

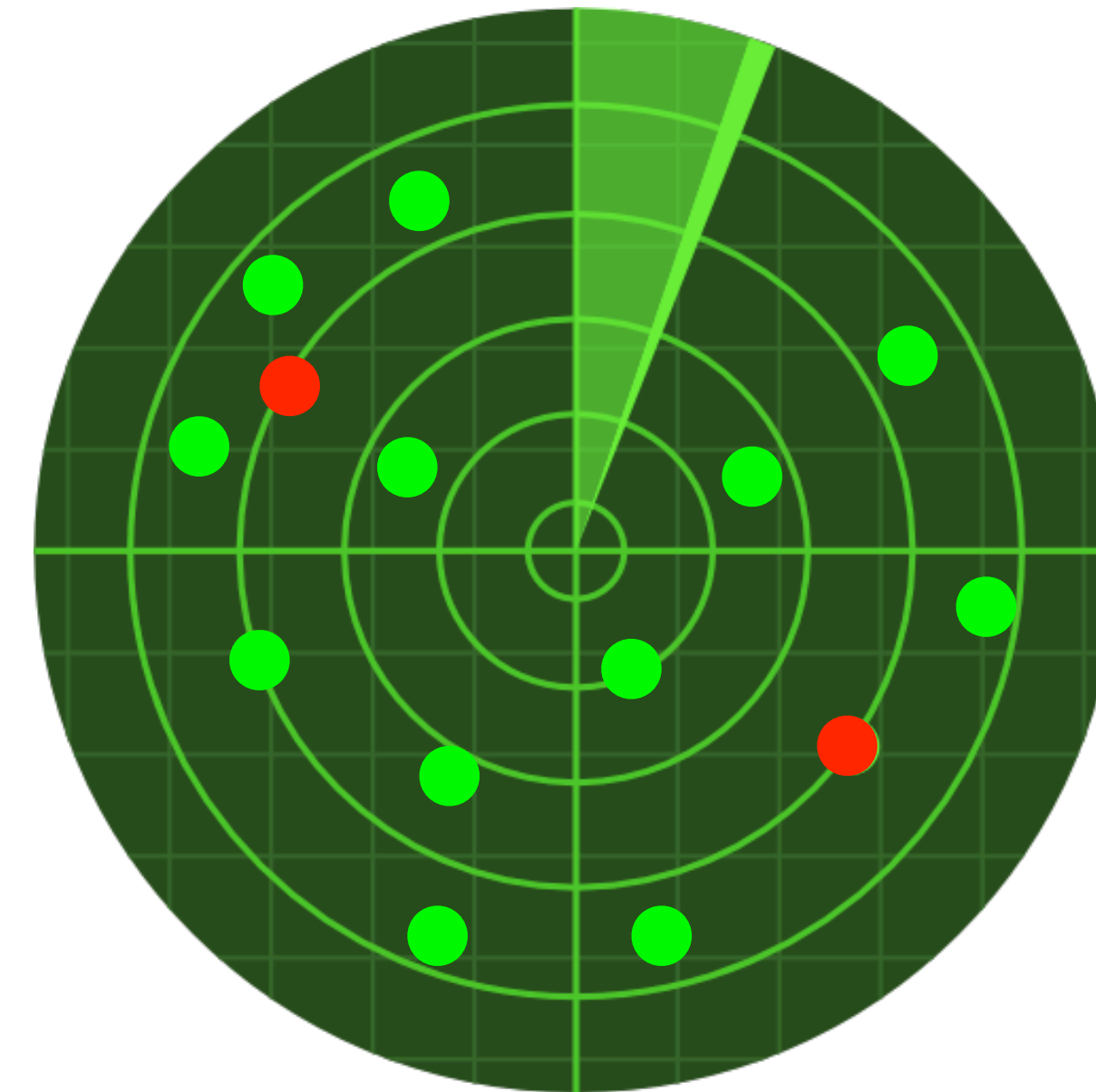
A System to Detect Forged-Origin Hijacks

<https://dfoh.uclouvain.be>

Thomas Holterbach

MANRS Ambassador 2023

University of Strasbourg



Joint work with:

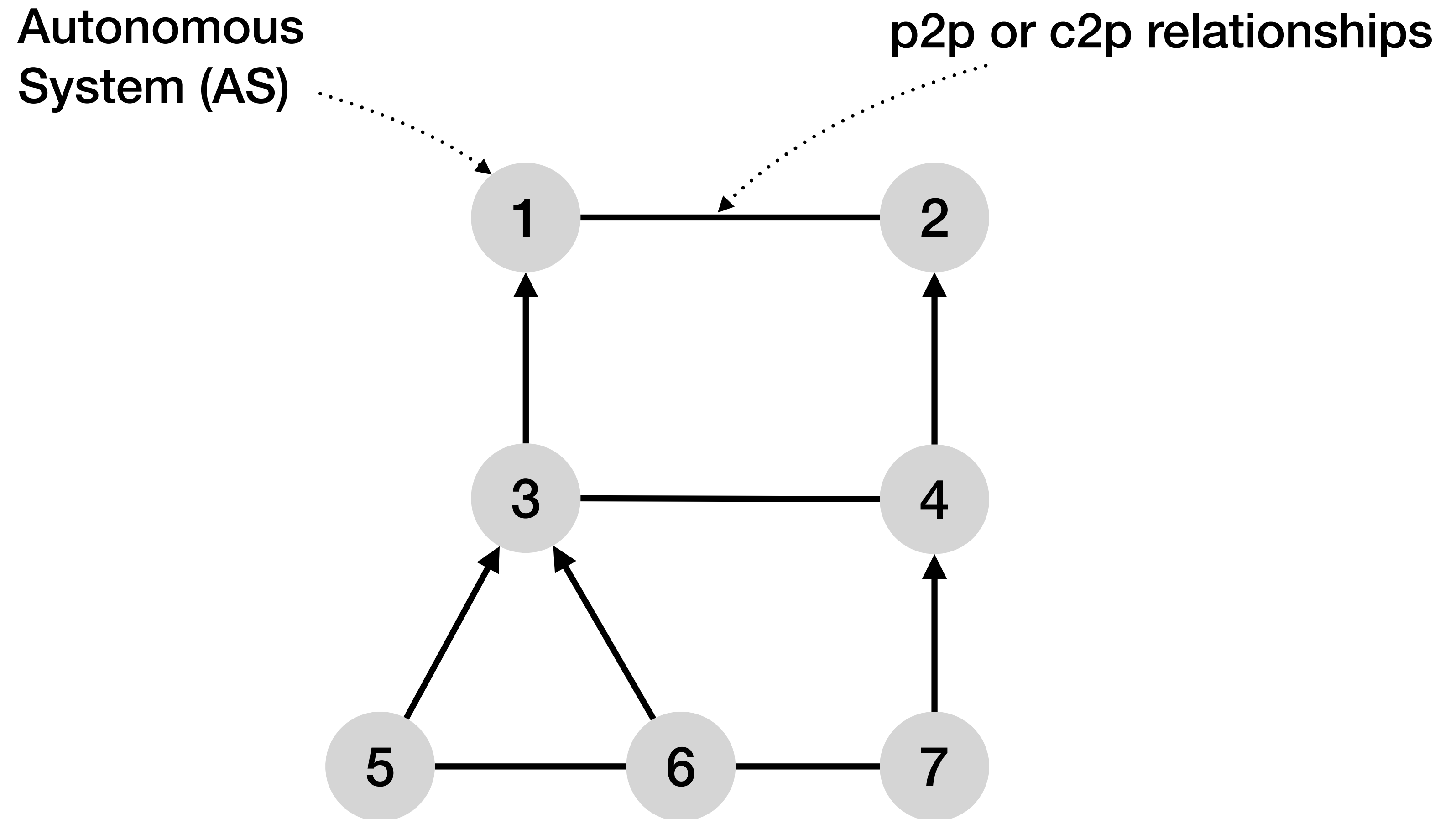
Thomas Alfroy

Alberto Dainotti

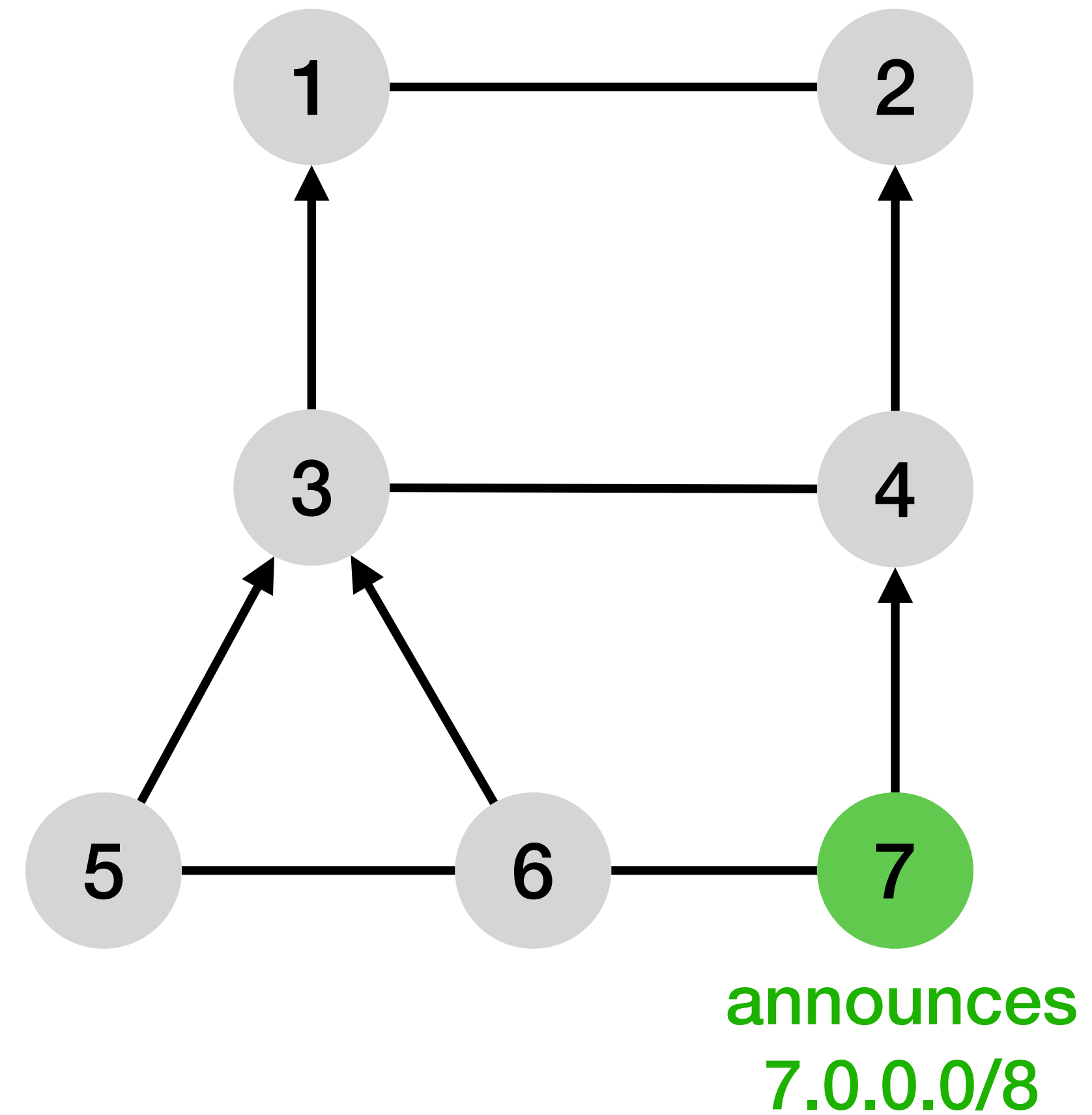
Amreesh D. Phokeer

Cristel Pelsser

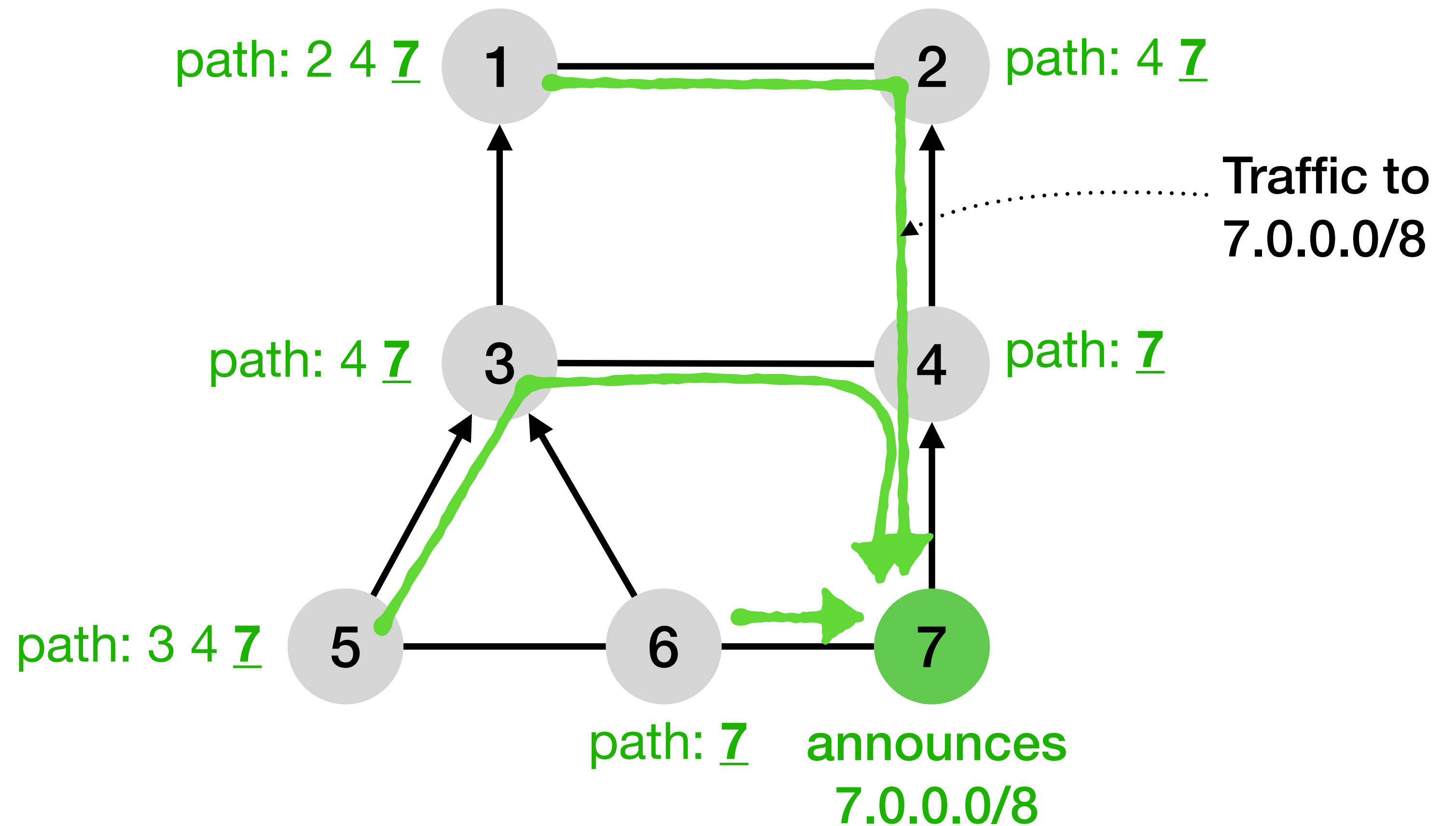
Internet routing (BGP) is vulnerable to traffic hijacking



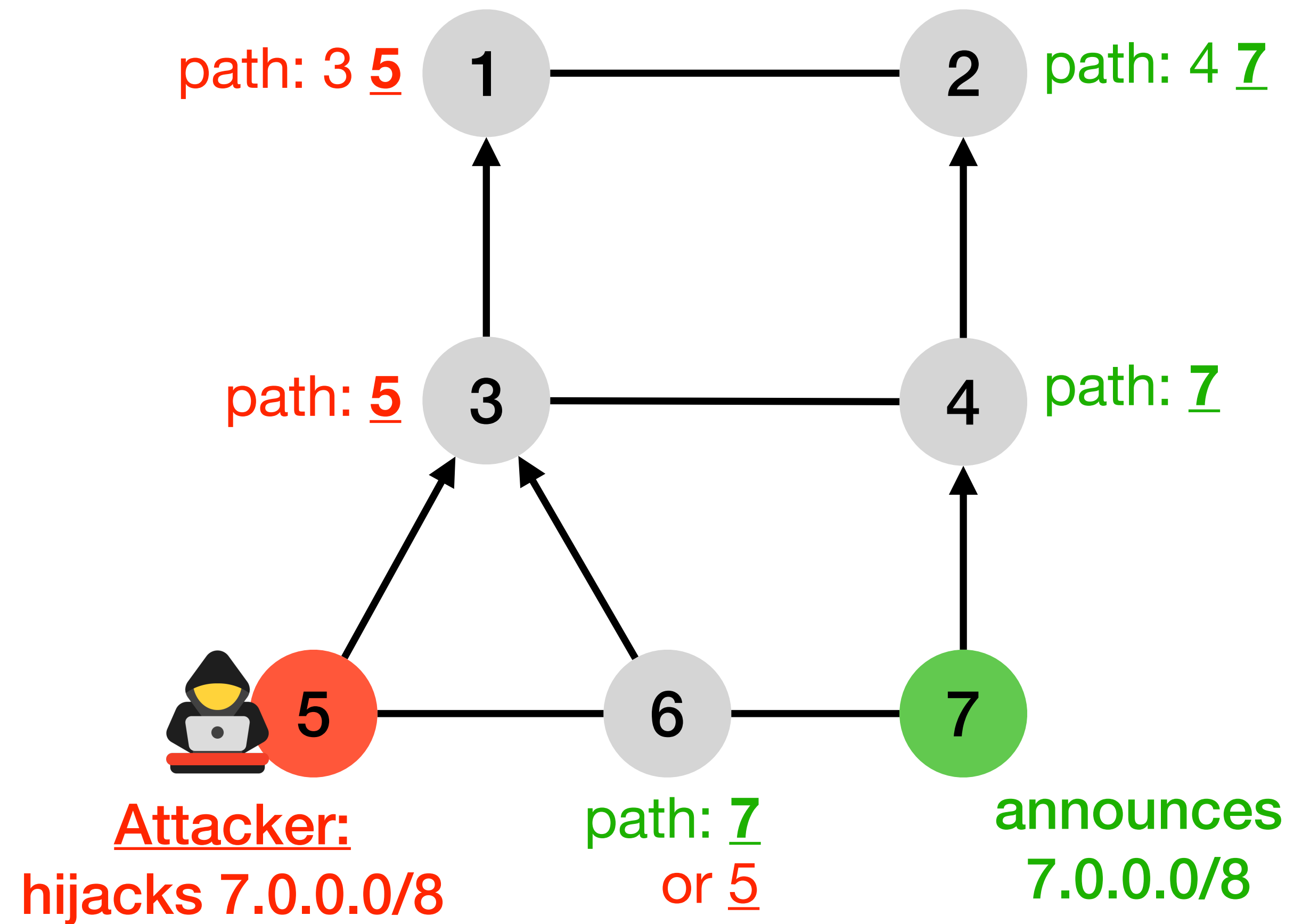
Internet routing (BGP) is vulnerable to traffic hijacking



