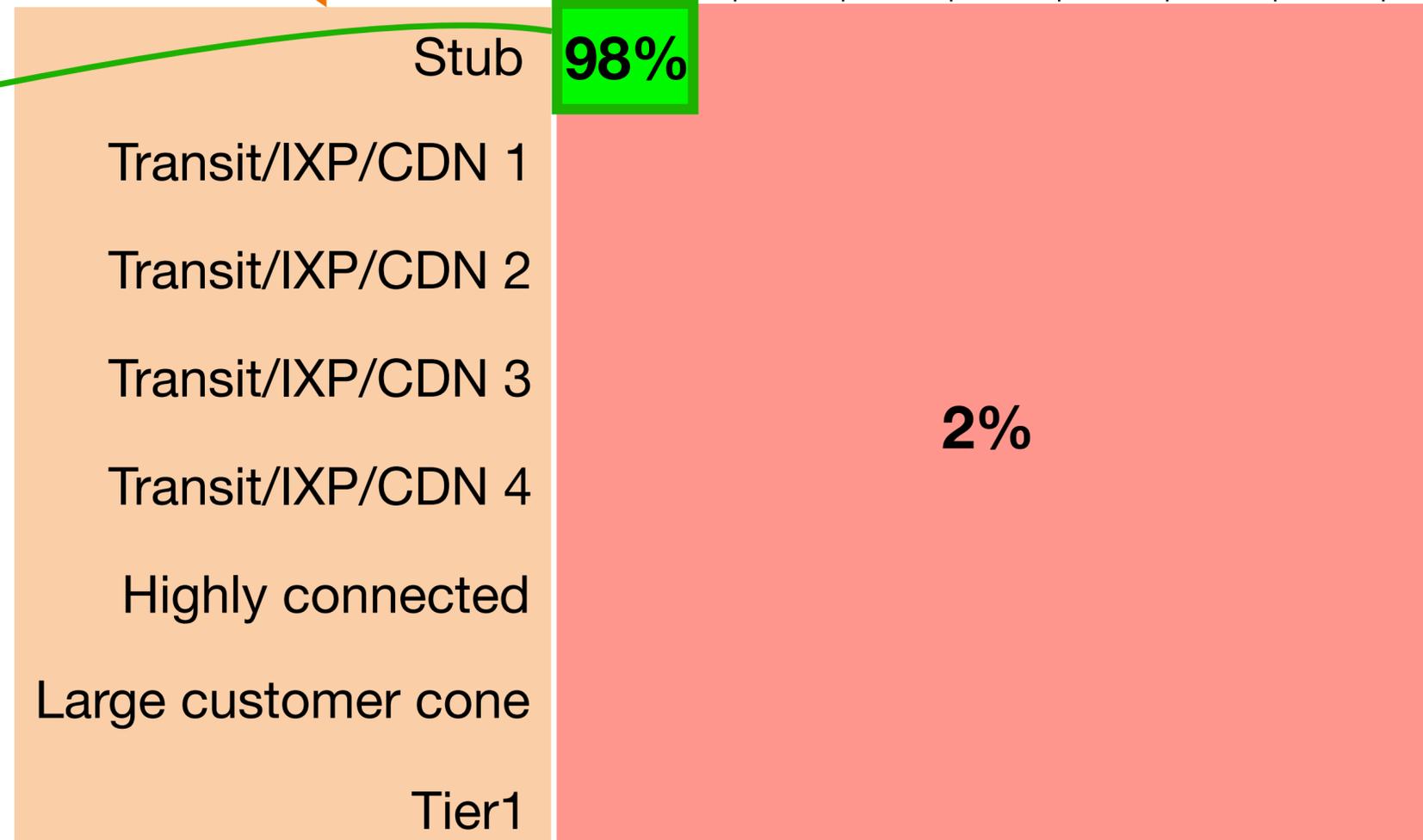
Proportion of sampled **nonexistent** AS links
(*random sampling*)

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



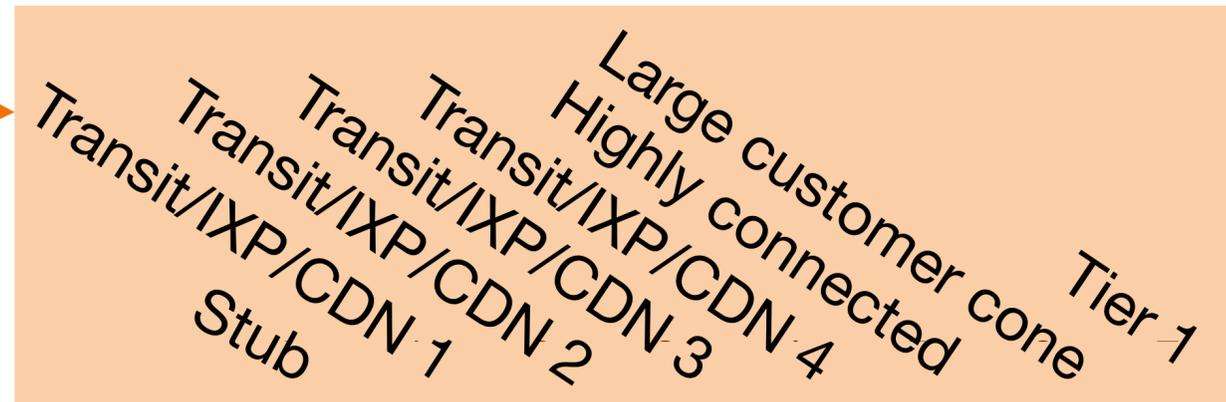
DFOH would perform well on scenarios involving two stubs