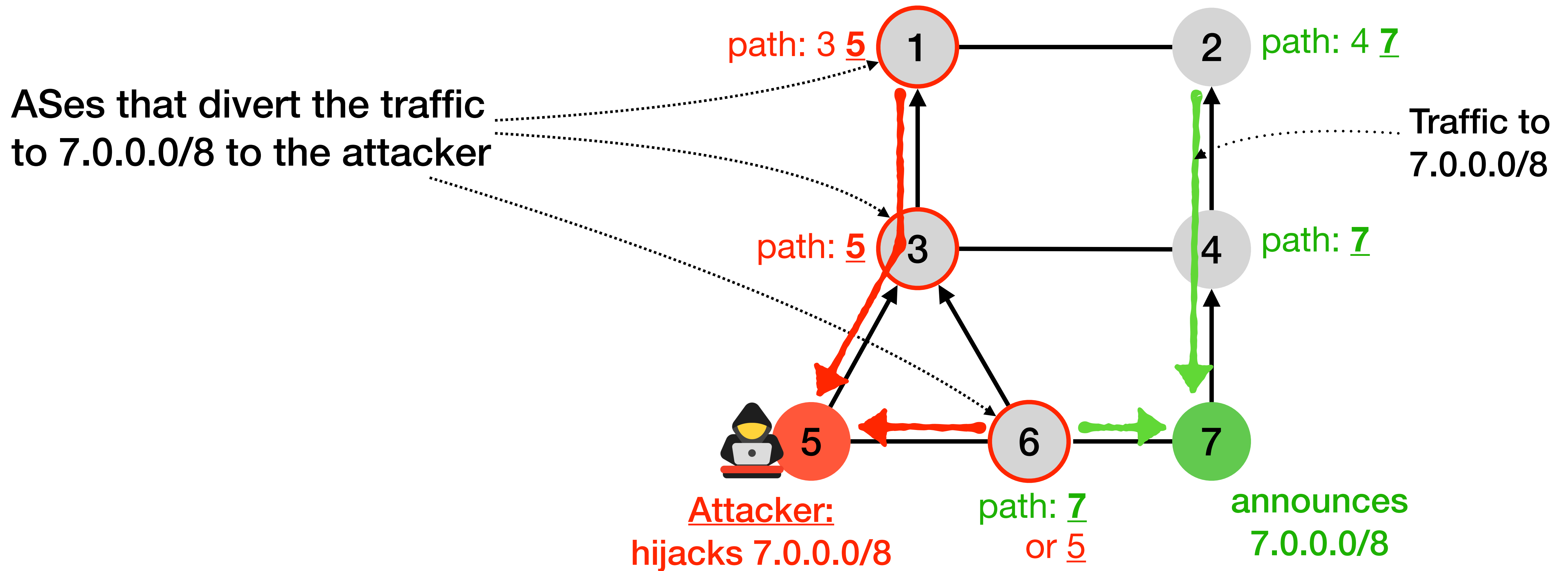
Internet routing (BGP) is vulnerable to traffic hijacking



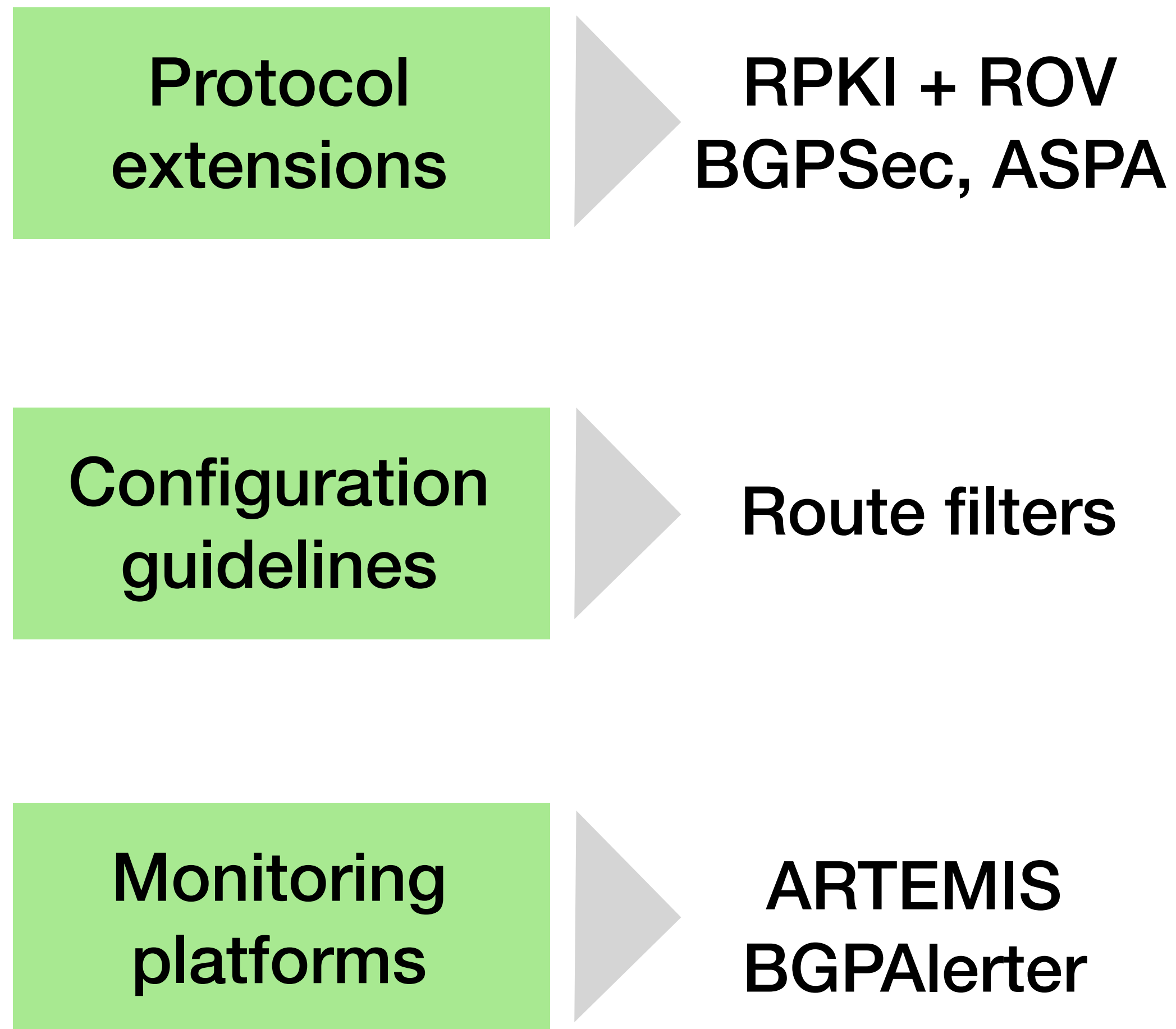
Internet routing (BGP) is vulnerable to traffic hijacking



Internet routing (BGP) is vulnerable to traffic hijacking

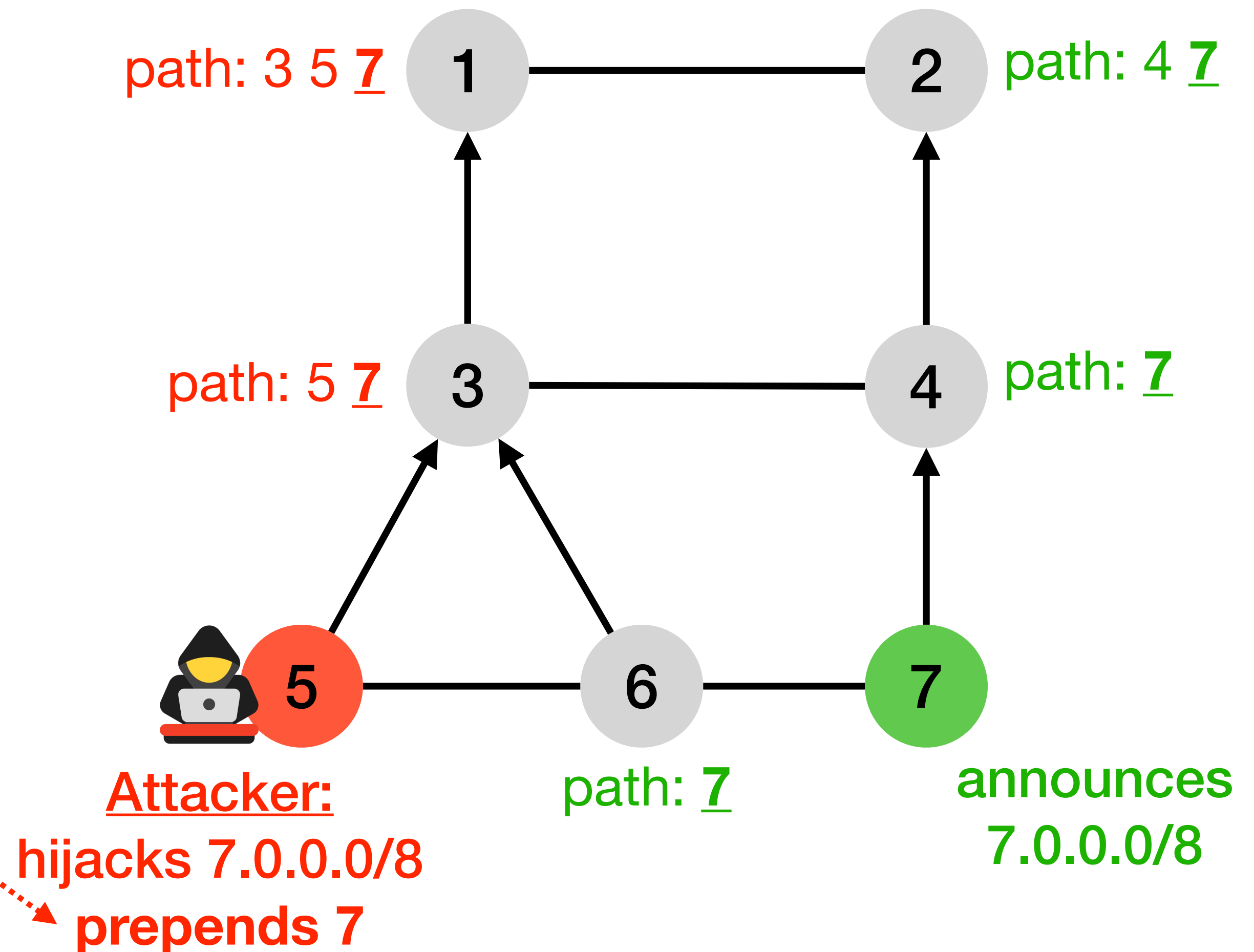


Fortunately, there are defenses against BGP hijacking



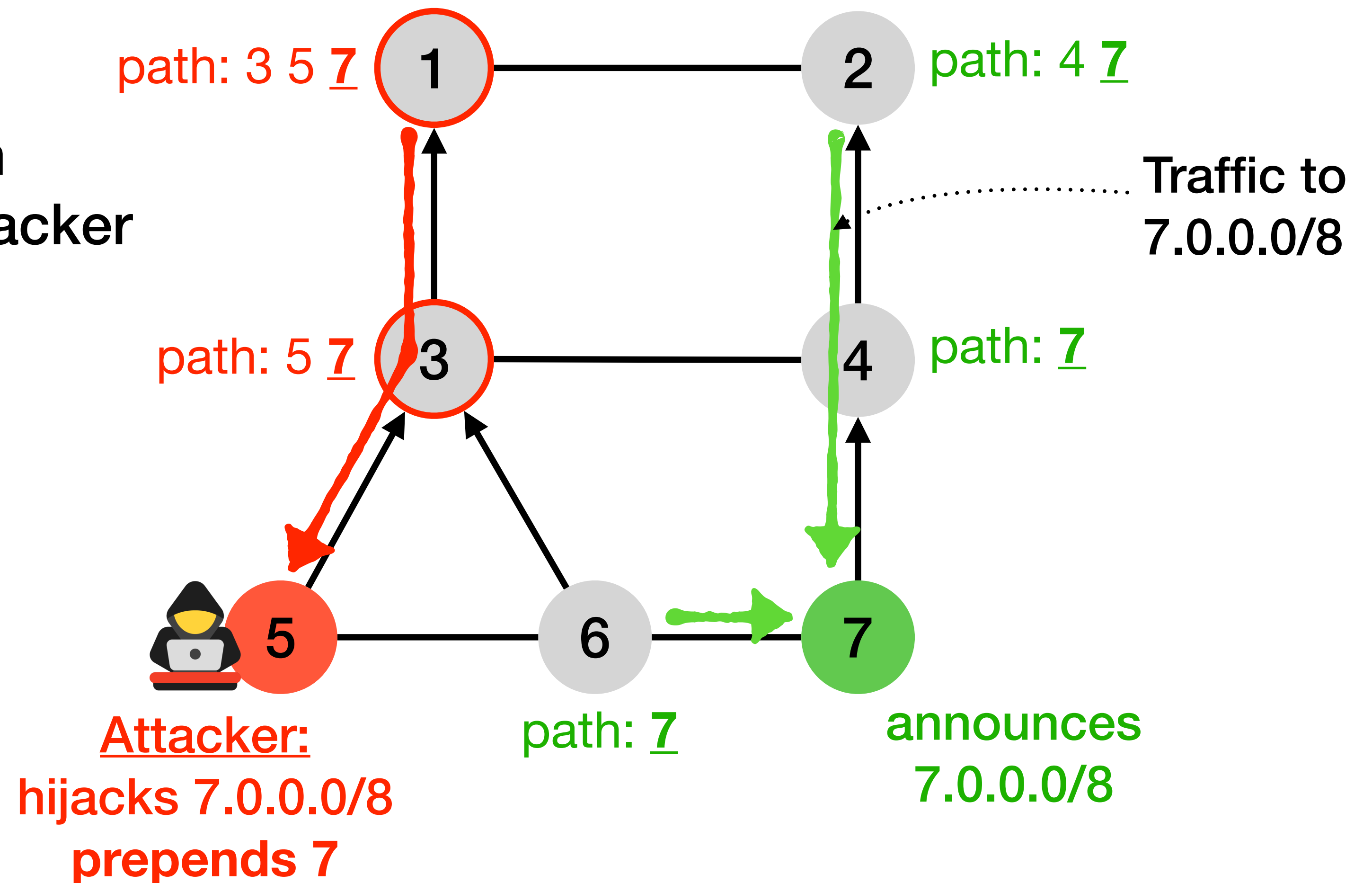
Despite the efforts, BGP is *still* vulnerable to **forged-origin hijacks**