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

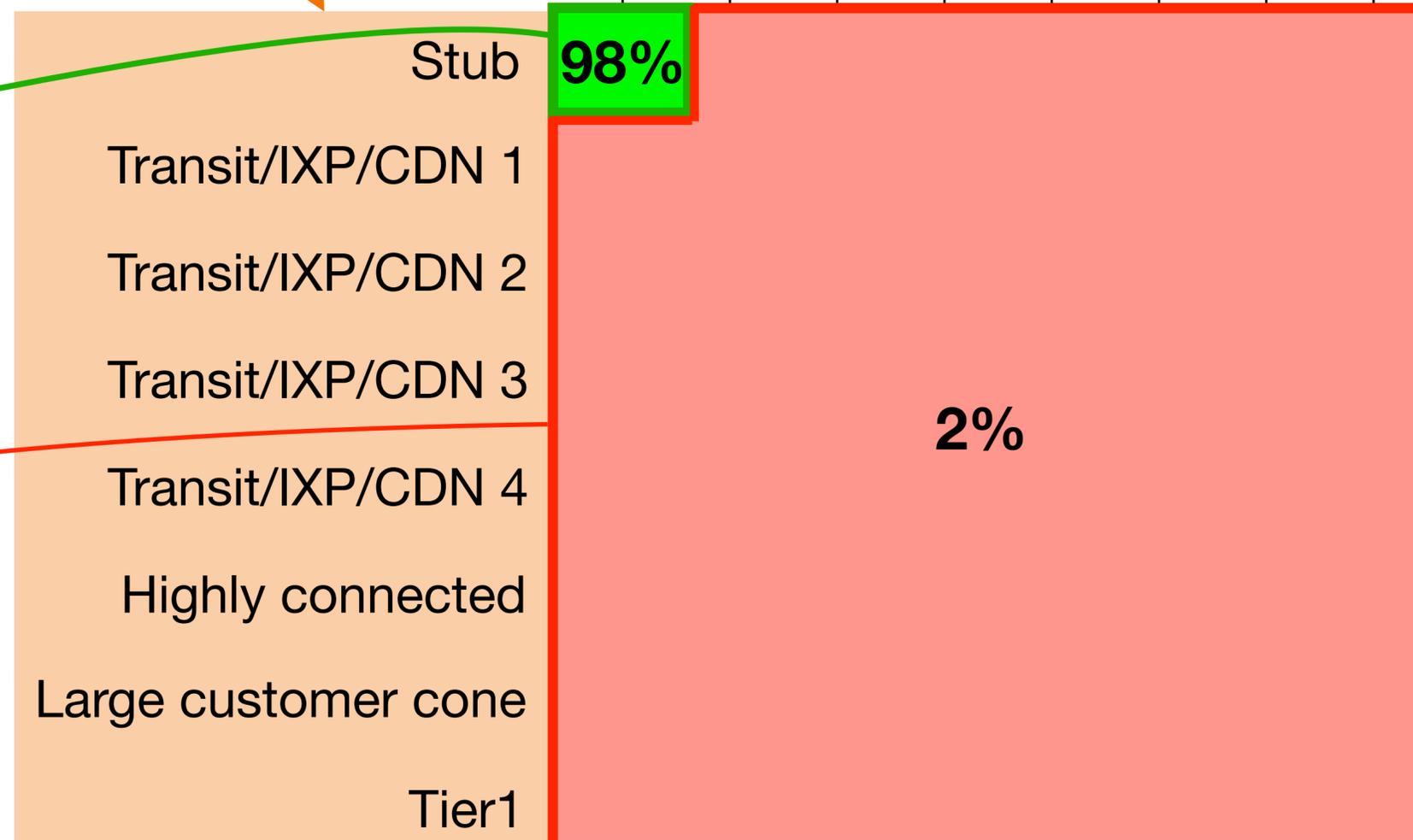
Problem: randomly sampling nonexistent links makes DFOH **skewed** towards stub-to-stub links as they are **overrepresented**