The attacker prepends the legitimate AS number to the AS path

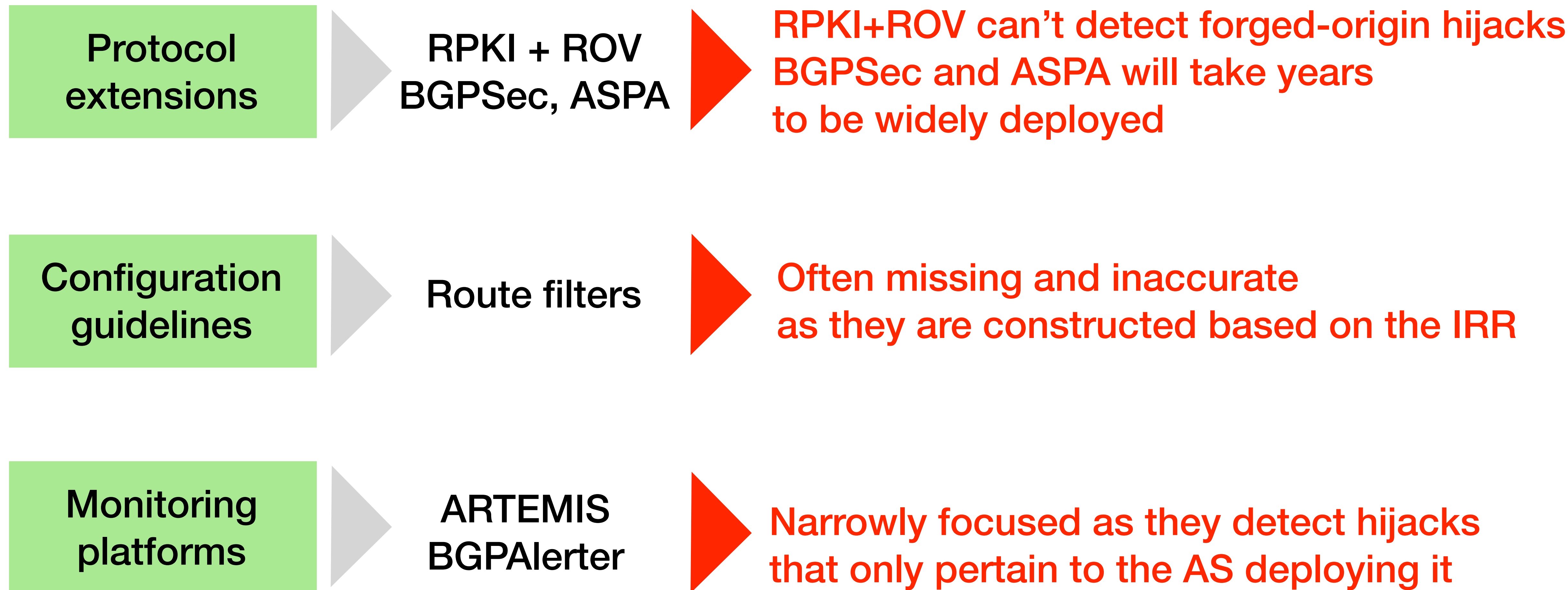


Despite the efforts, BGP is *still* vulnerable to **forged-origin hijacks**

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



Existing defenses poorly neutralise forged-origin hijacks



Forged-origin hijacks are actively used by attackers

August 17, 2022

The Record.
Recorded Future® News

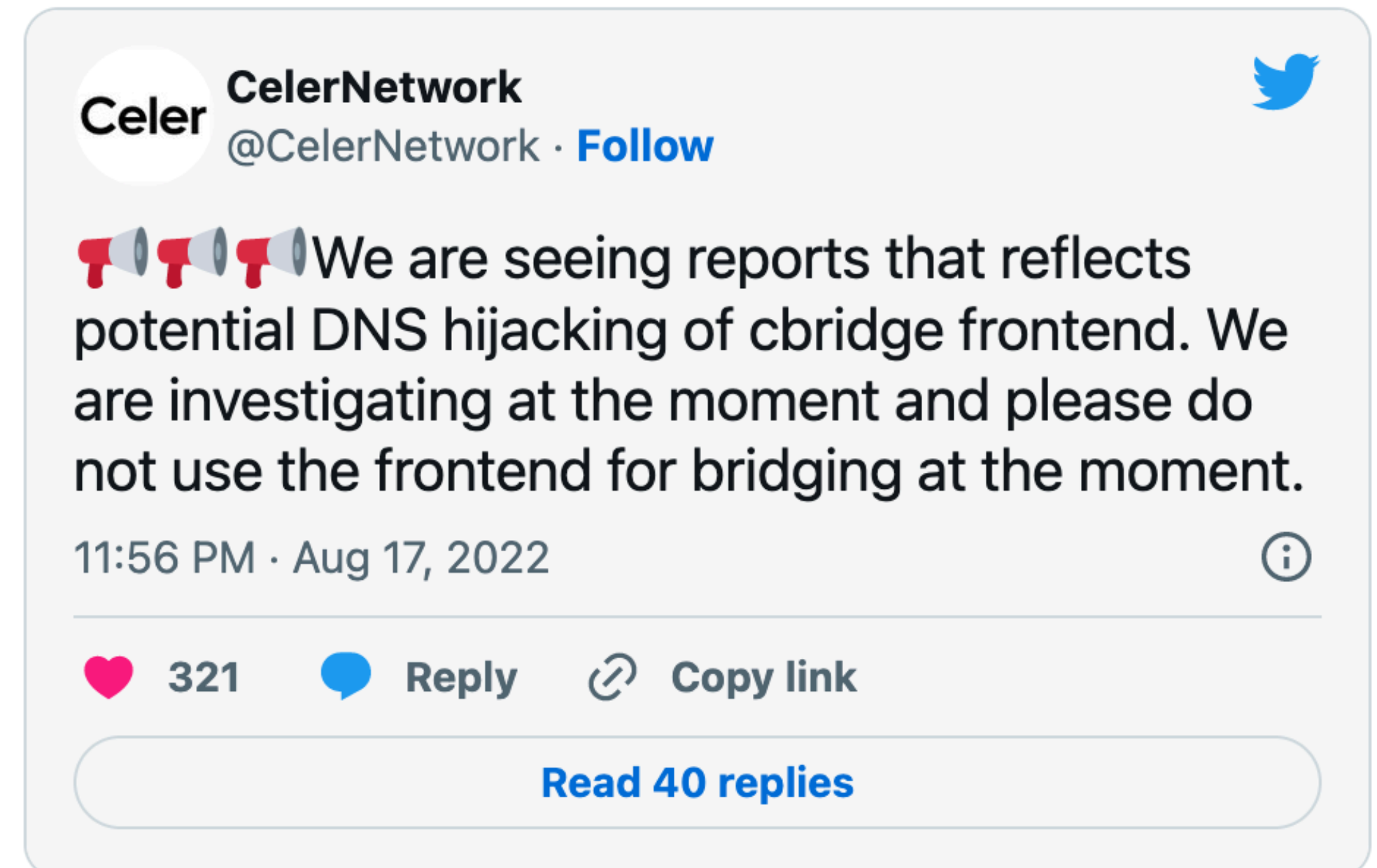
February 3, 2022

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

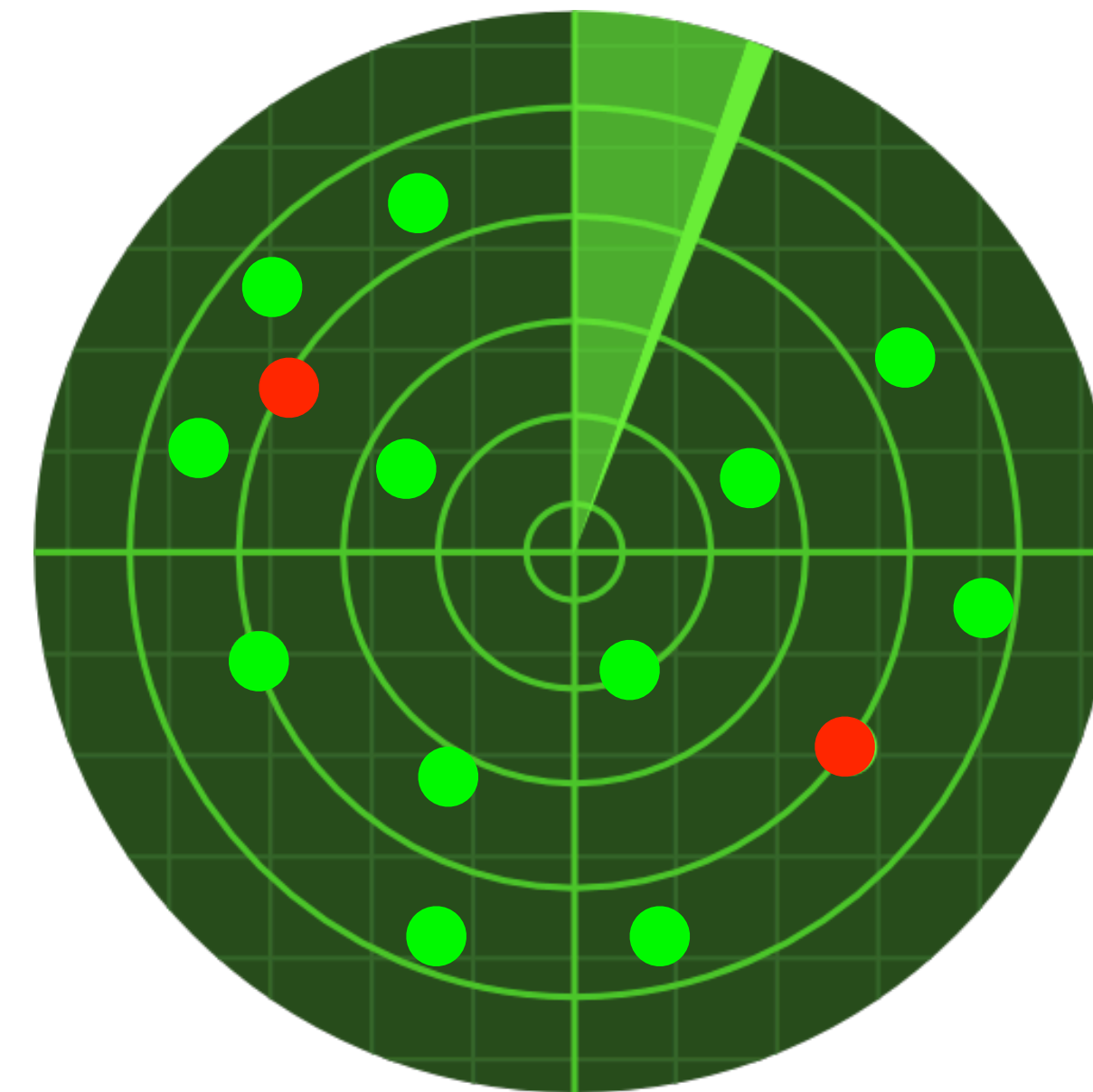
 321  Reply  Copy link

[Read 40 replies](#)

Both attacks are the result of a forged-origin hijack

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

Thomas Holterbach
University of Strasbourg



Joint work with:

Thomas Alfroy

Amreesh D. Phokeer

Alberto Dainotti

Cristel Pelsser

Outline

DFOH's main challenge

DFOH's inference pipeline

DFOH's inferences are accurate

DFOH is up and running

Outline

DFOH's main challenge

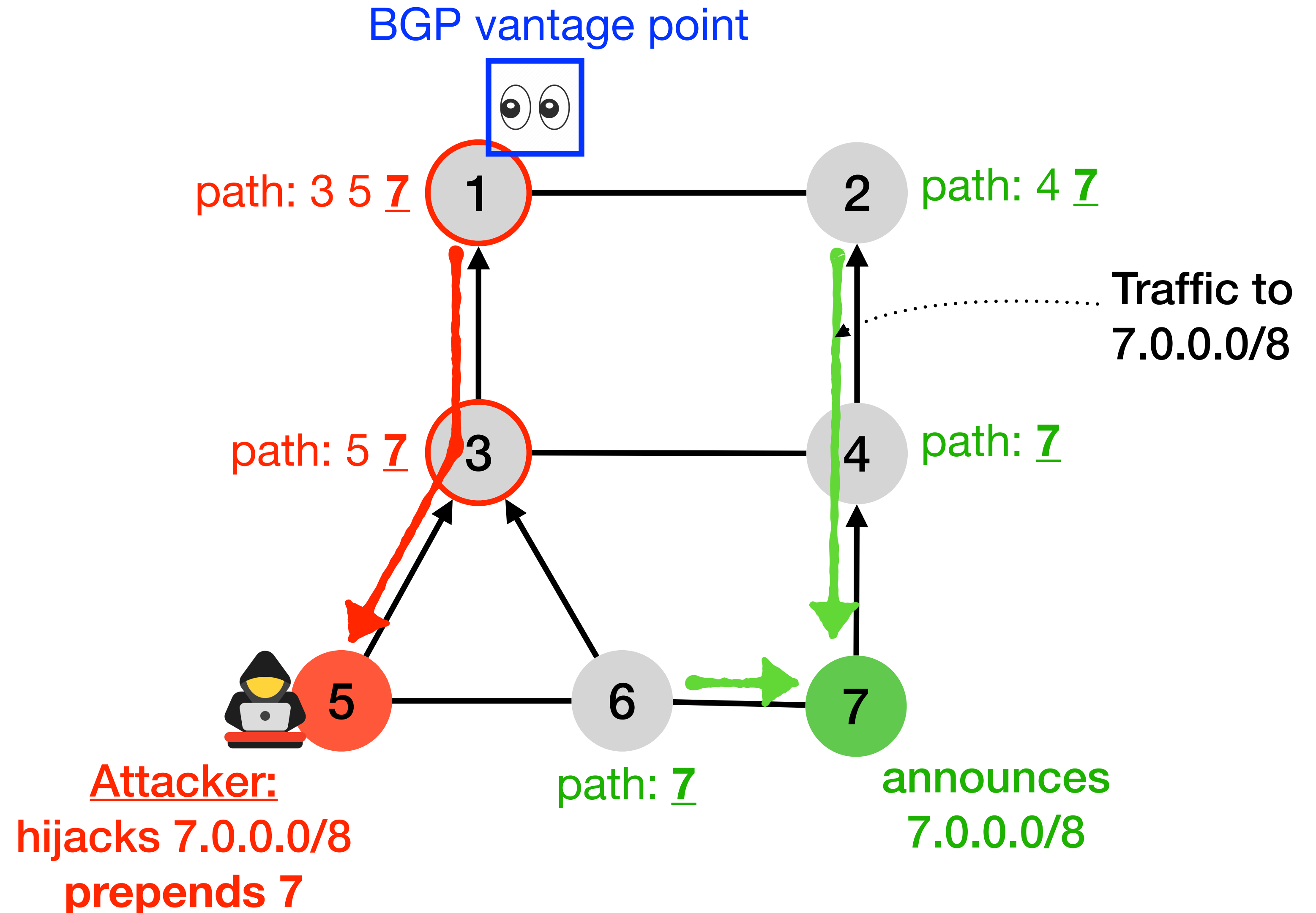
is to detect **fake** AS links

DFOH's inference pipeline

DFOH's inferences are accurate

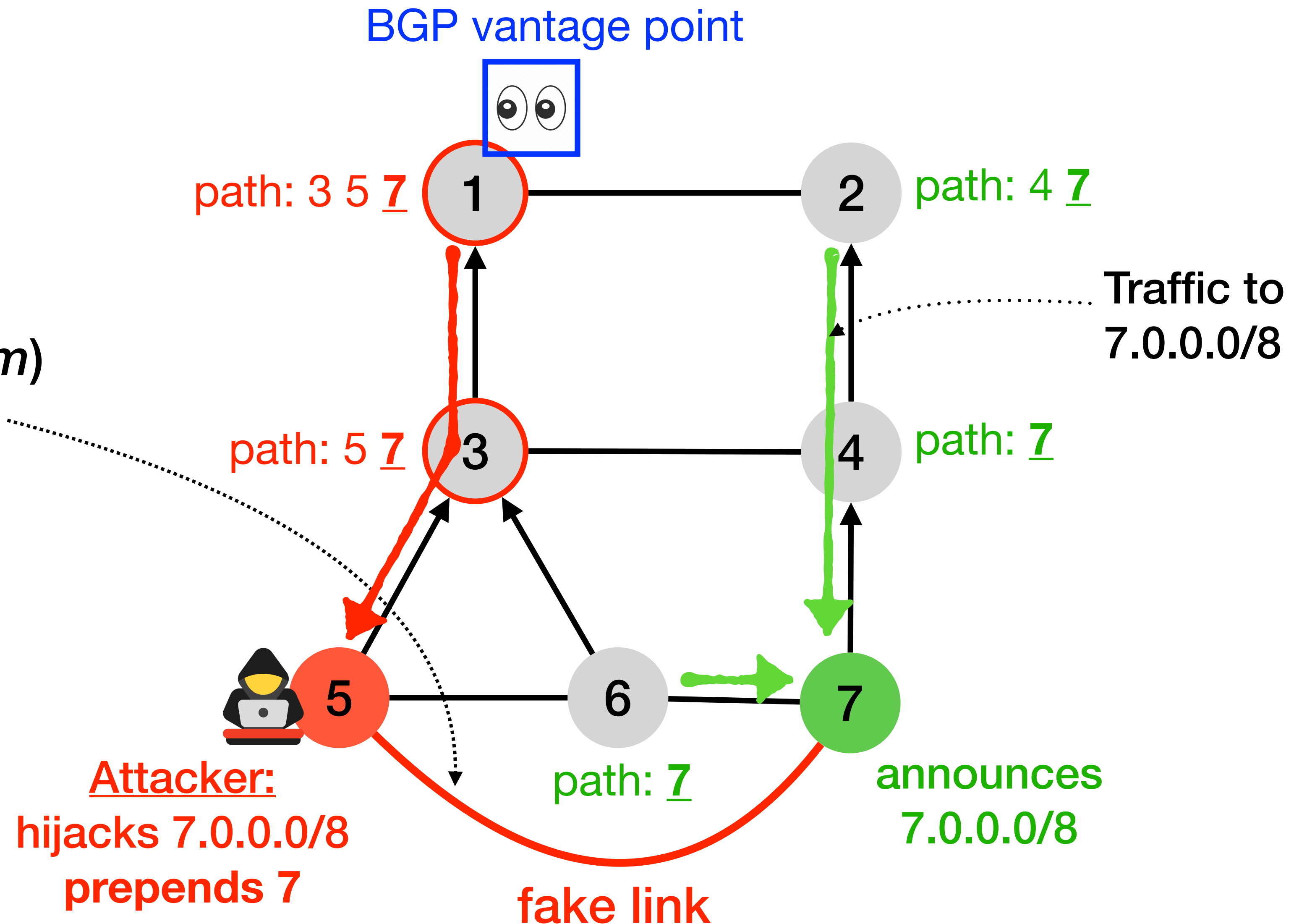
DFOH is up and running

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

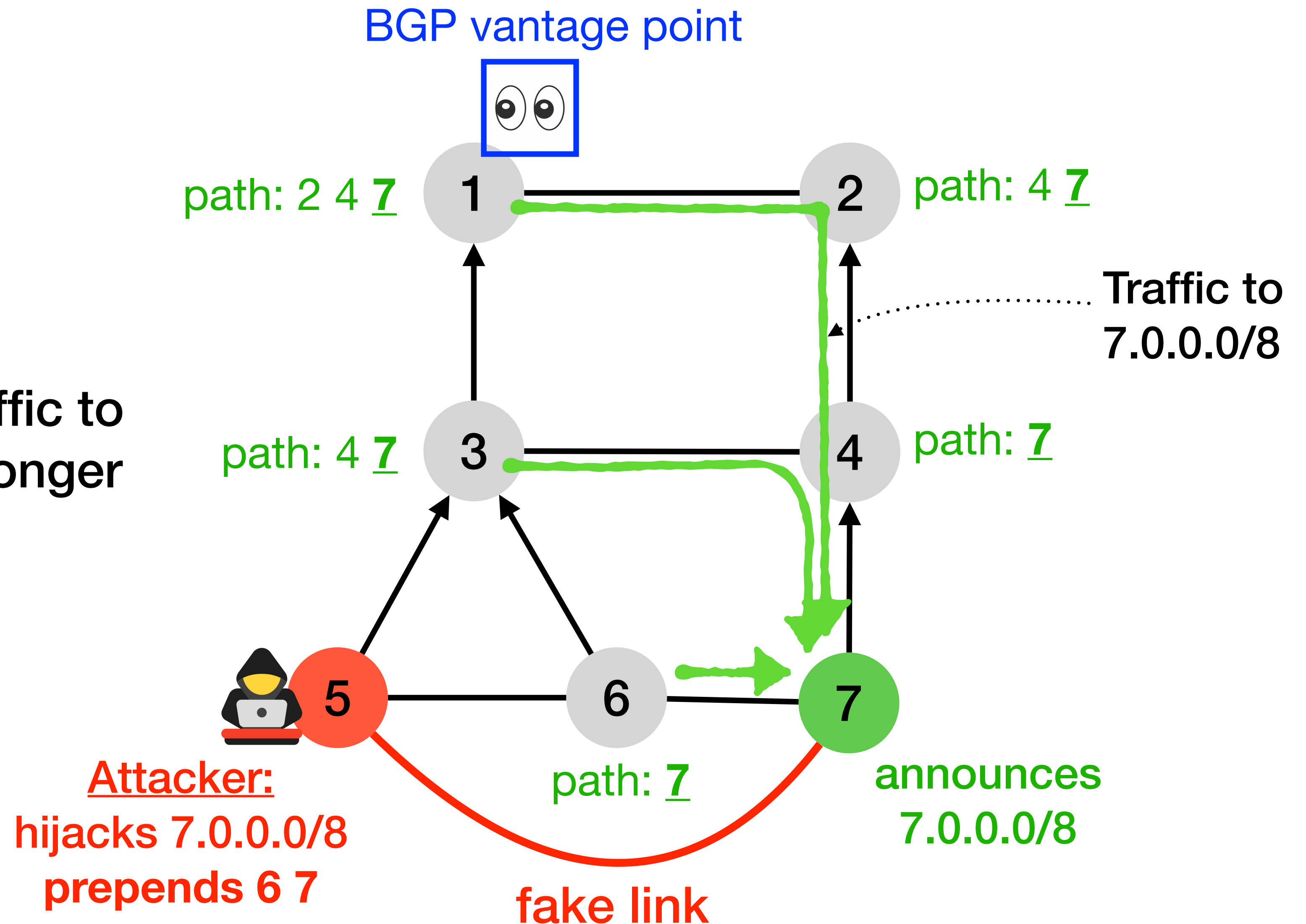
Upon the attack:
AS5 (attacker) and AS7 (victim)
appear directly connected



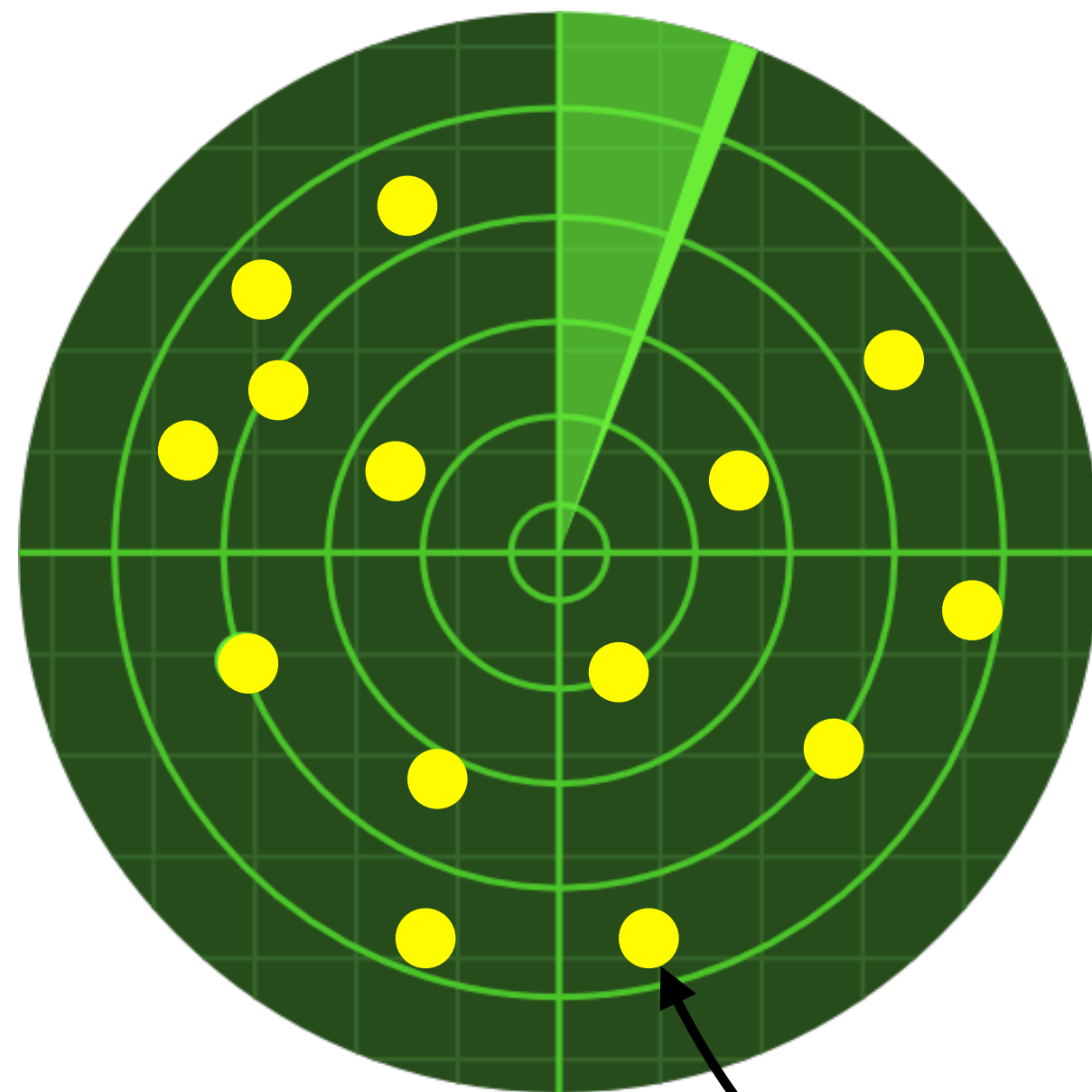
An attacker **cannot escape** from creating a new AS link without hampering the effectiveness of its attack