Clusters of ASes based on their degree and cone size



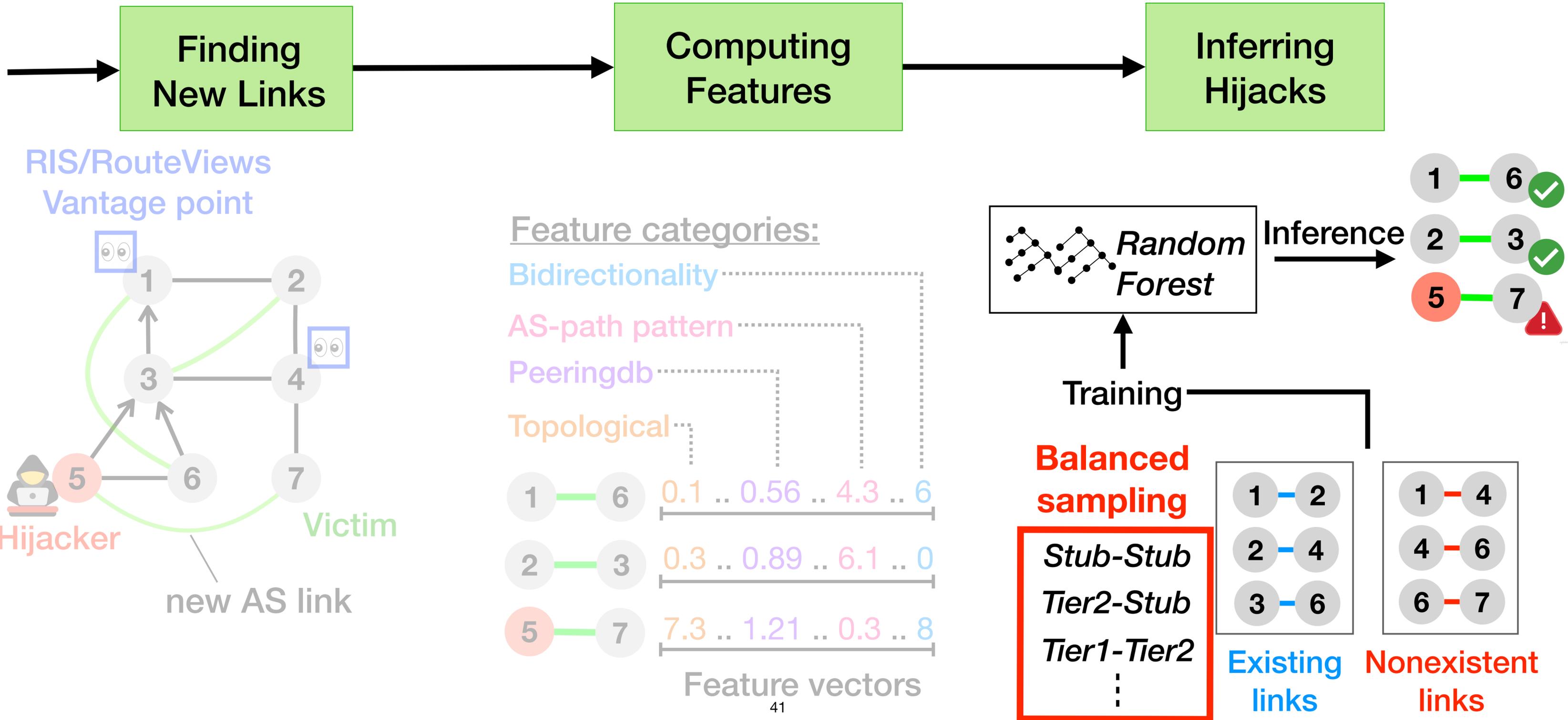
DFOH would perform well on scenarios involving two stubs

But not on the other scenarios



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

DFOH's fake AS links inference algorithm comprises three steps



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)