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

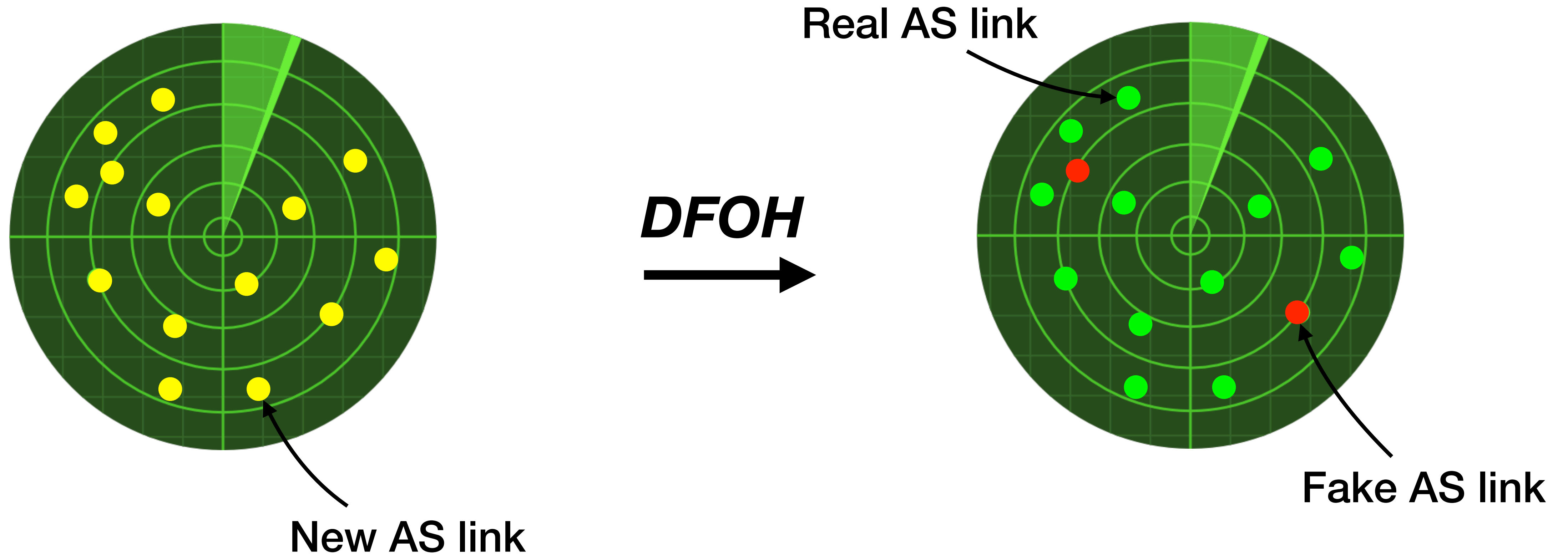


New AS link

**We find 166 new AS links
every day (median)**

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

discriminates fake AS links from the real ones

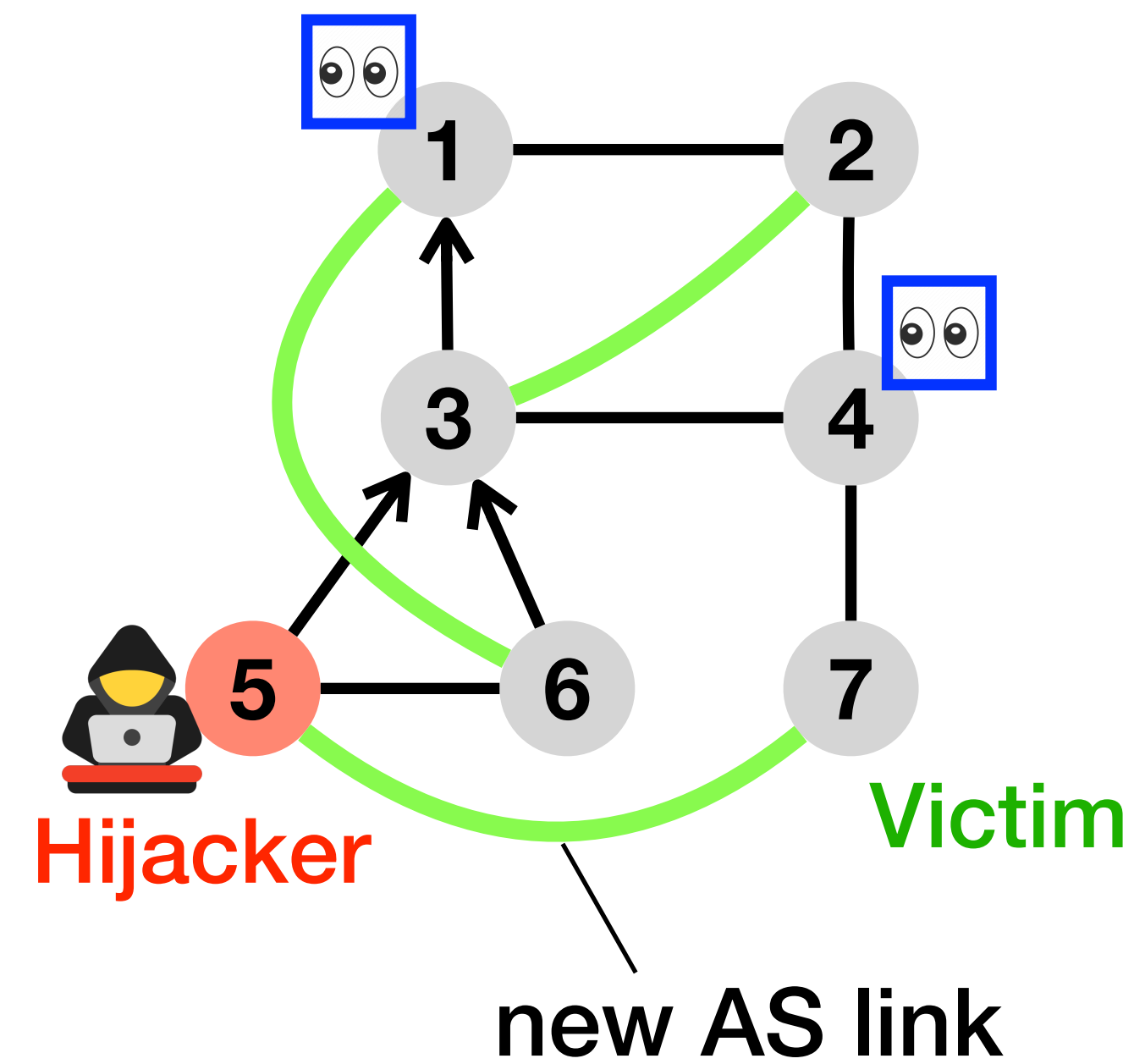
DFOH's inferences are accurate

DFOH is up and running

DFOH's fake AS links inference algorithm comprises three steps



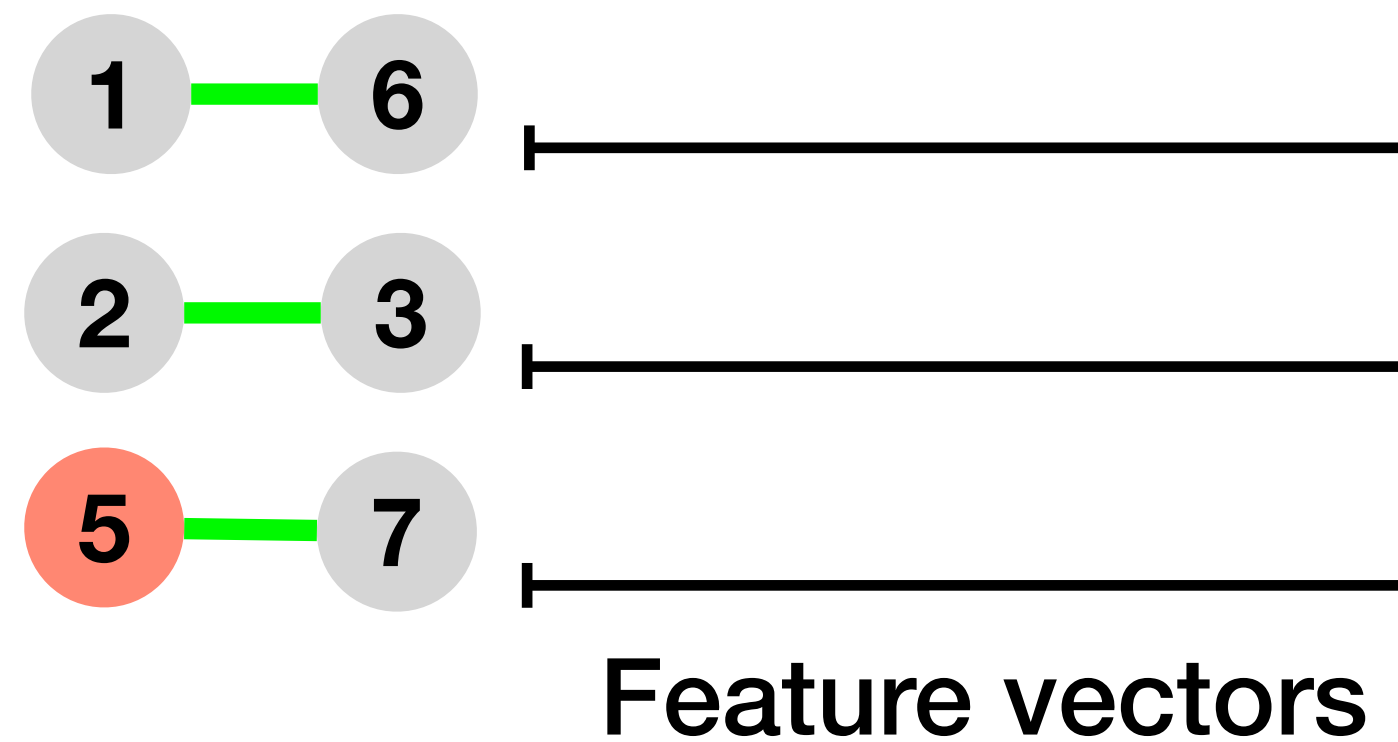
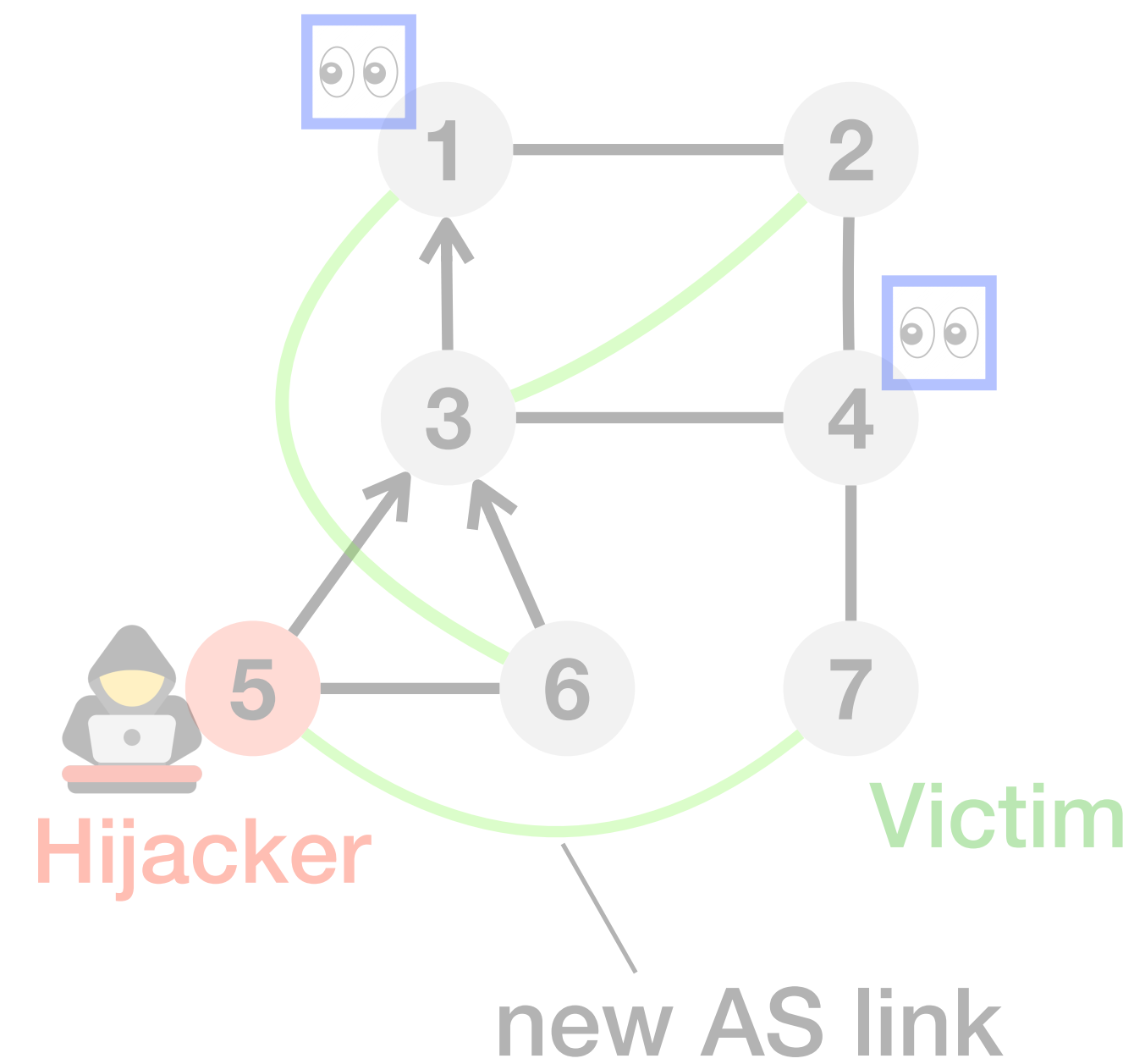
Vantage point



DFOH's fake AS links inference algorithm comprises three steps



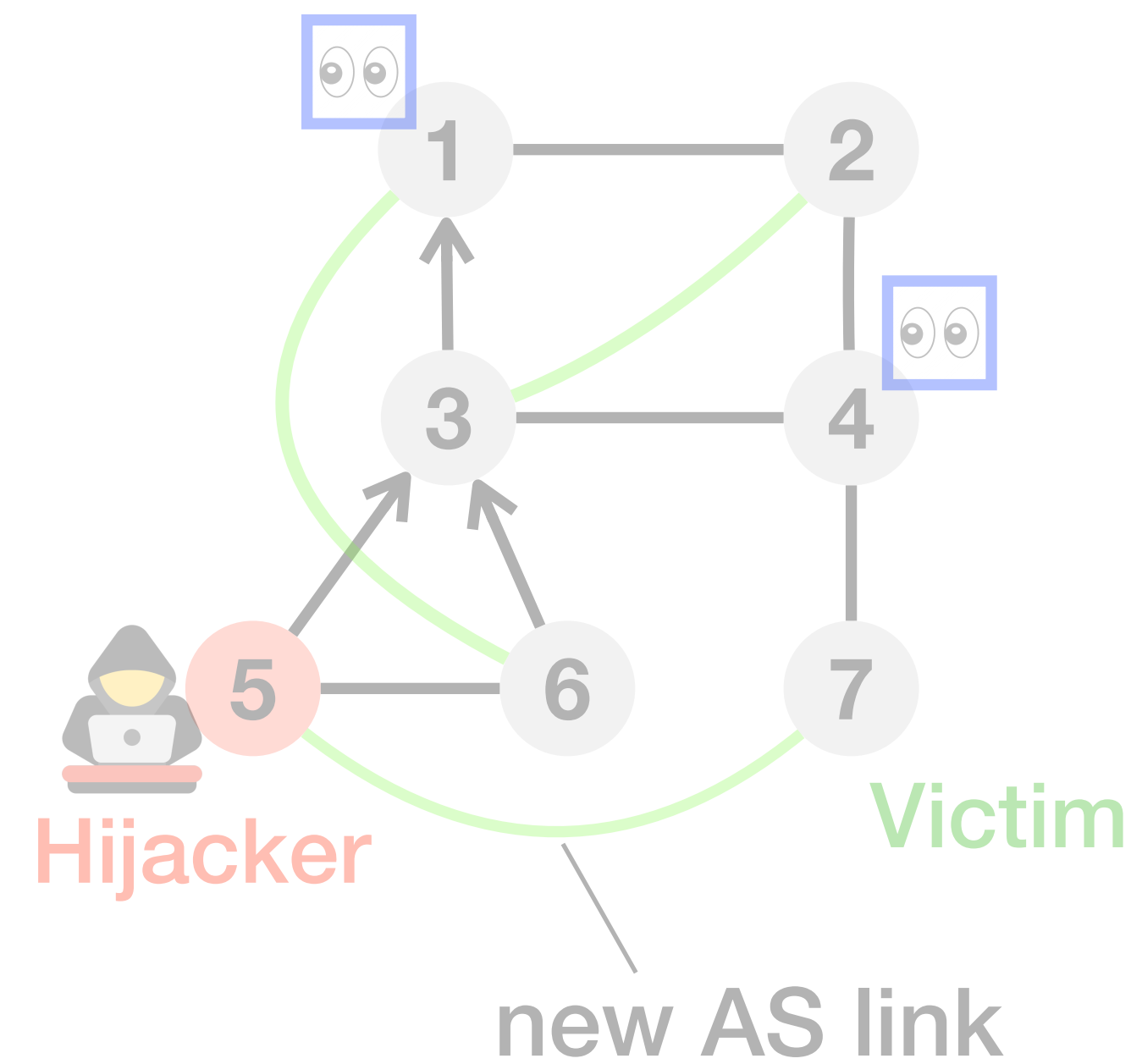
Vantage point



DFOH's fake AS links inference algorithm comprises three steps

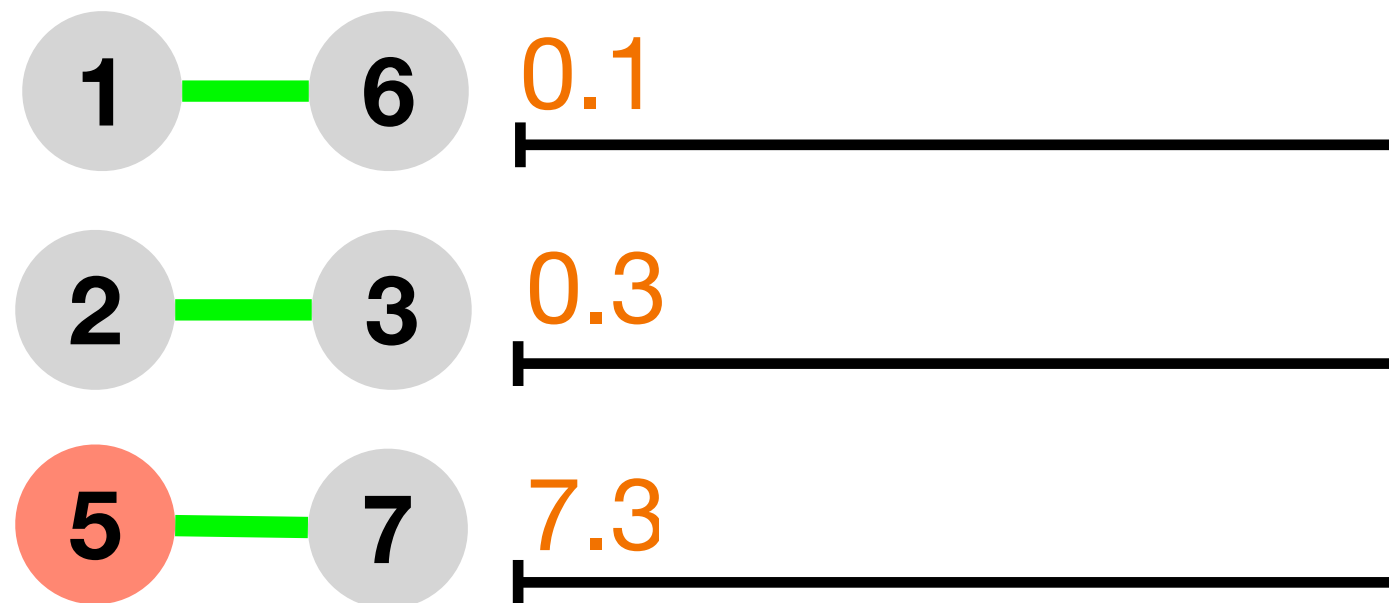


Vantage point



Feature categories:

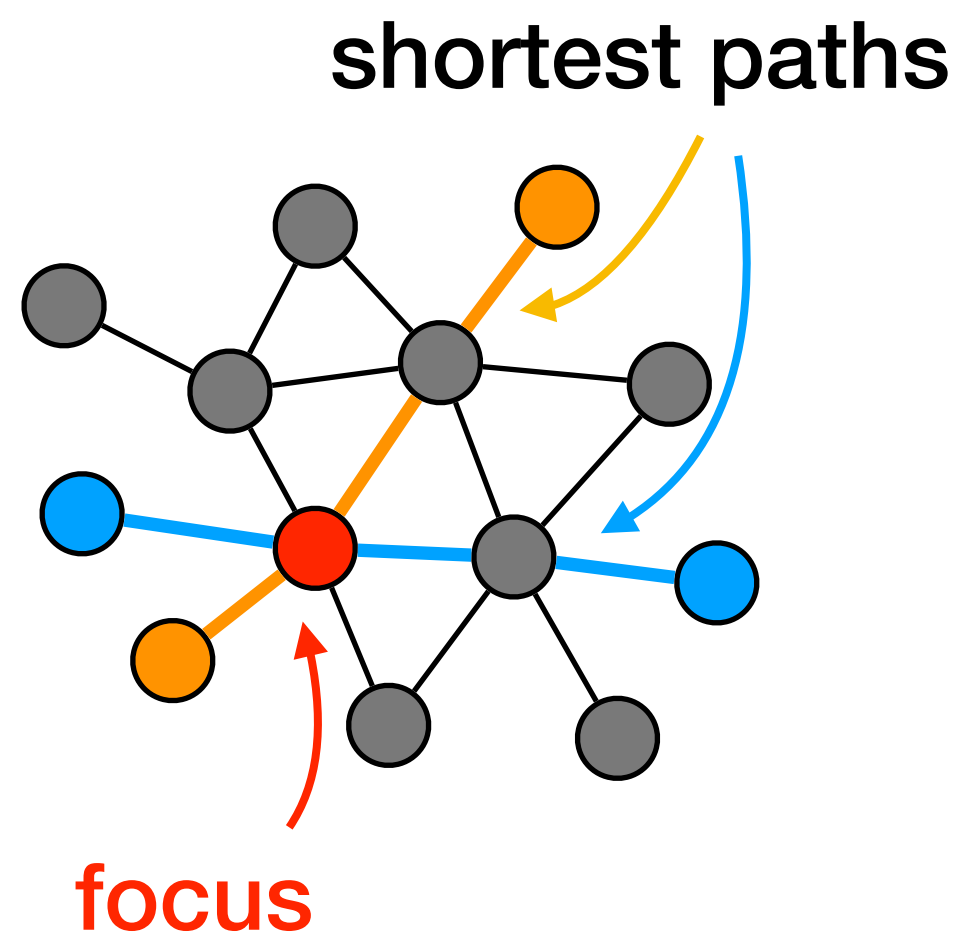
Topological



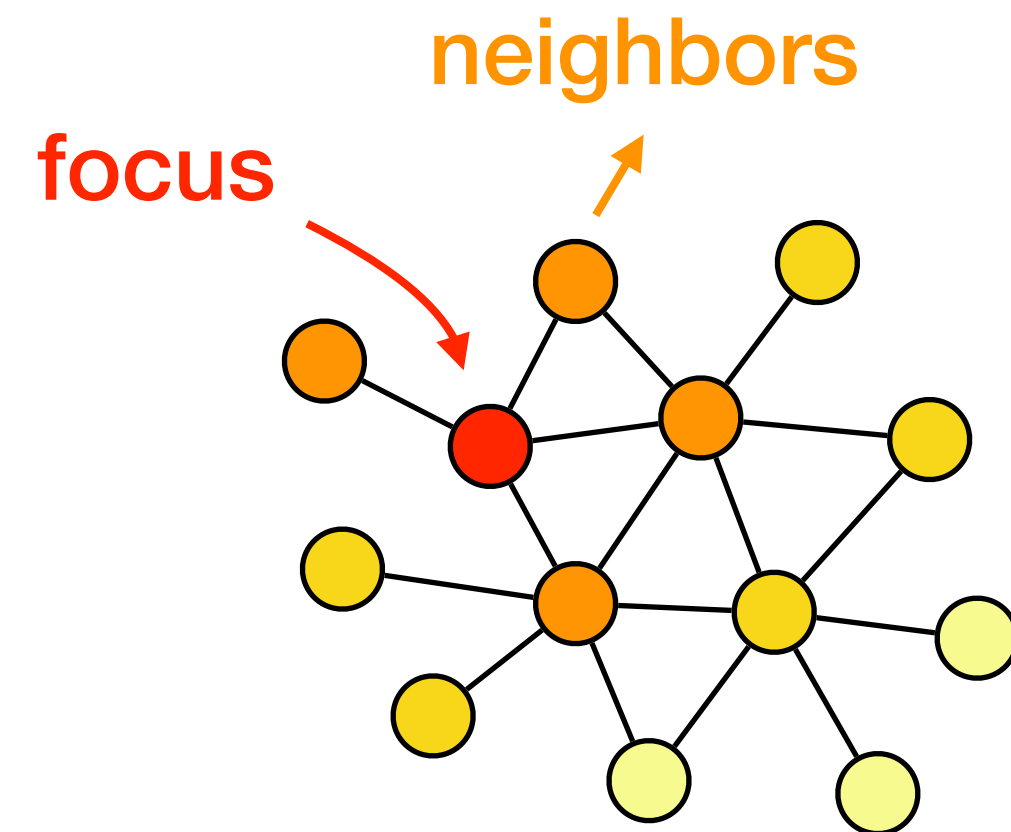
Feature vectors

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

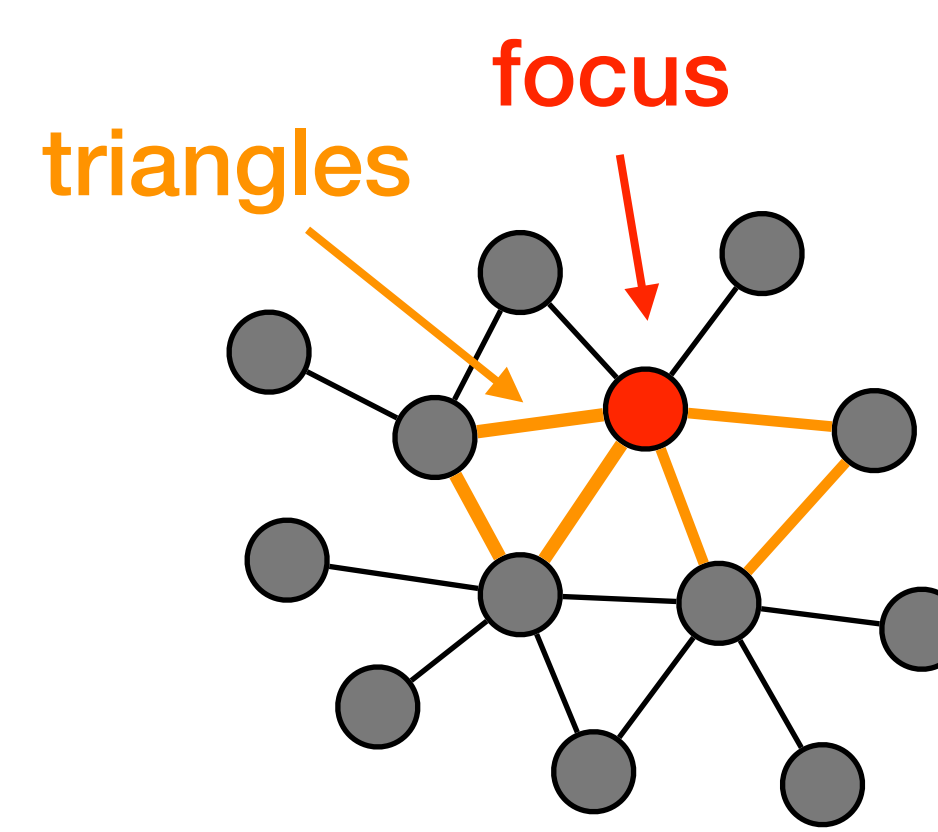
Node
centrality



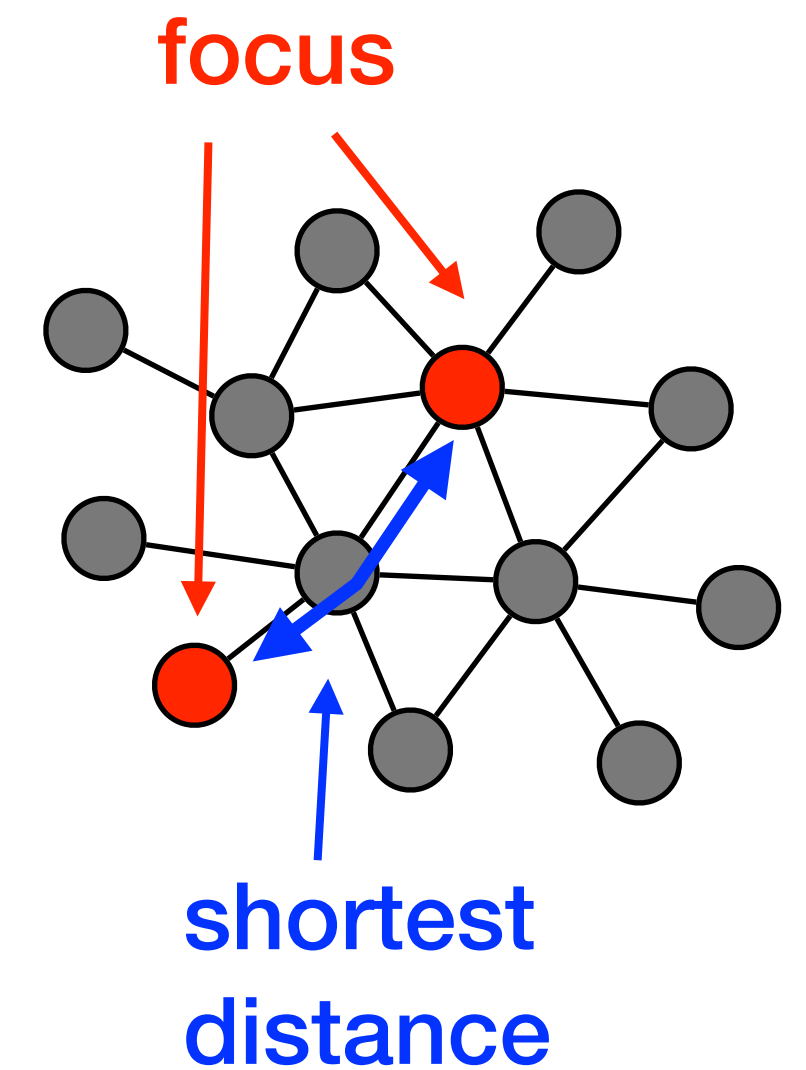
Neighborhood
richness



Topological
patterns



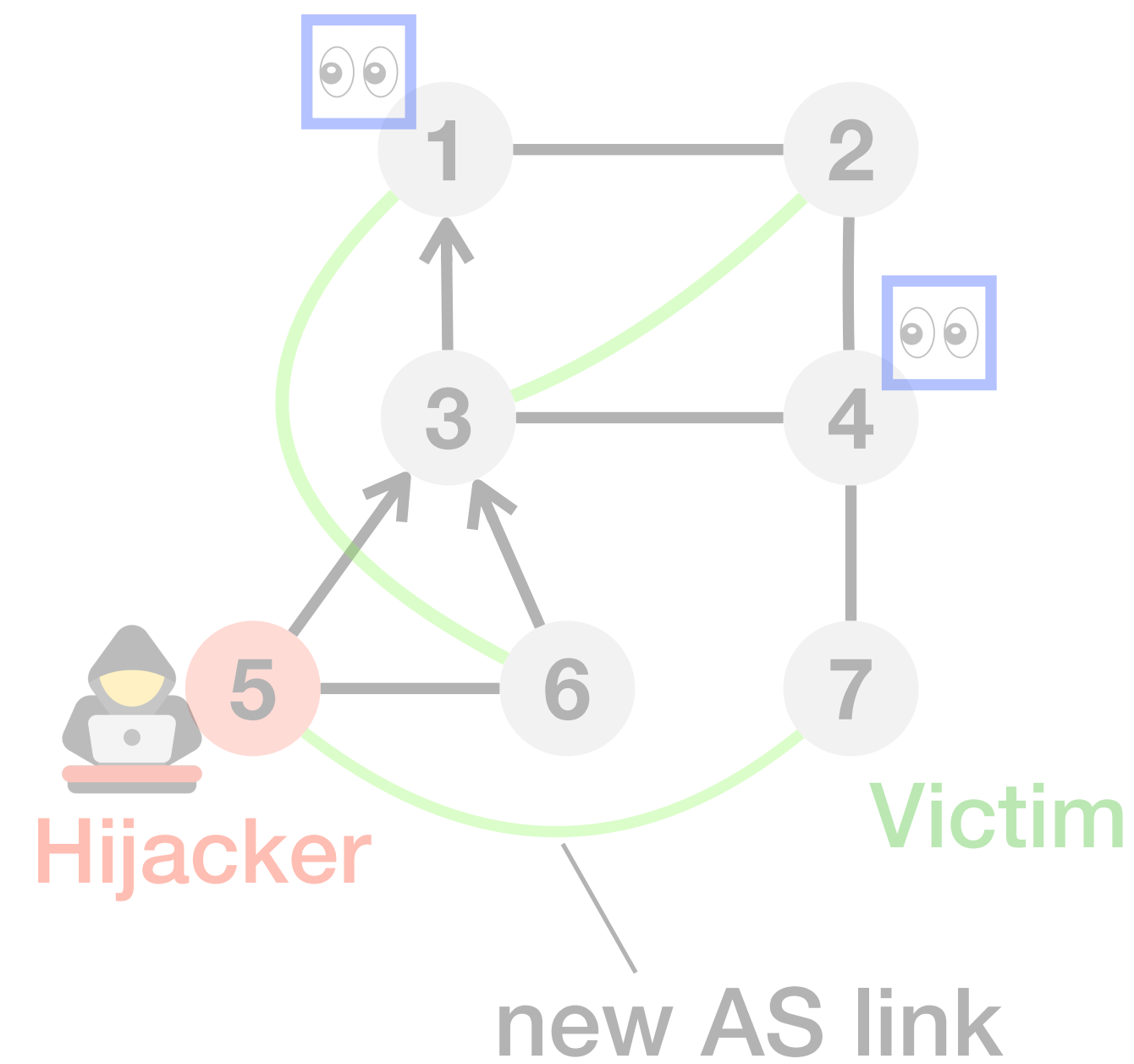
Closeness



DFOH's fake AS links inference algorithm comprises three steps



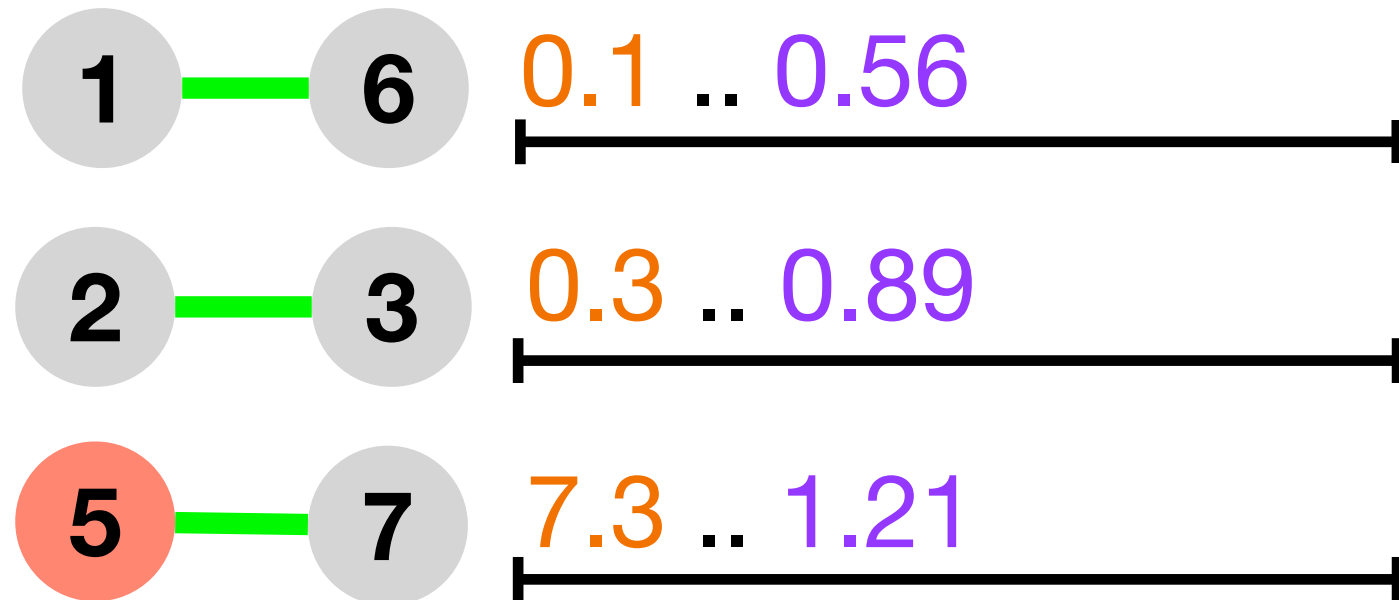
Vantage point



Feature categories:

Peeringdb

Topological



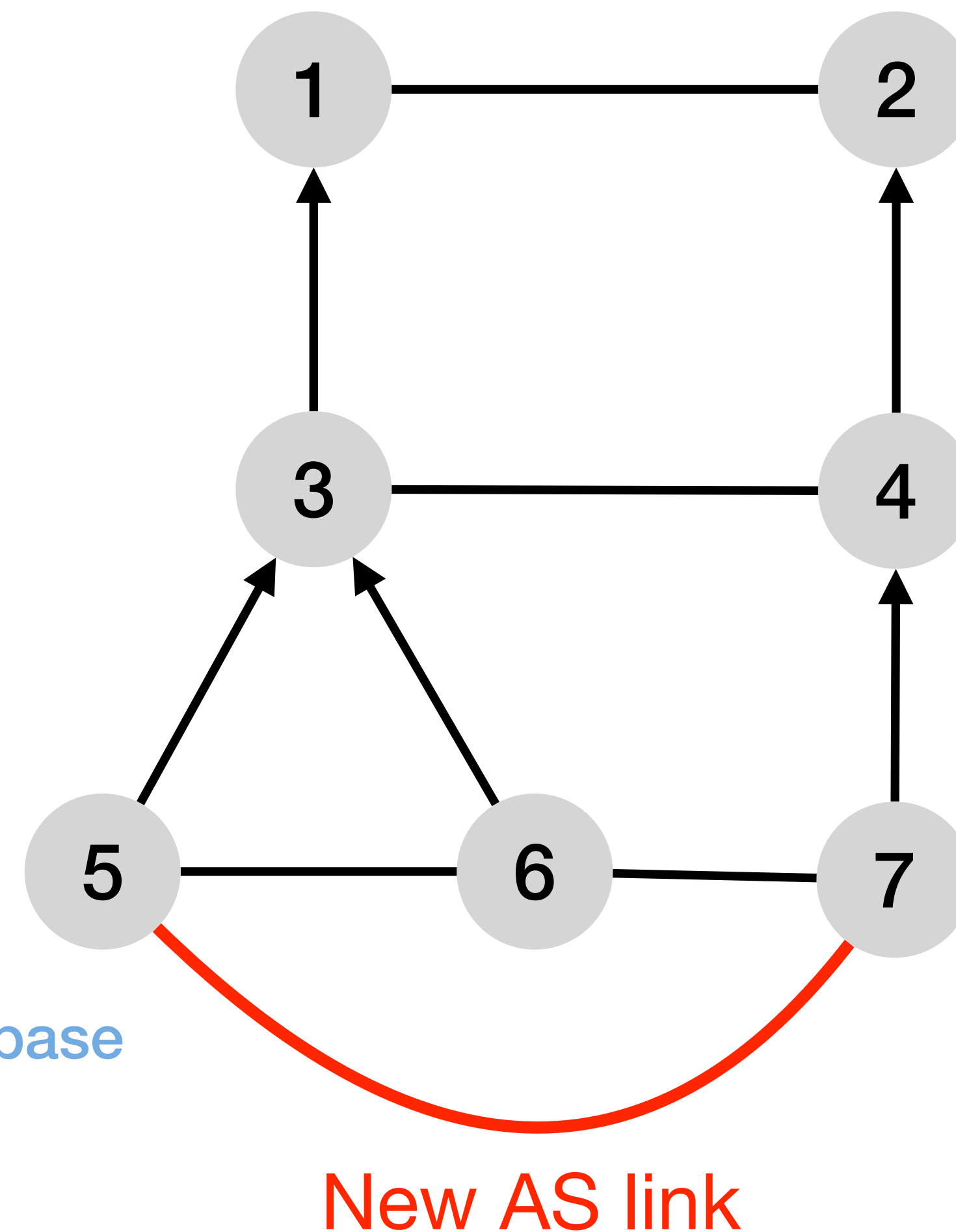
Feature vectors

DFOH leverages **correlations** in the public peering information

DFOH looks for three types
of information in PeeringDB:

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

Country: 
IXPs: AS-IX Cabase
Facilities: 
EQUINIX

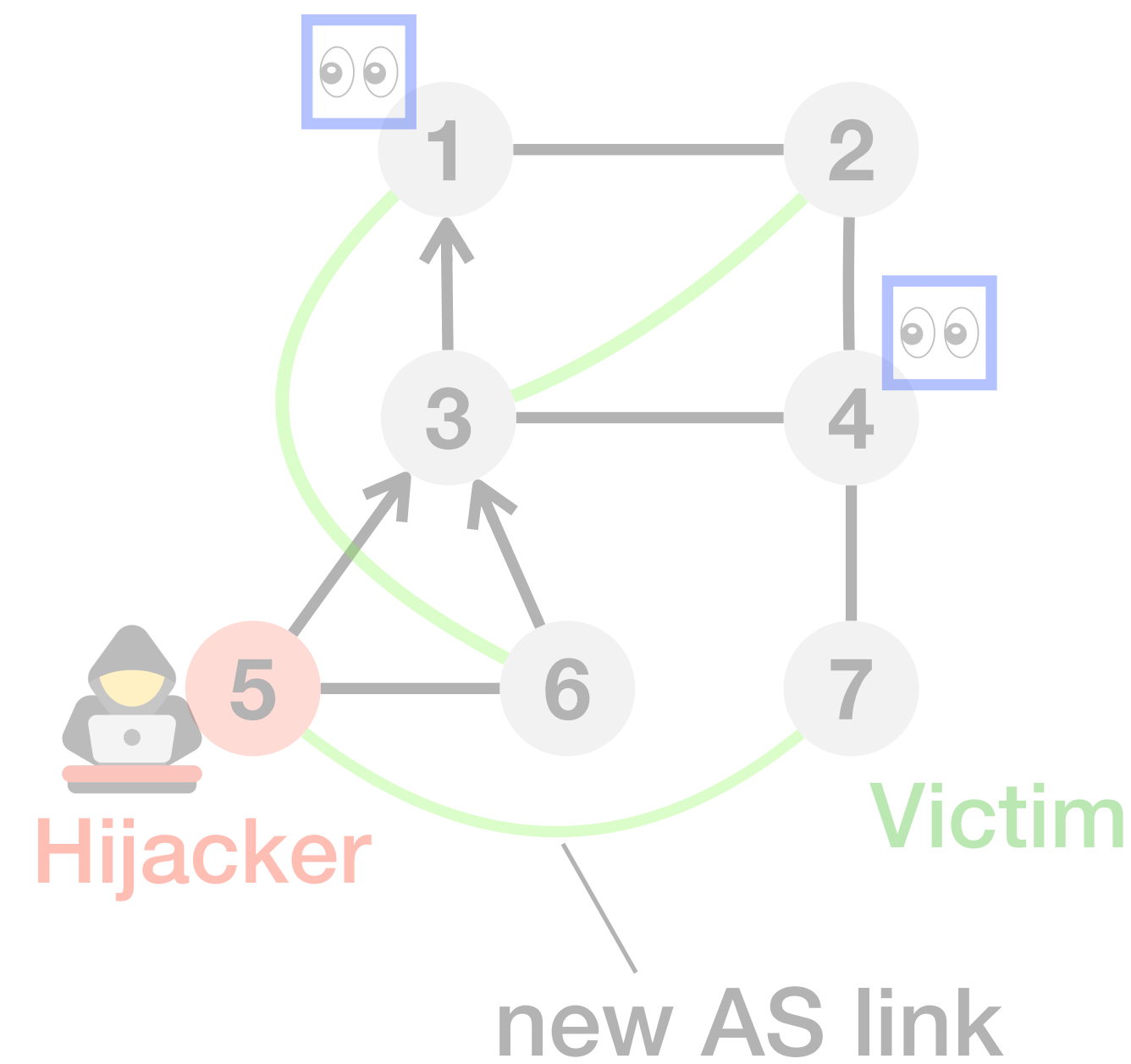


Country: 
IXPs: 
www.franceix.net

DFOH's fake AS links inference algorithm comprises three steps



Vantage point

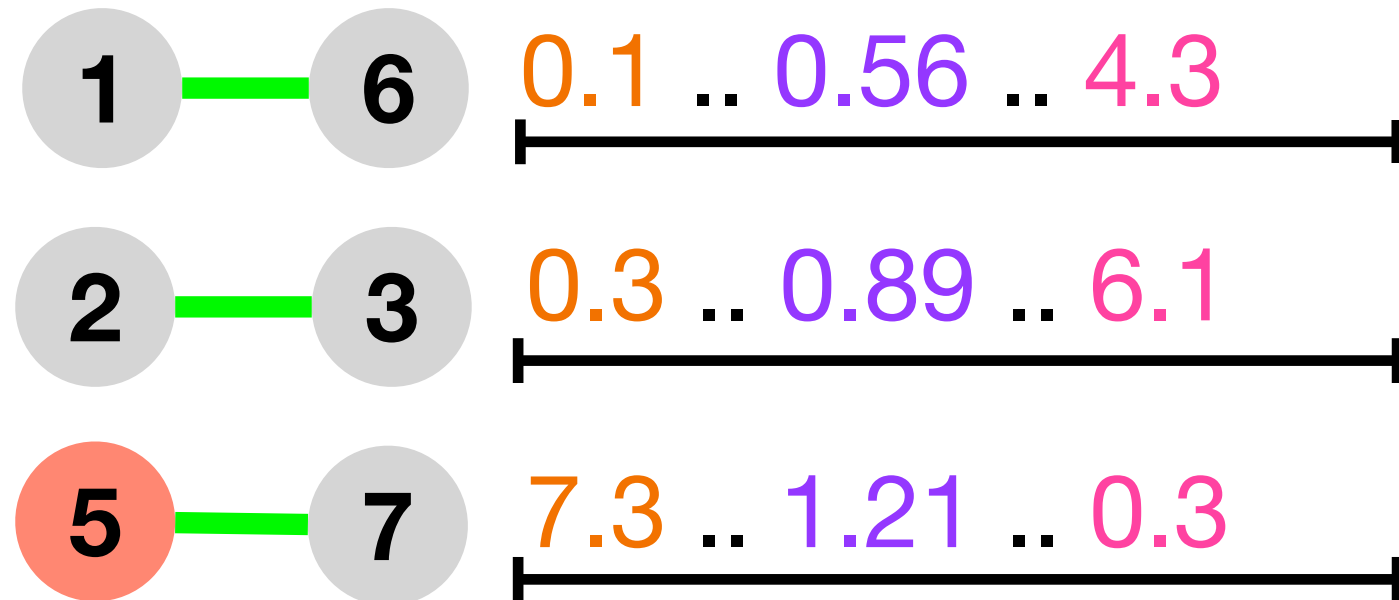


Feature categories:

AS-path pattern

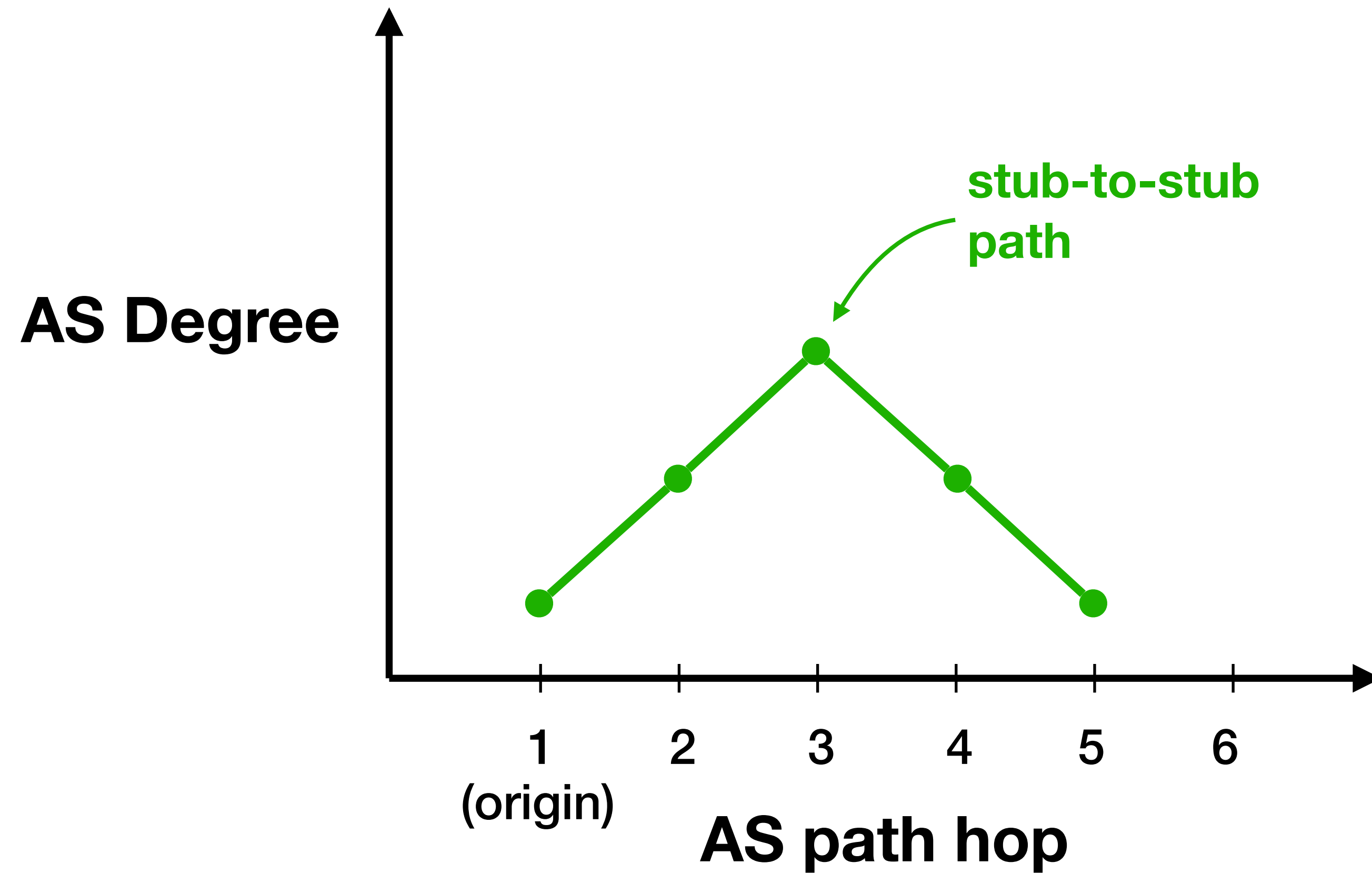
Peeringdb

Topological

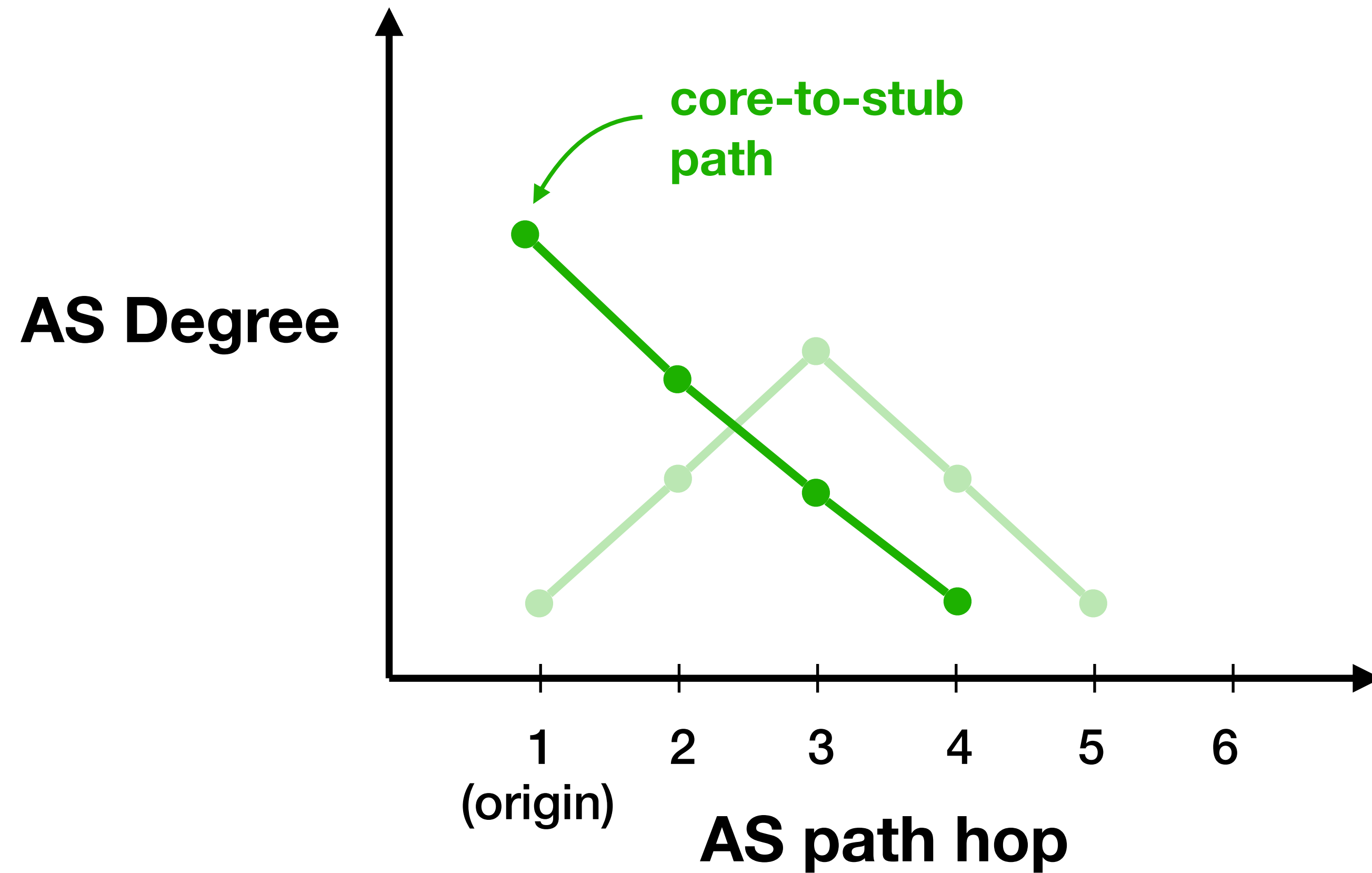


Feature vectors

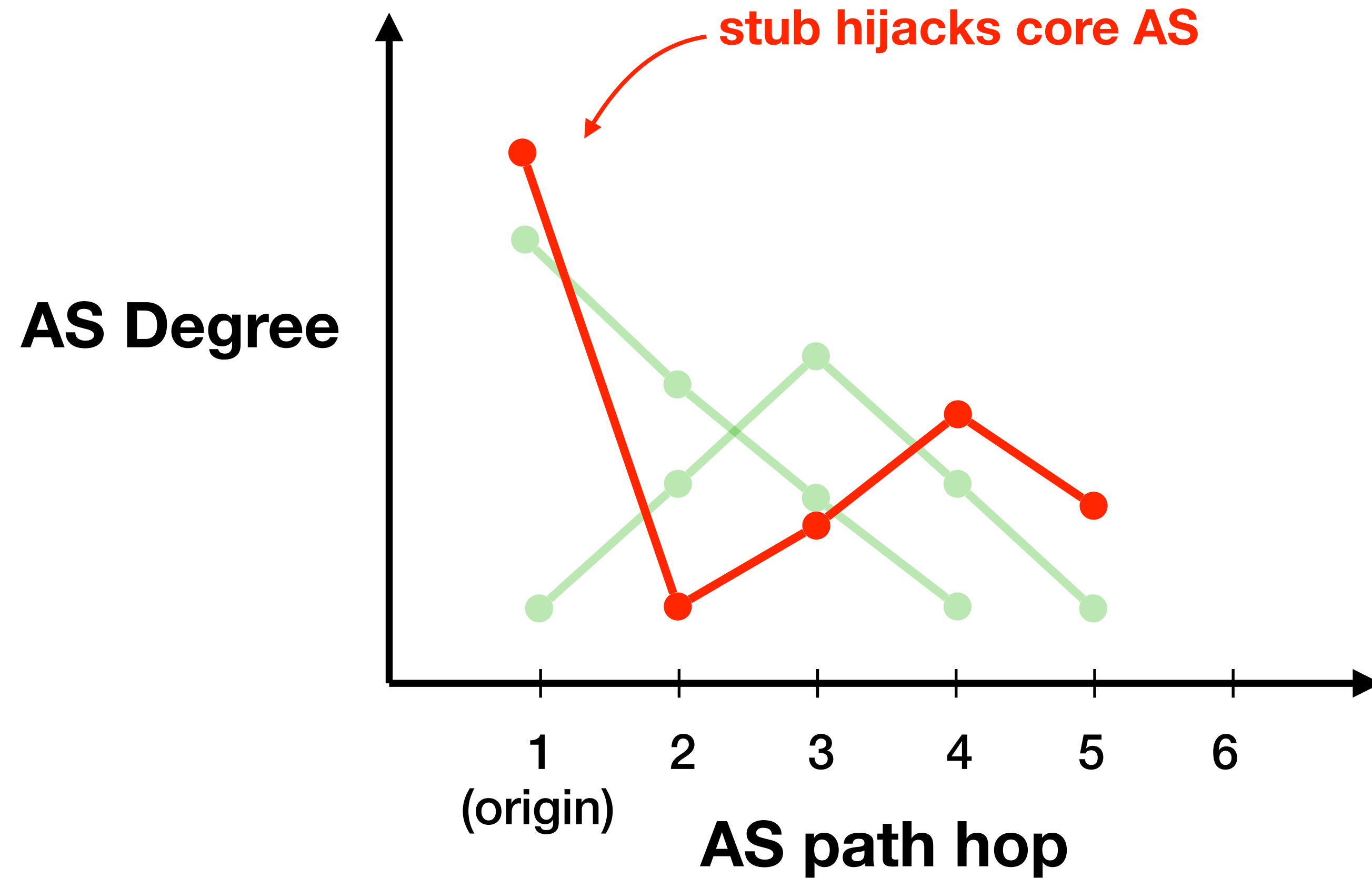
DFOH detects **fake AS paths** as they often violate patterns induced by business relationships



DFOH detects **fake AS paths** as they often violate patterns induced by business relationships



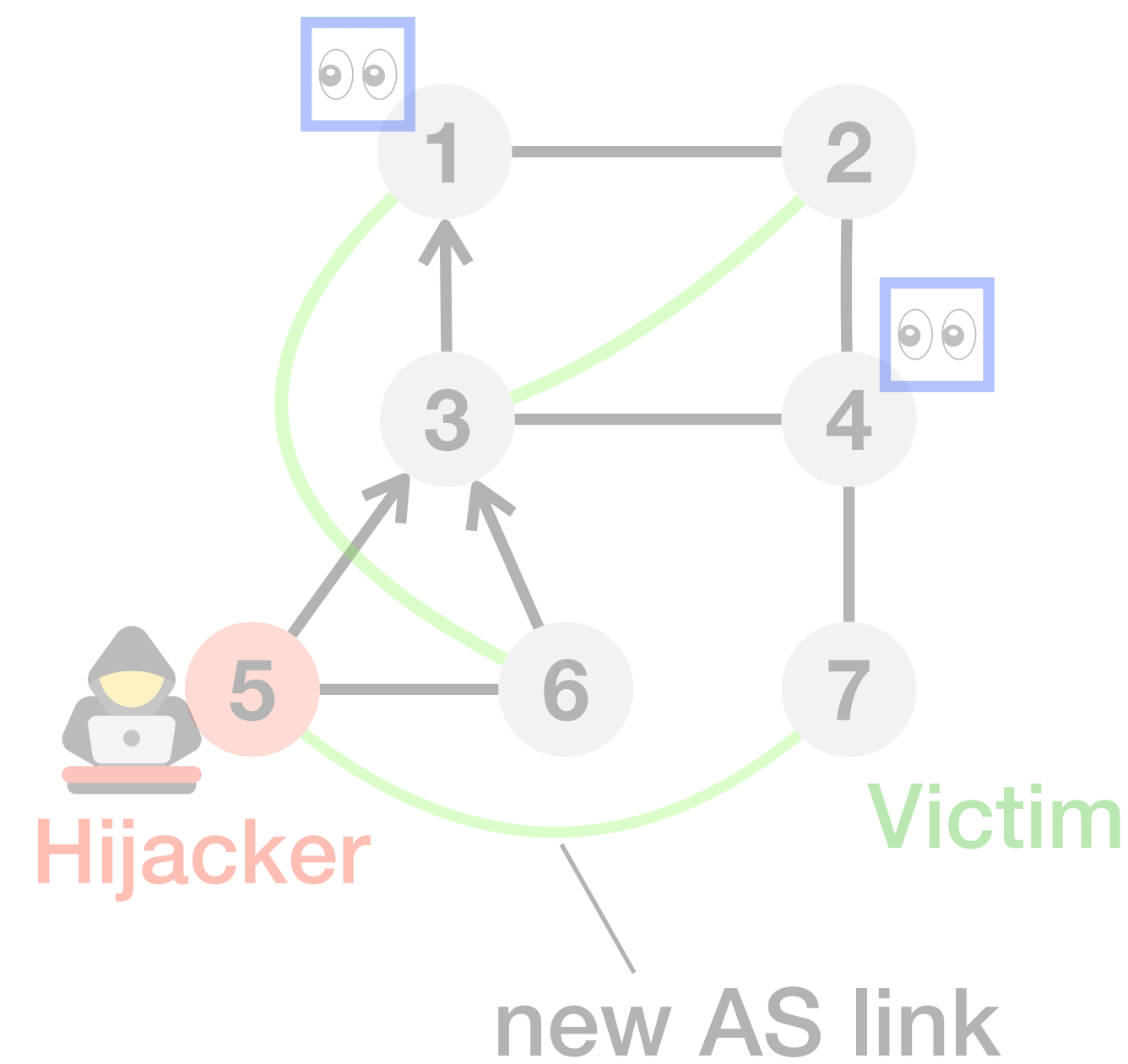
DFOH detects **fake AS paths** as they often violate patterns induced by business relationships



DFOH's fake AS links inference algorithm comprises three steps



Vantage point



Feature categories:

Bidirectionality

AS-path pattern

Peeringdb

Topological

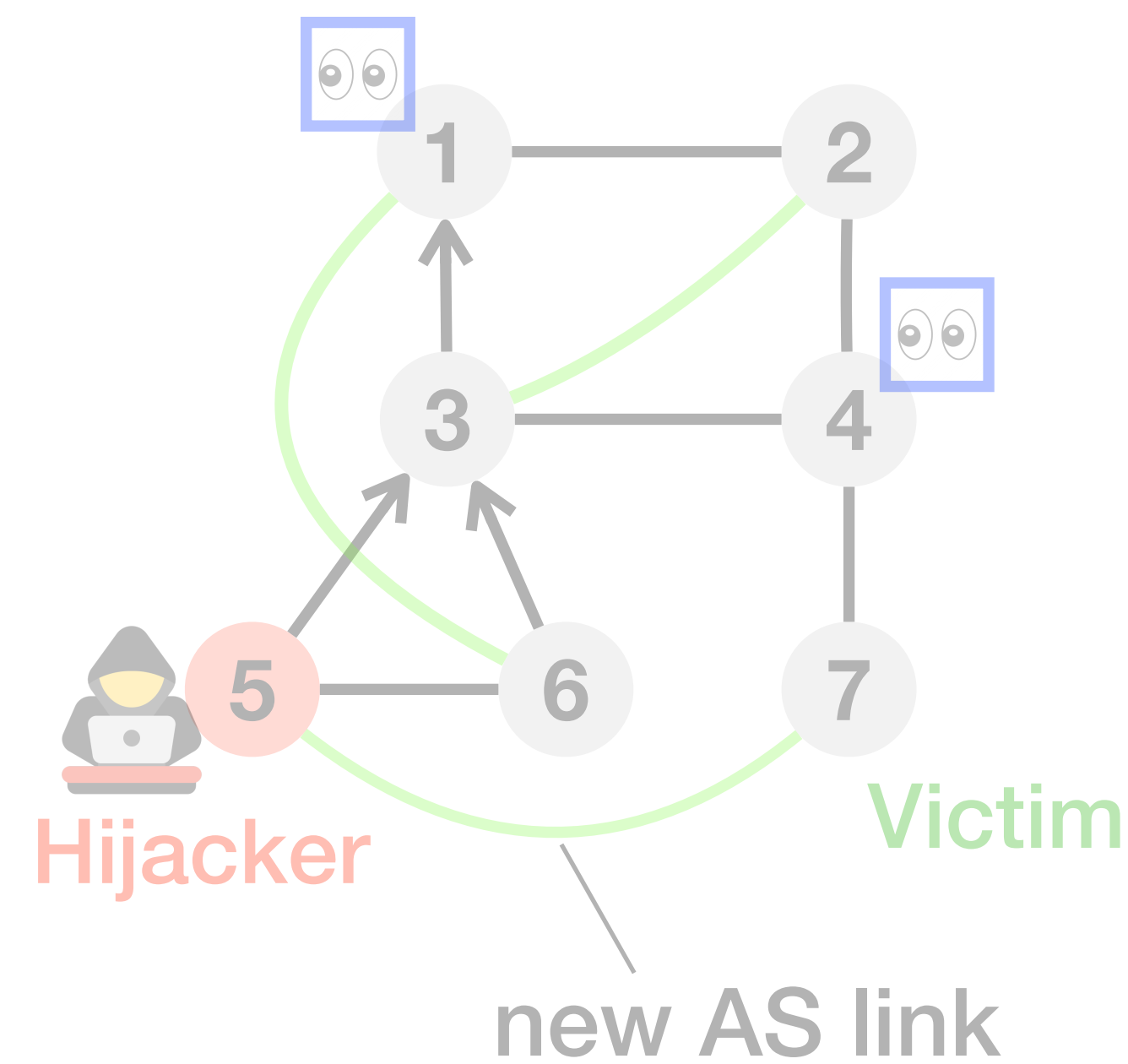
1	6	0.1	..	0.56	..	4.3	..	6
2	3	0.3	..	0.89	..	6.1	..	0
5	7	7.3	..	1.21	..	0.3	..	8

Feature vectors

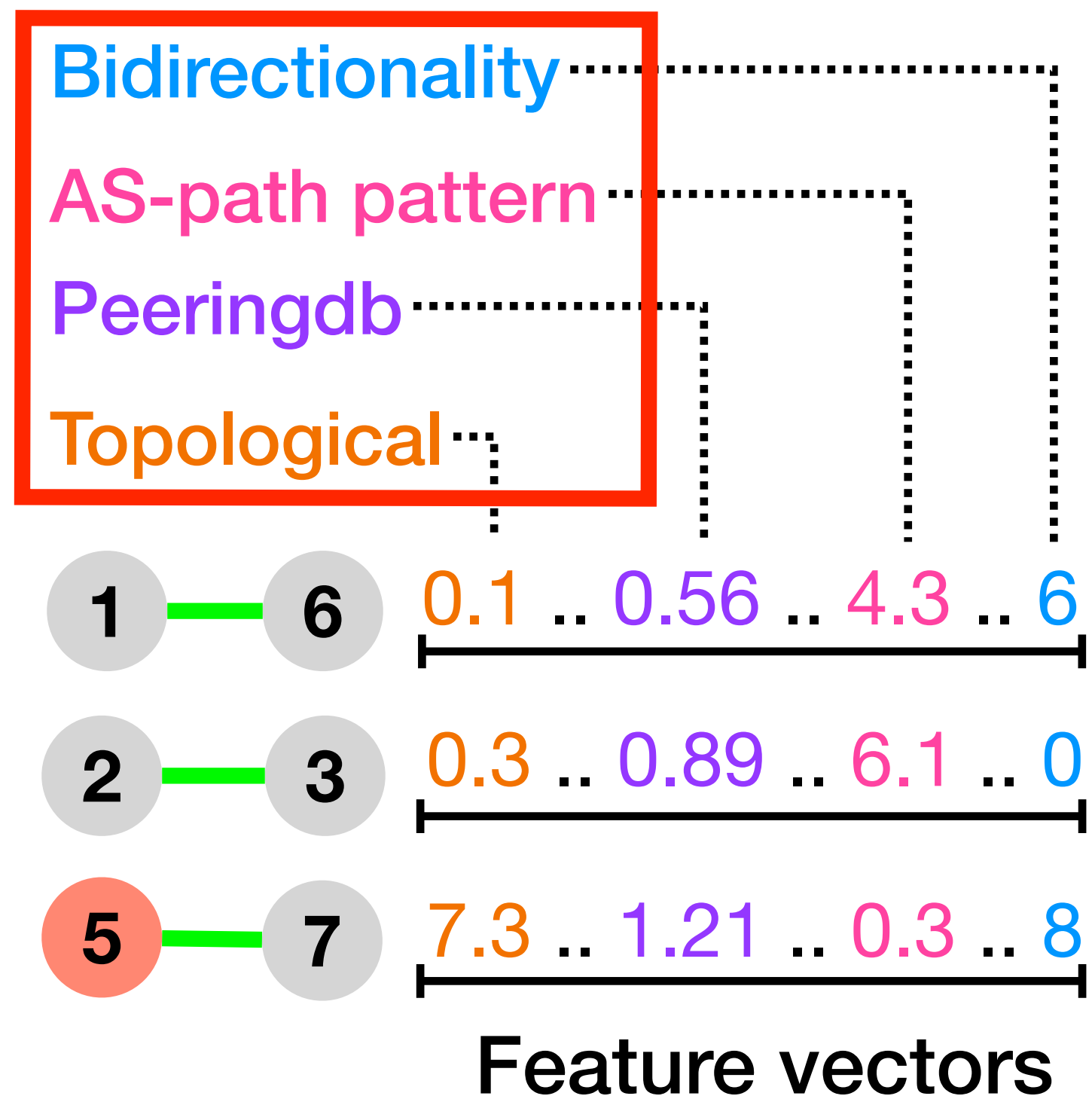
DFOH's fake AS links inference algorithm comprises three steps



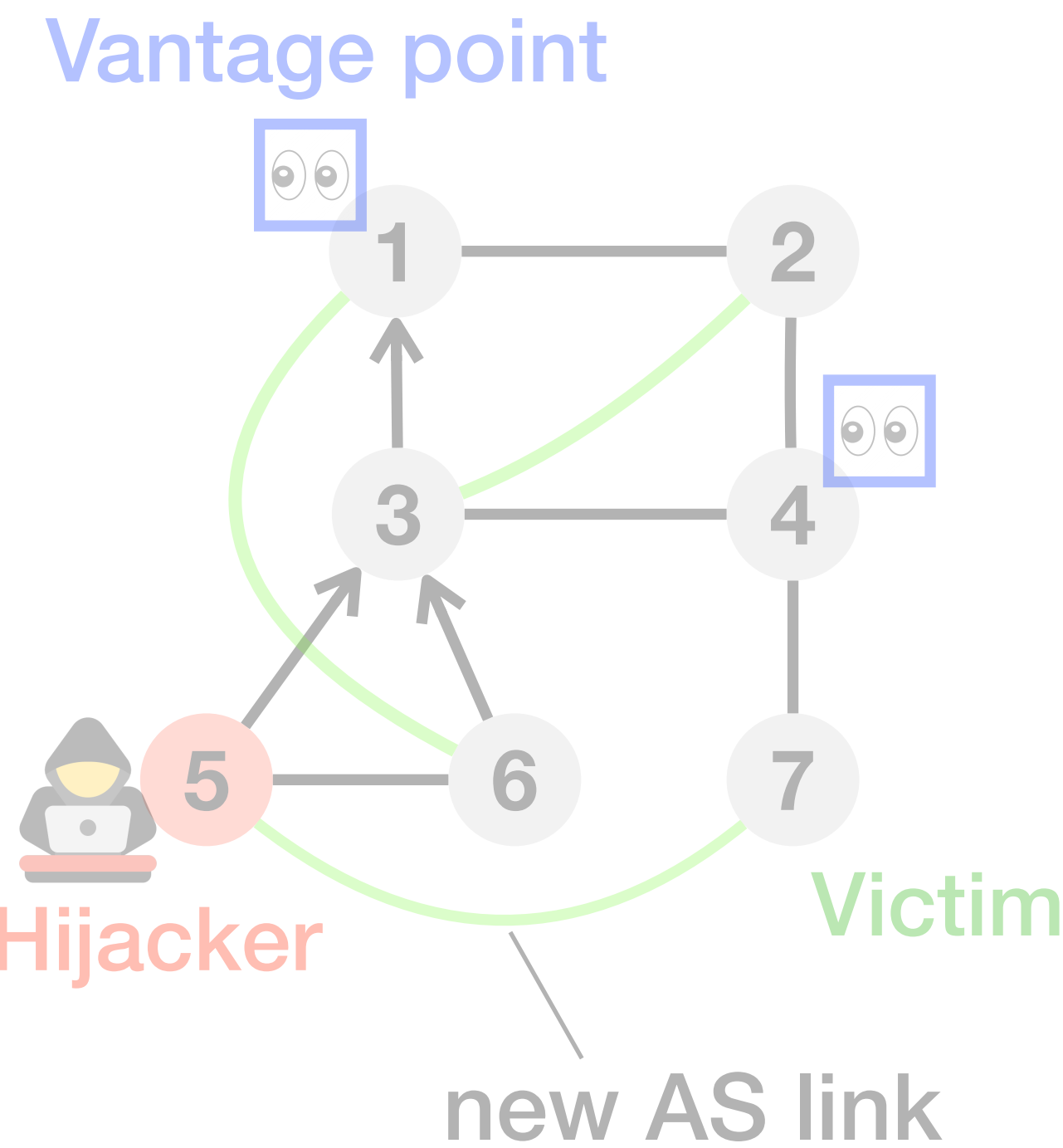
Vantage point



Key ingredient #1



DFOH's fake AS links inference algorithm comprises three steps



Feature categories:

Bidirectionality

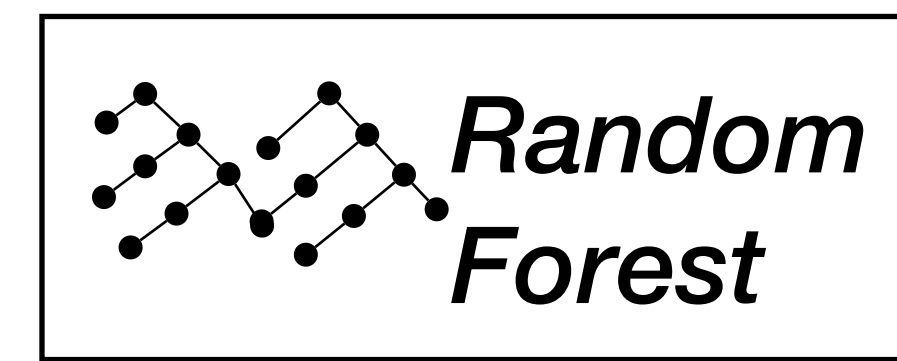
AS-path pattern

Peeringdb

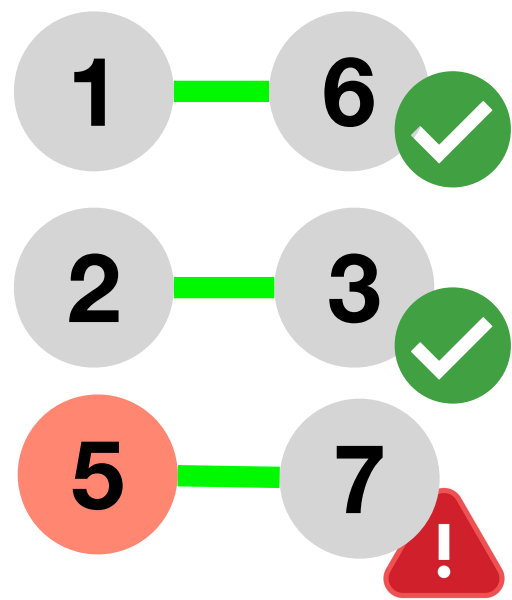
Topological

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

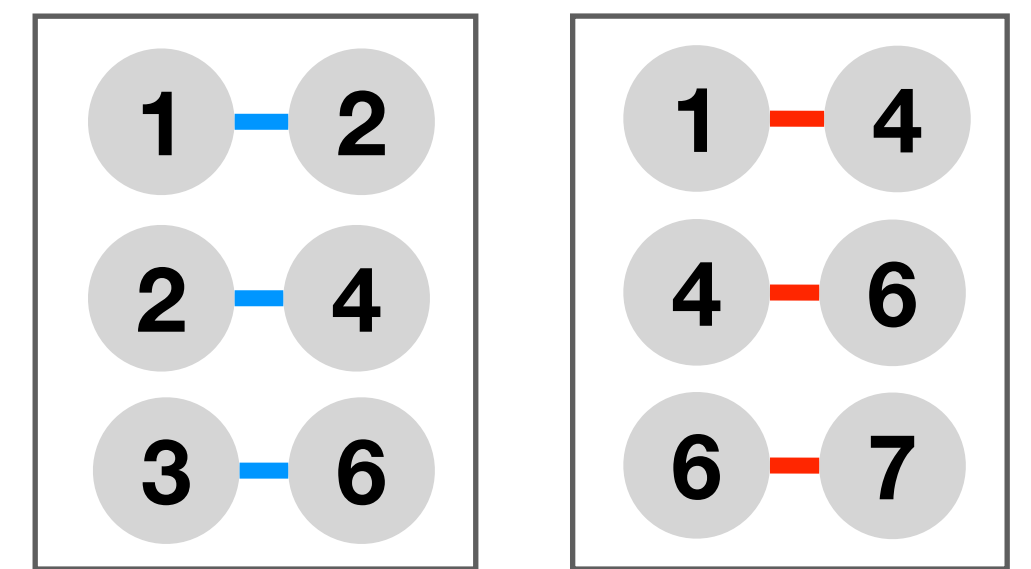


Inference



Training

Samples



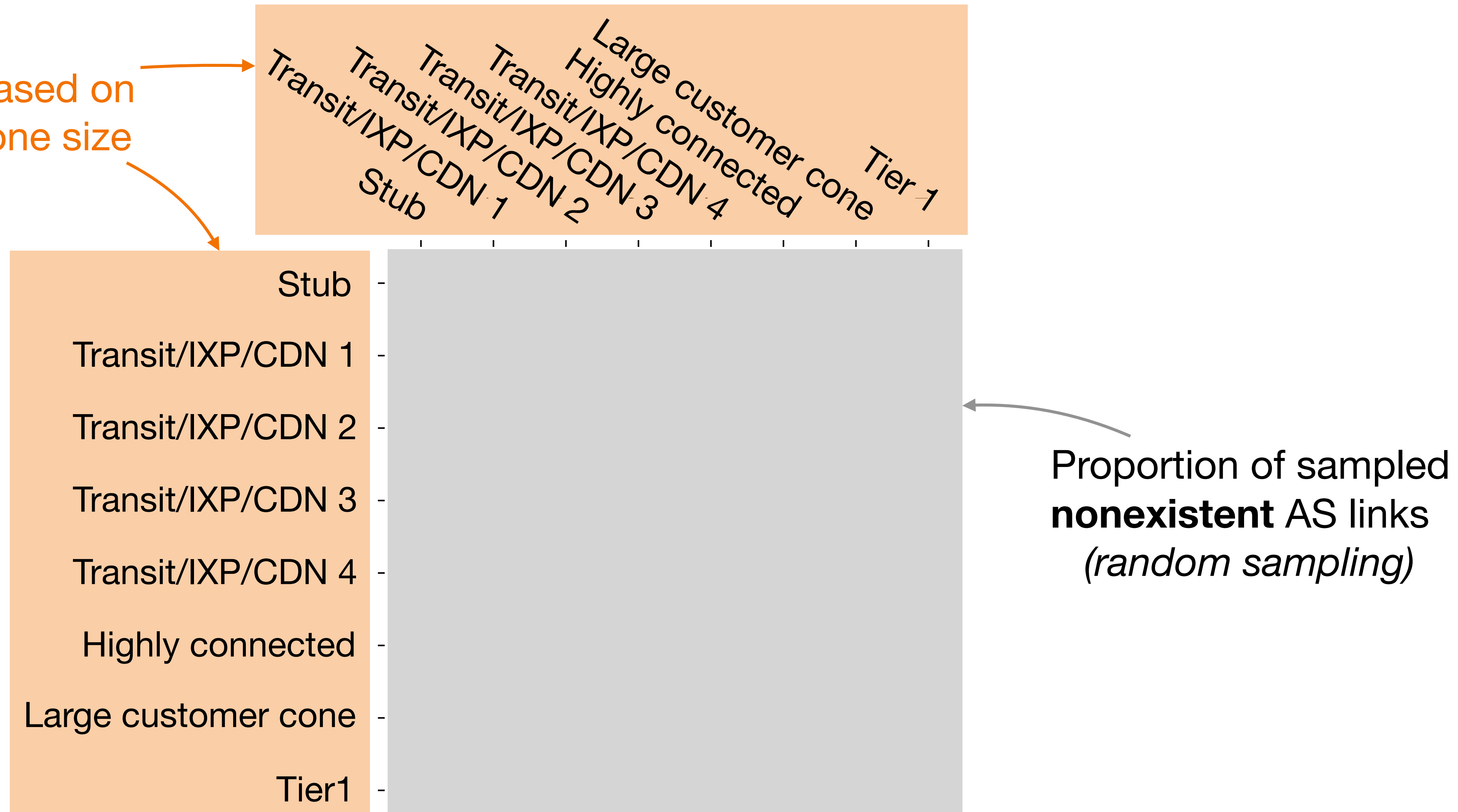
Existing links

Nonexistent links

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

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



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

	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	0.98	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 1	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Highly connected	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Large customer cone	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tier1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

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****

Large customer cone
Highly connected
Tier 1
Transit/IXP/CDN 4
Transit/IXP/CDN 3
Transit/IXP/CDN 2
Transit/IXP/CDN 1
Stub

Stub	0.98	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 1	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Highly connected	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Large customer cone	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tier1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

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

DFOH would perform well on scenarios involving two stubs

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

Large customer cone
Highly connected
Tier 1
Transit/IXP/CDN 4
Transit/IXP/CDN 3
Transit/IXP/CDN 2
Transit/IXP/CDN 1
Stub