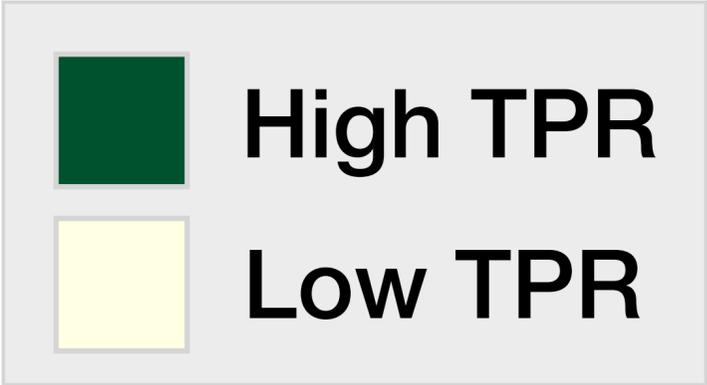
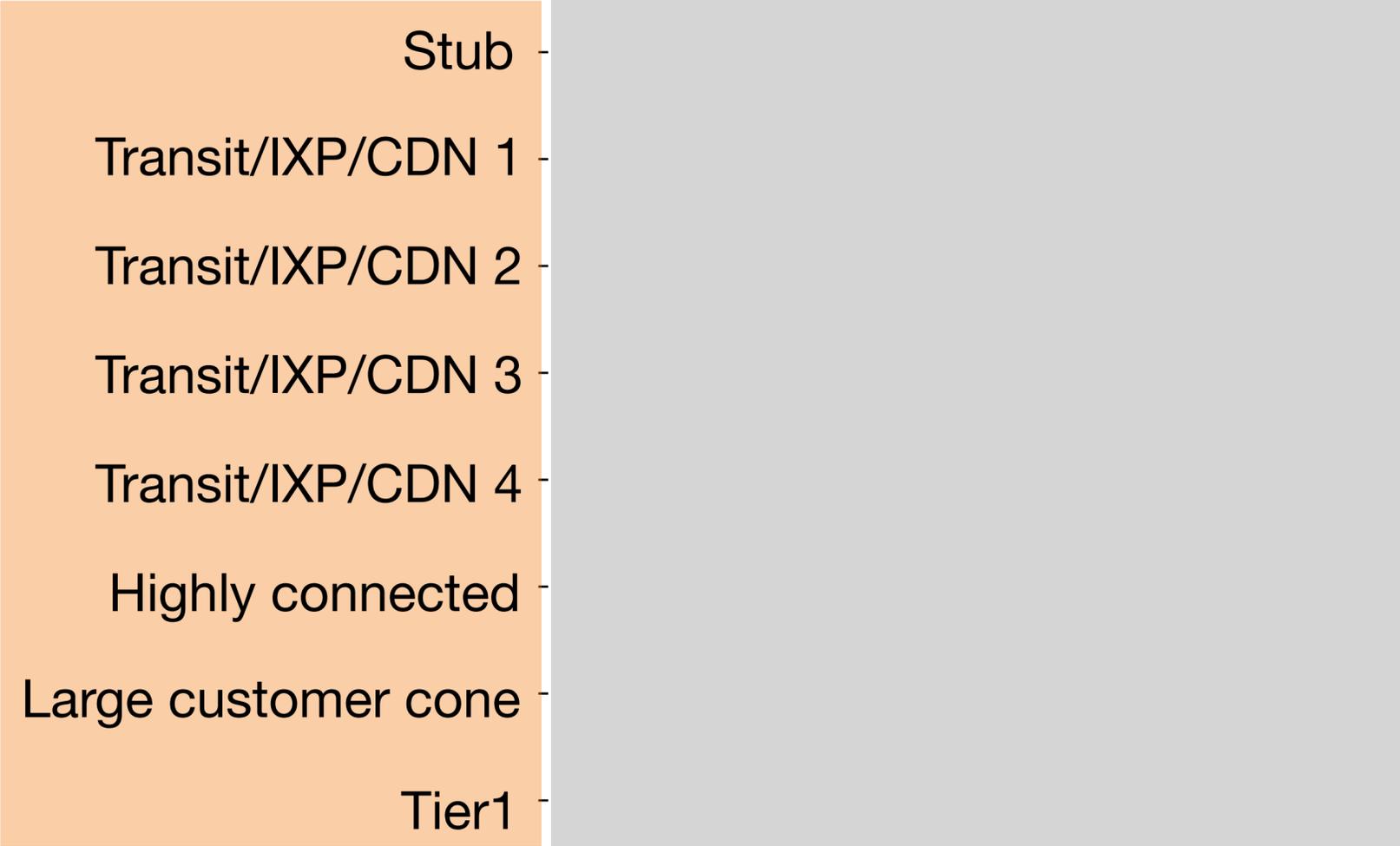
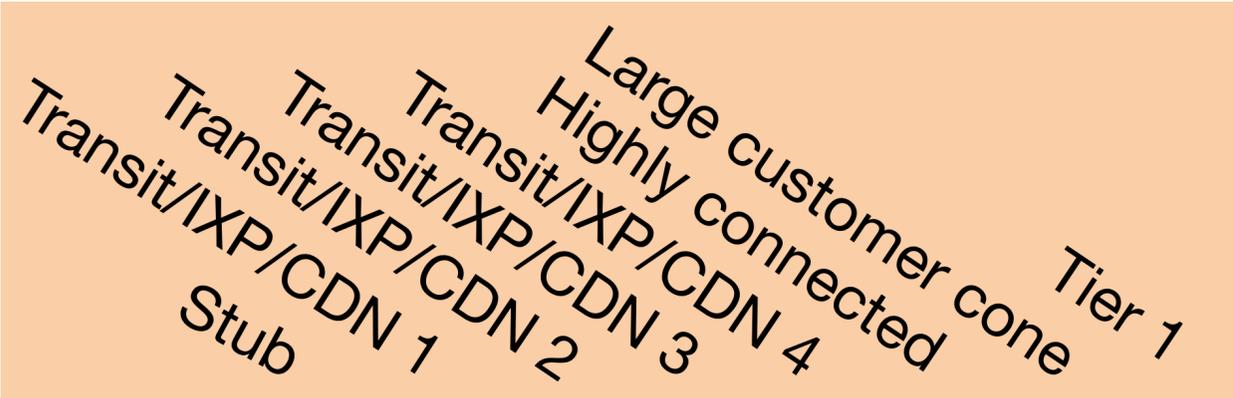
We focus on the **True Positive Rate (TPR)** and the **False Positive Rate (FPR)**

DFOH is **accurate** upon every attack scenario

Victim

True Positive Rate

Attacker



DFOH is **accurate** upon every attack scenario

Victim

True Positive Rate

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

DFOH is accurate upon every attack scenario

Victim

True Positive Rate

Attacker

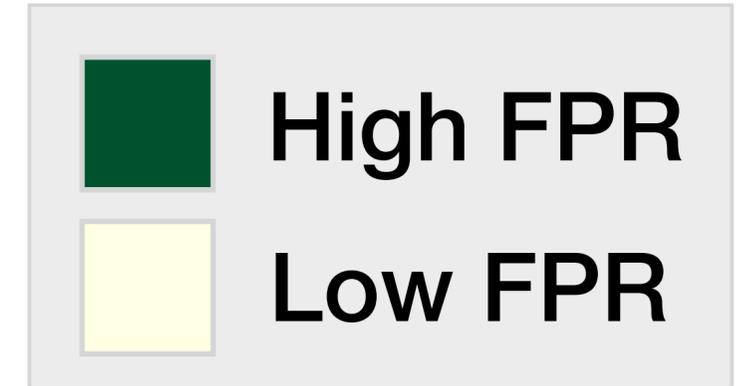
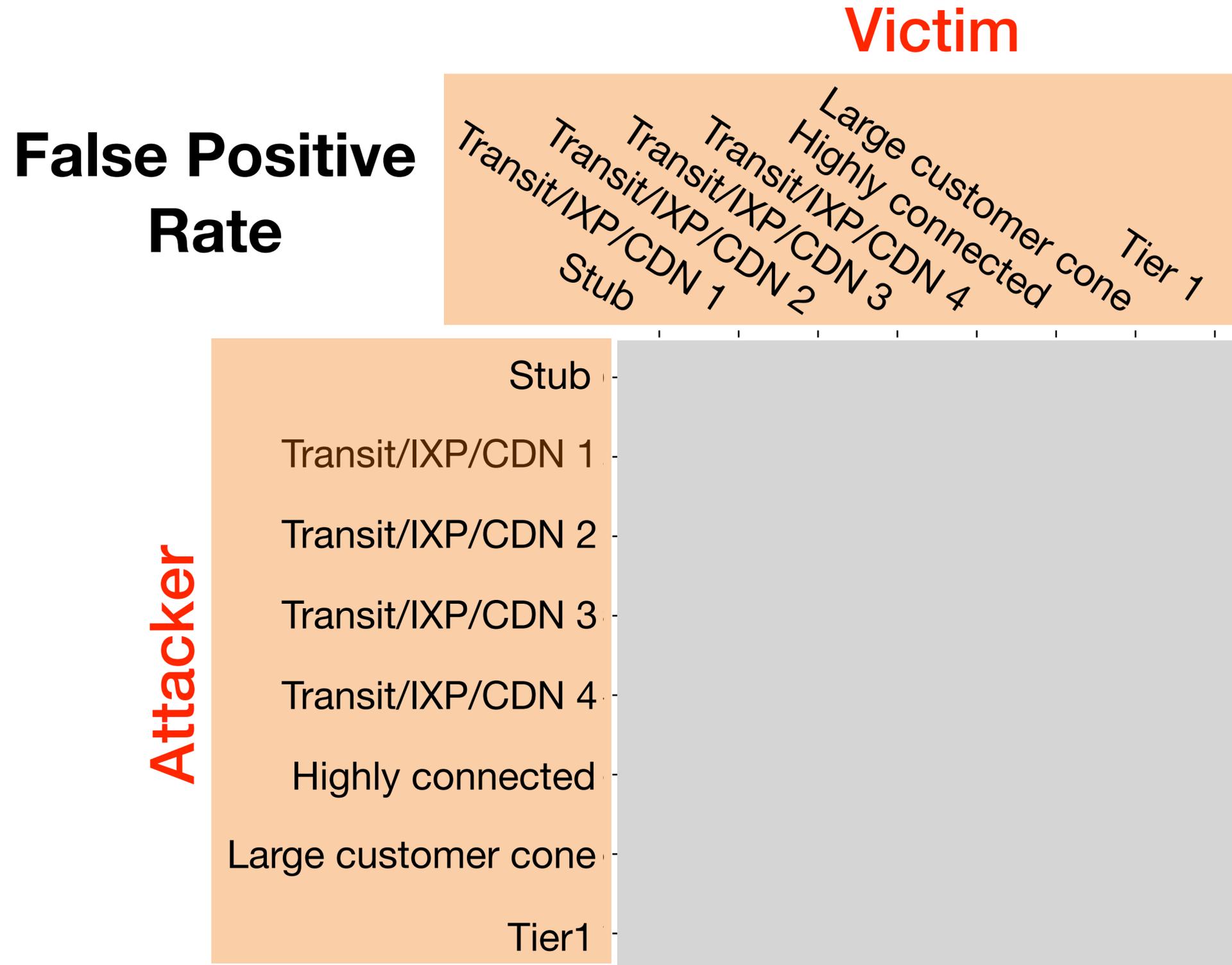
		Stub	Transit/IXP/CDN 1	Transit/IXP/CDN 2	Transit/IXP/CDN 3	Transit/IXP/CDN 4	Highly connected	Large customer cone	Tier 1
Attacker	Stub	0.97	0.86	0.91	0.96	0.94	0.95	0.95	0.84
	Transit/IXP/CDN 1	0.86	0.73	0.90	0.97	0.82	0.96	0.83	0.73
	Transit/IXP/CDN 2	0.91	0.90	0.85	0.95	0.99	0.99	0.90	0.83
	Transit/IXP/CDN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

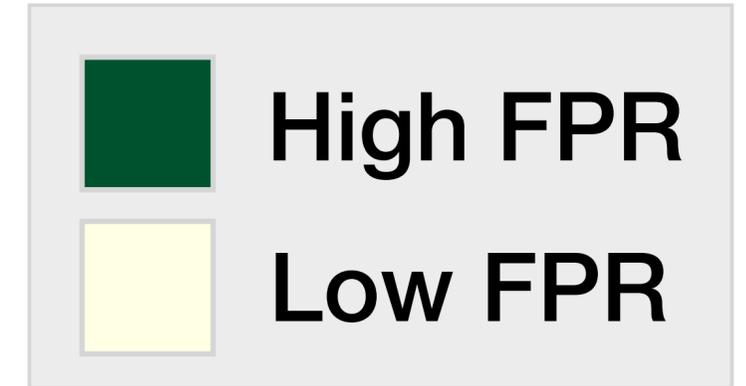
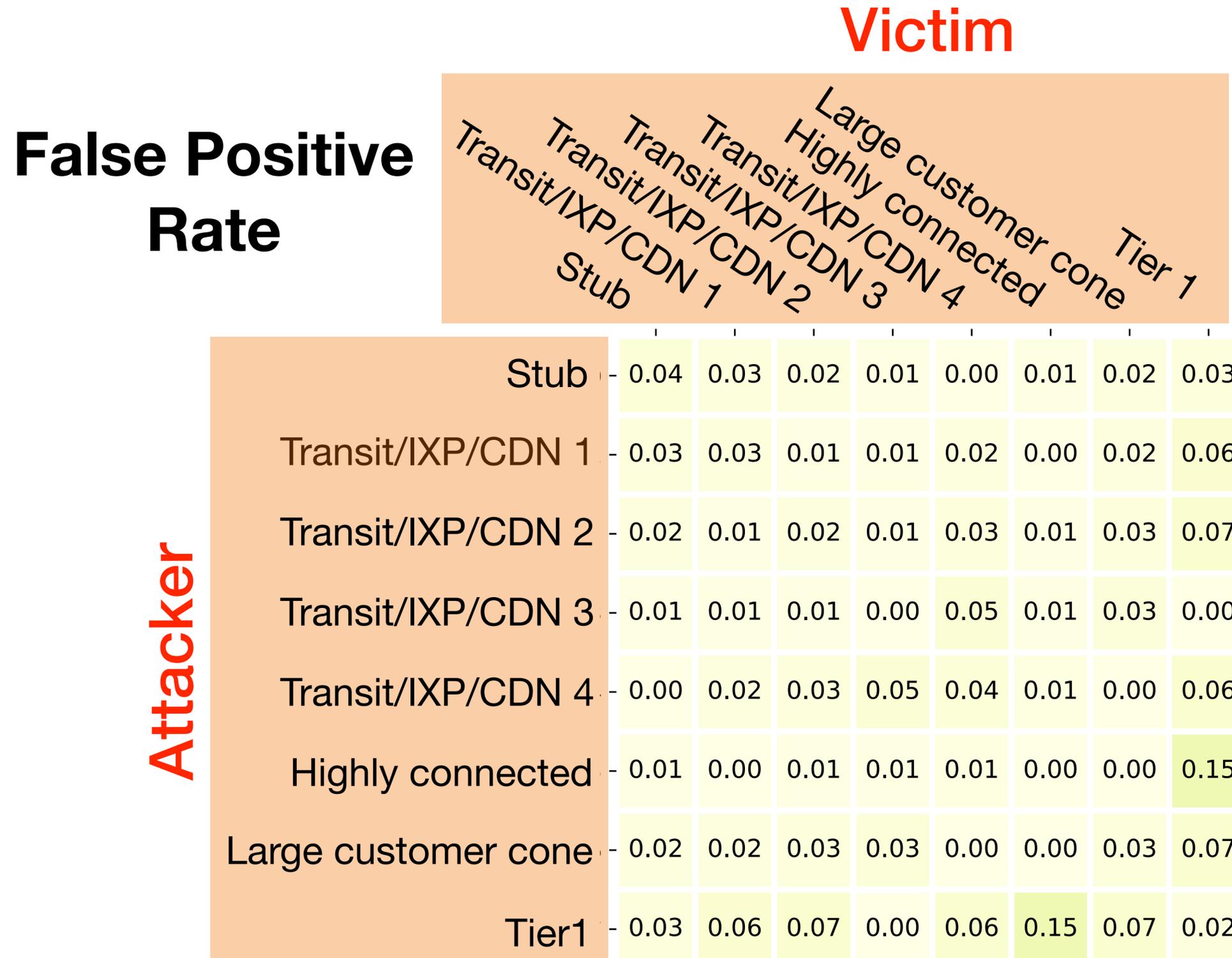
Low TPR

The minimum TPR is 0.73

DFOH is **accurate** upon every attack scenario



DFOH is **accurate** upon every attack scenario



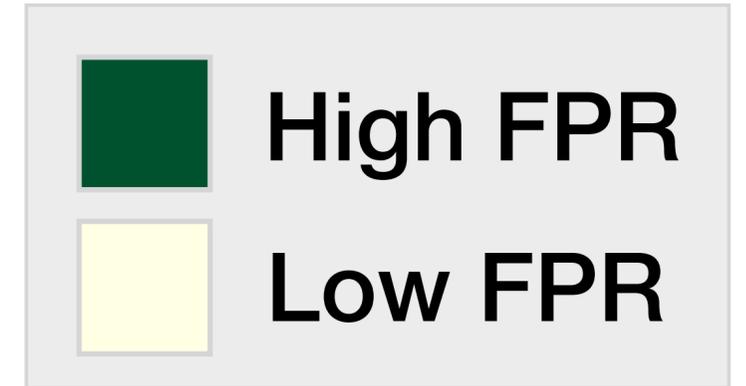
DFOH is **accurate** upon every attack scenario

False 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	0.04	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



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 Holterbach¹, Thomas Alftoy², Amresh Phokeer³, Alberto Dainotti¹, Cristel Peissier⁴
¹University of Strasbourg, ²Internet Society, ³Georgia Tech, ⁴UCLouvain

Abstract

Despite global efforts to secure Internet routing, attackers still successfully exploit the lack of strong BGP security mechanisms. This paper focuses on an attack vector that is frequently used: forged-origin hijacks, a type of BGP hijack where the attacker manipulates the AS path to make it immune to RPKI-ROV filters and appear as legitimate routing updates from a BGP monitoring standpoint. Our contribution is DFOH, a system that quickly and consistently detects forged-origin hijacks in the whole Internet. Detecting forged-origin hijacks boils down to inferring whether the AS path in a BGP route is legitimate or has been manipulated. We demonstrate that current state-of-art approaches to detect BGP anomalies are insufficient to deal with forged-origin hijacks. We identify the key properties that make the inference of forged AS paths challenging, and design DFOH to be robust against real-world factors (e.g., data biases). Our inference pipeline includes two key ingredients: (i) a set of strategically selected features, and (ii) a training scheme adapted to topological biases. DFOH detects 80% of the forged-origin hijacks within only 5 days. In addition, it only reports ≈ 17.5 suspicious cases every day for the whole Internet, a small number that allows operators to investigate the reported cases and take countermeasures.

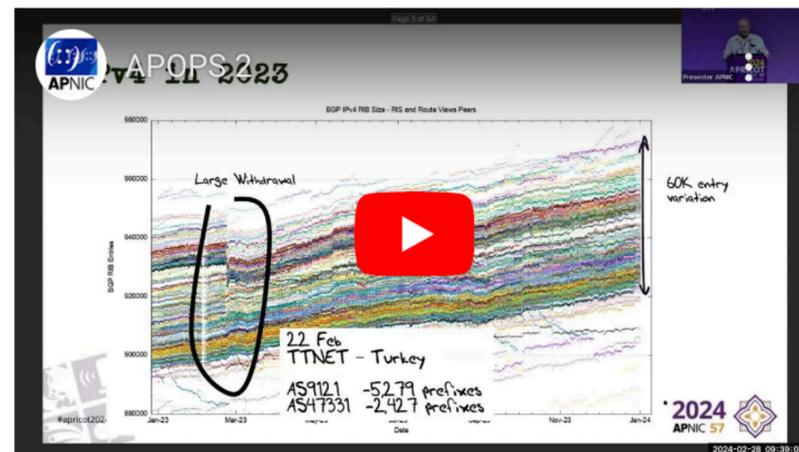
1 Introduction

On 3 February 2022, the cryptocurrency platform KLAYswap was targeted by hackers who stole \$1.9 million worth of digital assets [9]. More recently, on 17 August 2022, an attack to 48,000+ crypto-user wallets—affected 73 victims, who lost \$235,000 [4]. Both attacks were the result of a forged-origin BGP hijack, a type of routing hijack where the attacker announces forged AS paths towards a victim prefix by prepending the victim's origin AS number in order to make them appear legitimate. Clearly, BGP hijacking attacks are not a surprise anymore. They repeatedly make the headlines [1, 2] and are known as attack vectors to steal cryptocurrency [8], obtain bogus certificates [15], or deanonymize Tor users [6].

The vulnerability they exploit is simply the result of BGP being designed without security in mind. An attacker can manipulate every attribute in a BGP message (including the AS path and its origin AS) and illegitimately announce a prefix owned by its victim so as to divert the traffic to its network. Proactive solutions against BGP hijacks are being gradually deployed. However, forged-origin hijacks have been left uncovered by such solutions—despite these attacks being actively used in the wild. In fact, network operators attempt to proactively thwart BGP hijacks by configuring their routers to filter hijacked routes [46] using (i) RPKI-based Route Origin Validation (ROV) and (ii) data from Internet Routing Registries (IRR). Unfortunately, RPKI-ROV filters do not help to detect forged-origin hijacks, since the forged origin in the AS path is actually valid, while IRR-based filters are known to be inaccurate, incomplete [27], and far too often missing given the increasing number of observed BGP hijacks [7]. Today, network operators do not have many options left other than waiting for the deployment of new security extensions to BGP to consistently prevent forged-origin hijacks [44]. Such deployment—if it will happen at all—might take more than a decade, as in the case of RPKI-ROV [1].

In this paper, we present DFOH, the first locally-deployable system that widely, quickly, and accurately Detects Forged-Origin Hijacks on the Internet. With a single deployment of DFOH on a commodity server, any attacker performing a forged-origin hijack is likely to be quickly detected, the hijack publicly reported, and the victim immediately notified. Being aware of the attack, the victim can apply countermeasures and the community can take actions to prevent similar attacks from happening again. Additionally, DFOH can detect past attacks, allowing the community to measure the frequency of such attacks or profile forged-origin hijackers.

DFOH is a passive system that processes the AS paths observed in publicly collected BGP routes to detect forged-origin hijacks. The problem of detecting forged-origin hijacks can be reduced to identifying whether a link between two ASes is real or fake. Unfortunately, there are multiple reasons why two ASes might connect, whereas there is no simple



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)

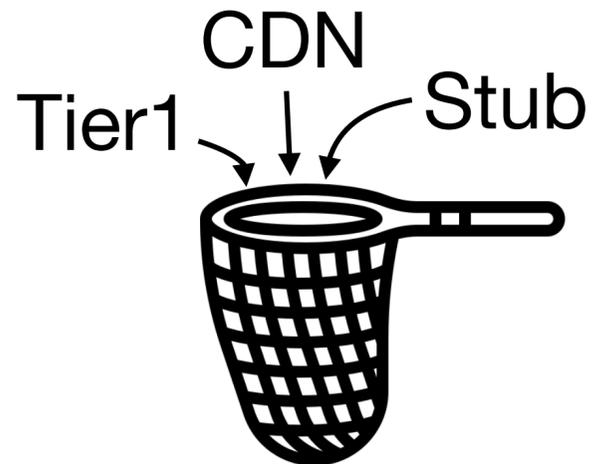
***DFOH*: A System to Detect Forged-Origin Hijacks**



DFOH runs in a commodity server



DFOH detects hijacks on the whole Internet

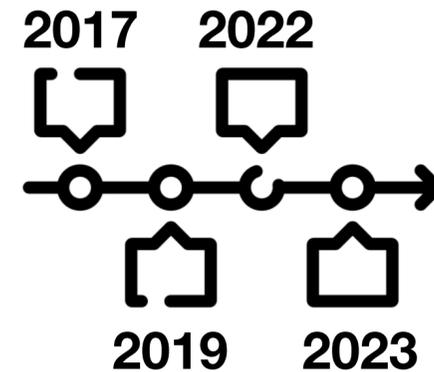


DFOH is accurate in every attack scenario

DFOH: A System to Detect Forged-Origin BGP Hijacks



DFOH runs in a commodity server



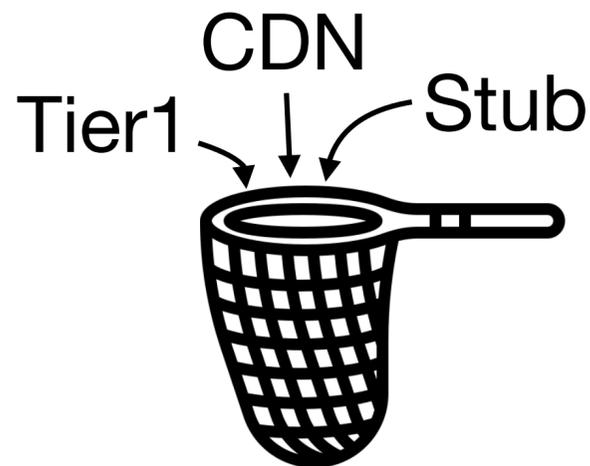
DFOH detects past hijacks



DFOH detects hijacks on the whole Internet



DFOH provides near-real-time detection



DFOH is accurate in every attack scenario



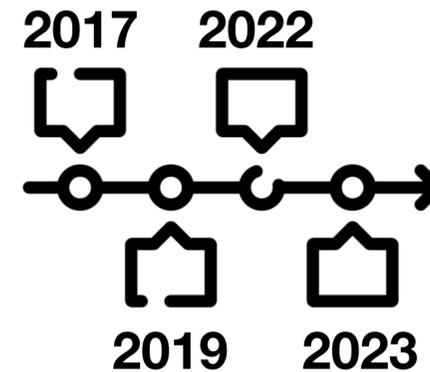
DFOH is robust against adversarial inputs

DFOH: A System to Detect Forged-Origin Hijacks

<https://dfoh.uclouvain.be>



DFOH runs in a commodity server



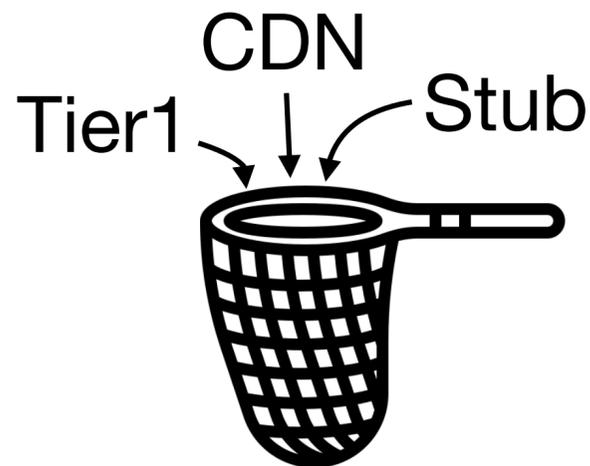
DFOH detects past hijacks



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