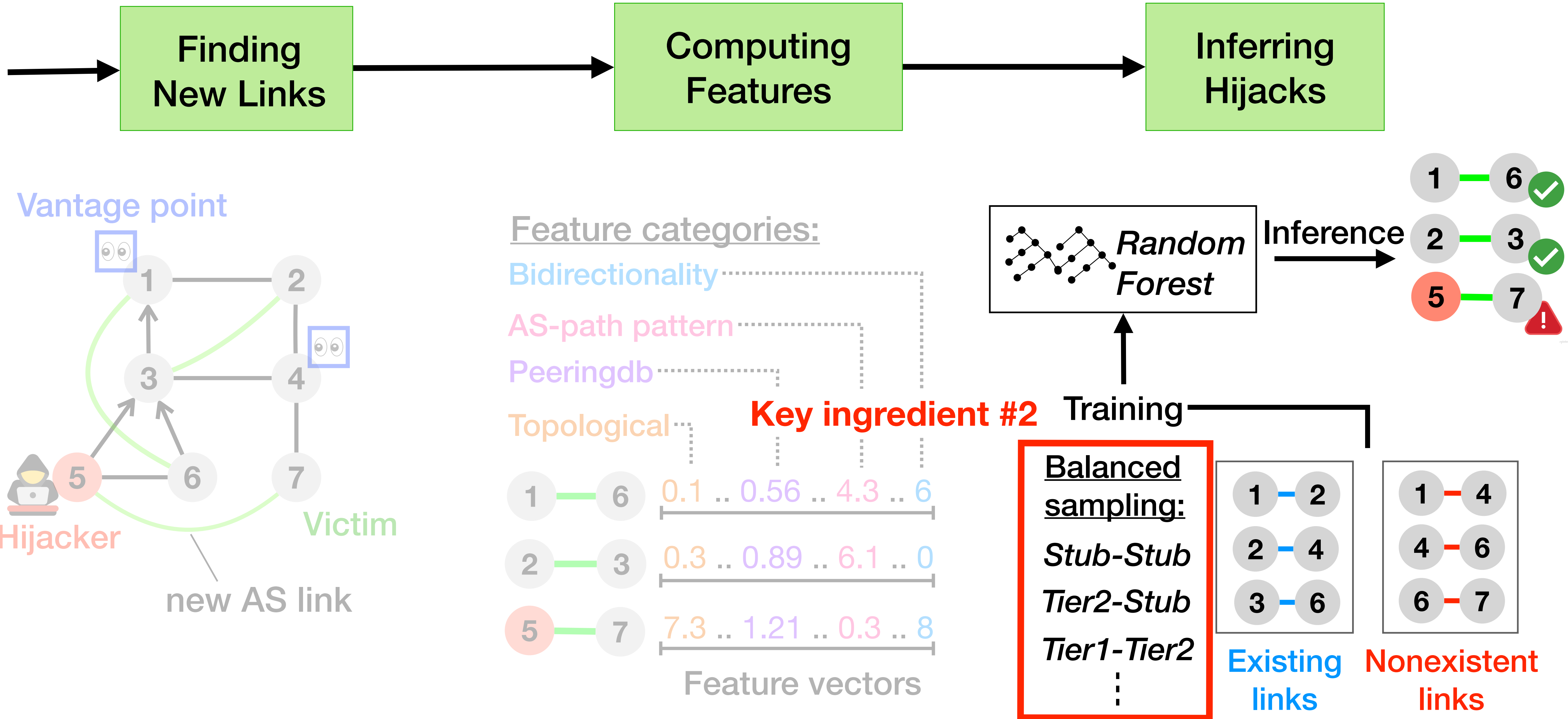
Stub	0.98	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 1	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Transit/IXP/CDN 4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Highly connected	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Large customer cone	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tier1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

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

discriminates fake AS links from the real ones

DFOH's inferences are accurate

in **every** attack scenario

DFOH is up and running

We evaluate *DFOH* on **artificially created** forged-origin hijacks and measure its accuracy upon every attack scenario

Methodology:

Step #1: We take existing AS paths and prepend a new origin to create a new link

Step #2: We consider 9k cases where the new link exists (*legitimate cases*) and 9k cases where the new link does not exist (*malicious cases*)

We evaluate *DFOH* on **artificially created** forged-origin hijacks and measure its accuracy upon every attack scenario

Methodology:

Step #1: We take existing AS paths and prepend a new origin to create a new link

Step #2: We consider 9k cases where the new link exists (*legitimate cases*) and 9k cases where the new link does not exist (*malicious 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

	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

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

The minimum TPR is 0.73

DFOH is **accurate** upon every attack scenario

Victim

False 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	0.04	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

DFOH is **accurate** upon every attack scenario

False Positive Rate		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
Attacker	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

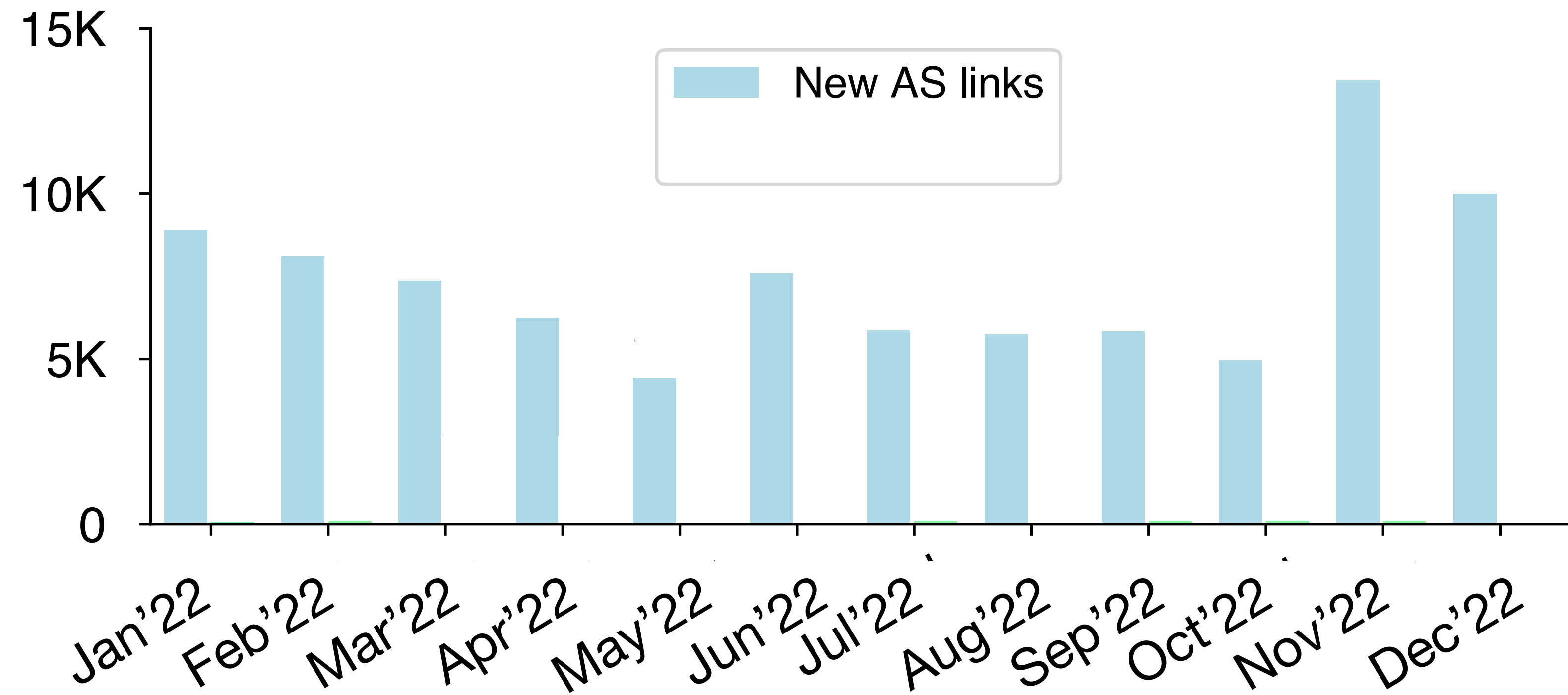
in **every** attack scenario

DFOH is up and running

and **useful** for operators

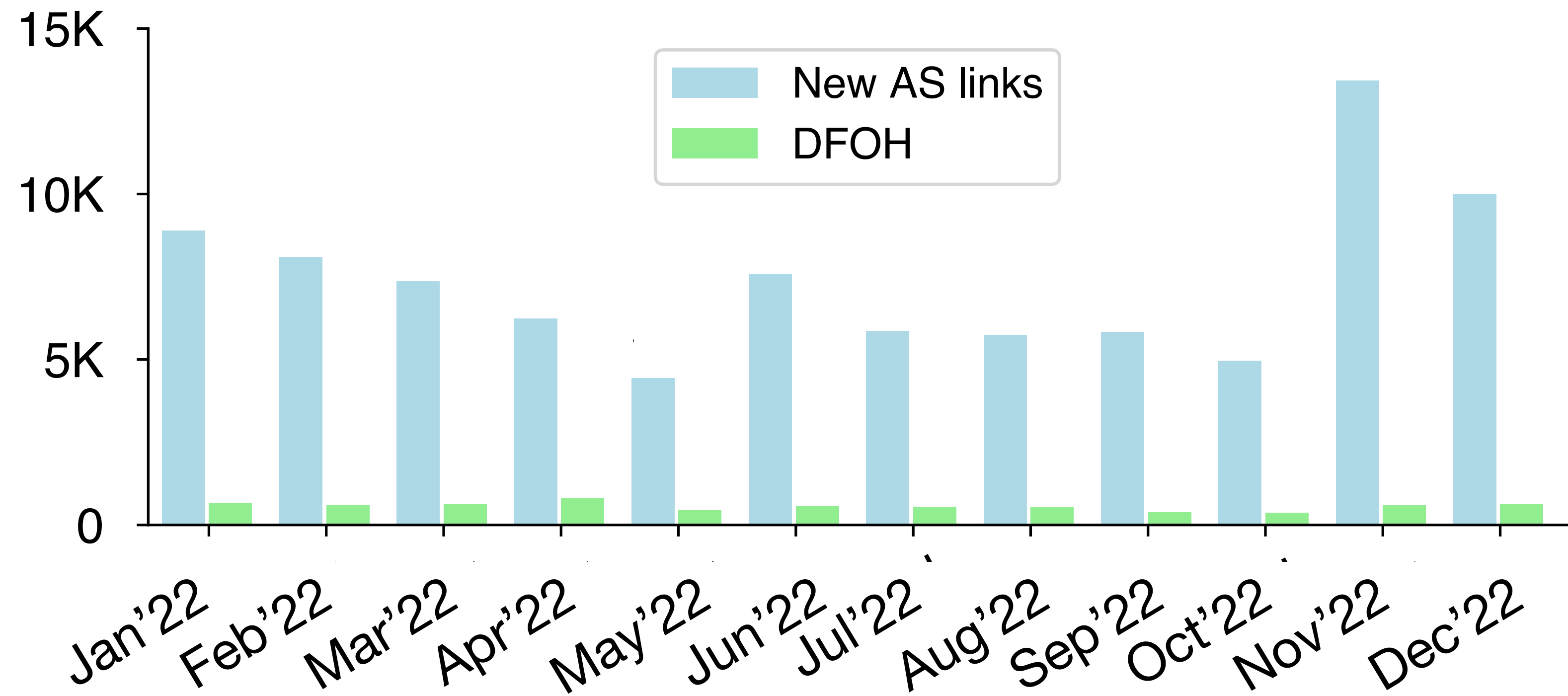
DFOH makes the detection of forged-origin hijacks **practical** for operators

Number of reported cases



DFOH makes the detection of forged-origin hijacks **practical** for operators

Number of reported cases



DFOH is up and running at <https://dfoh.uclouvain.be/>

DFOH
A System to Detect Forged-Origin 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 quickly and with high confidence know when their IP prefixes are being hijacked.

How DFOH works

DFOH detects forged-origin hijacks using a three-steps approach depicted in the figure below.

- DFOH finds new AS links by parsing the AS paths of the collected BGP routes. DFOH zooms on these new AS links because a forged-origin hijack likely triggers the appearance of a new AS link between the victim and attacker (in case of a Type-1 hijack). Yet, most of the new AS links are the result of legitimate events such as new peering agreements. Thus, the following steps aim to discriminate the fake AS links from the legitimate ones.
- DFOH computes a set of features for every new AS links detected. The features can be divided into four categories (topological, AS-path pattern, peering and bidirectionality) that complement each other. We carefully chose this set of features to make sure DFOH is accurate upon various attack scenarios and is robust against adversarial inputs.
- DFOH uses an inference pipeline that is similar to what generic state-of-the-art link prediction frameworks use. Yet, the training part of the pipeline relies on a balanced sampling algorithm that makes DFOH robust against the routing biases and avoids DFOH to overfit for one particular attack scenario.

New Link Zooming

Feature Computation

Hijack Inference

Autonomous System

Hijacker

Victim

new AS link

Feature categories:

- Bidirectionality
- AS-path pattern
- Peeringdb
- Topological

Feature vectors

Balanced sampling:

	Existing links	Nonexistent links
Stub-Stub	6-7	6-8
Tier2-Stub	4-7	5-7
Tier1-Tier2	2-5	1-5

Training

Random Forest

Inference

1-5 ✓

4-8 ✓

6-9 ⚠



We provide the paper, presentations and source code

We showcase *DFOH* with APNIC's prefix and ASN!

Attacker	<input type="text"/>
Victim	<input type="text" value="4608"/> ← APNIC's ASN
Start date	<input type="text" value="2022-01-01"/>
End date	<input type="text" value="2023-12-01"/>
<input checked="" type="checkbox"/> Only show the suspicious cases	

We showcase *DFOH* with APNIC's prefix and ASN!

Attacker

Victim ← APNIC's ASN

Start date

End date

Only show the suspicious cases

There suspicious cases reported over two years

Date	AS link	# of AS paths	DFOH inference	Confidence level
2022-07-10	4608 147028	1	suspicious	2
2022-07-22	4608 9269	1	suspicious	2
2022-07-25	3257 4608	27	suspicious	1

We showcase *DFOH* with APNIC's prefix and ASN!

Attacker

Victim ← APNIC's ASN

Start date

End date

Only show the suspicious cases

There suspicious cases reported over two years

Date	AS link	# of AS paths	DFOH inference	Confidence level
2022-07-10	4608 147028	1	suspicious	2
2022-07-22	4608 9269	1	suspicious	2
2022-07-25	3257 4608	27	suspicious	1

Time	Prefix	AS path	Vantage points
2022-07-10 07:12:37	103.0.0.0/16	44393 147028 4608	RRC00 49.12.70.222

We showcase *DFOH* with APNIC's prefix and ASN!

Attacker

Victim ← APNIC's ASN

Start date

End date

Only show the suspicious cases

There suspicious cases reported over two years

Date	AS link	# of AS paths	DFOH inference	Confidence level
2022-07-10	4608 147028	1	suspicious	2
2022-07-22	4608 9269	1	suspicious	2
2022-07-25	3257 4608	27	suspicious	1

APNIC's Prefix

Time	Prefix	AS path	Vantage points
2022-07-10 07:12:37	103.0.0.0/16	44393 147028 4608	RRC00 49.12.70.222

APNIC's ASN

We showcase *DFOH* with APNIC's prefix and ASN!

Attacker

Victim ← APNIC's ASN

Start date

End date

Only show the suspicious cases

There suspicious cases reported over two years

Date	AS link	# of AS paths	DFOH inference	Confidence level
2022-07-10	4608 147028	1	suspicious	2
2022-07-22	4608 9269	1	suspicious	2
2022-07-25	3257 4608	27	suspicious	1

Time	Prefix	AS path	Vantage points
2022-07-10 07:12:37	103.0.0.0/16	44393 147028 4608	RRC00 49.12.70.222

APNIC's Prefix

Suspicious AS (stub)

Securebit (BGP tunnel broker)

APNIC's ASN

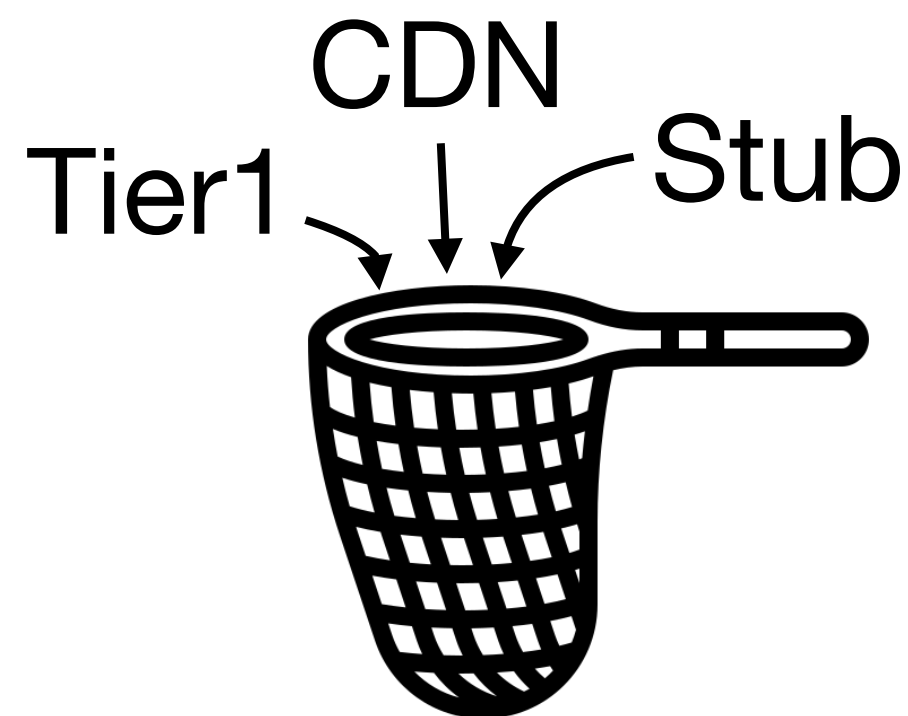
***DFOH*: A System to Detect Forged-Origin Hijacks**



DFOH runs in a commodity server



DFOH detects hijacks on the whole Internet

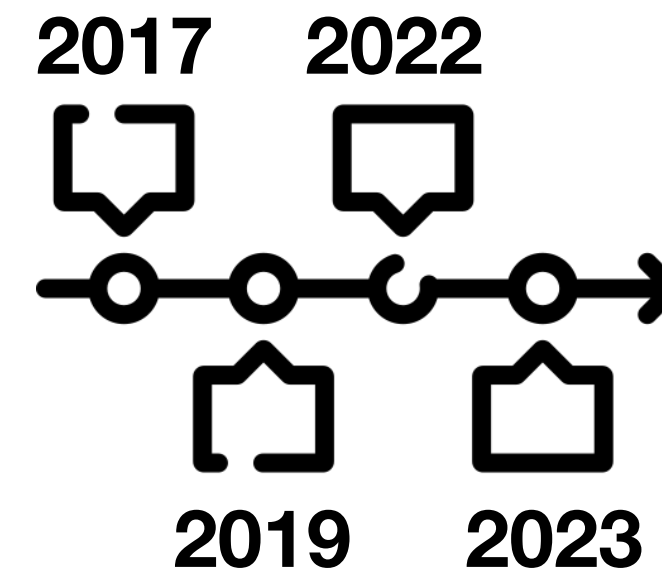


DFOH is accurate in every attack scenario

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



***DFOH* runs in a commodity server**



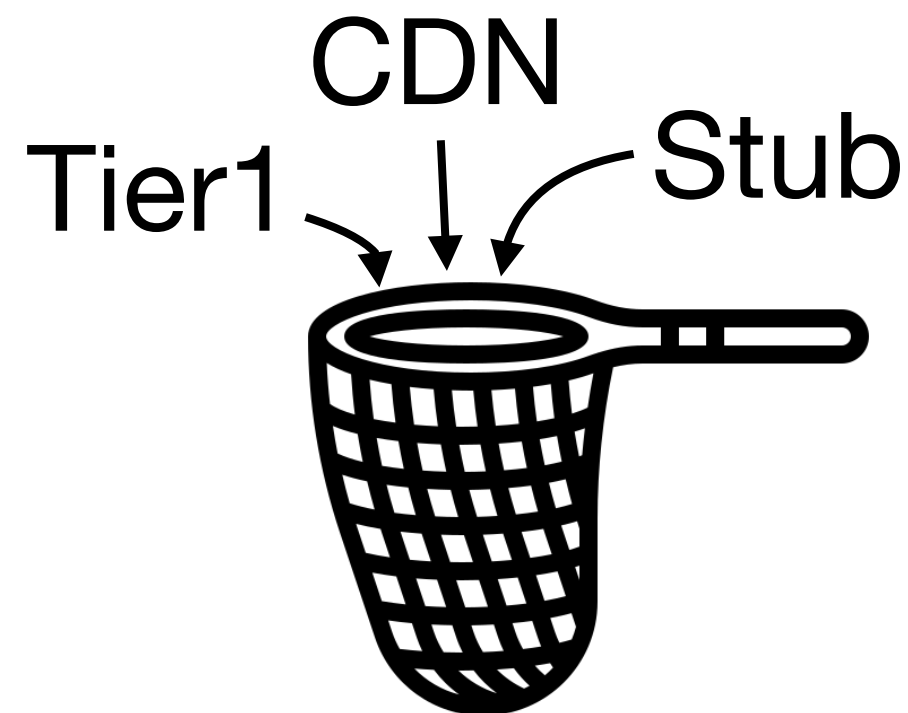
***DFOH* detects past hijacks**



***DFOH* detects hijacks on the whole Internet**



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



***DFOH* is accurate in every attack scenario**



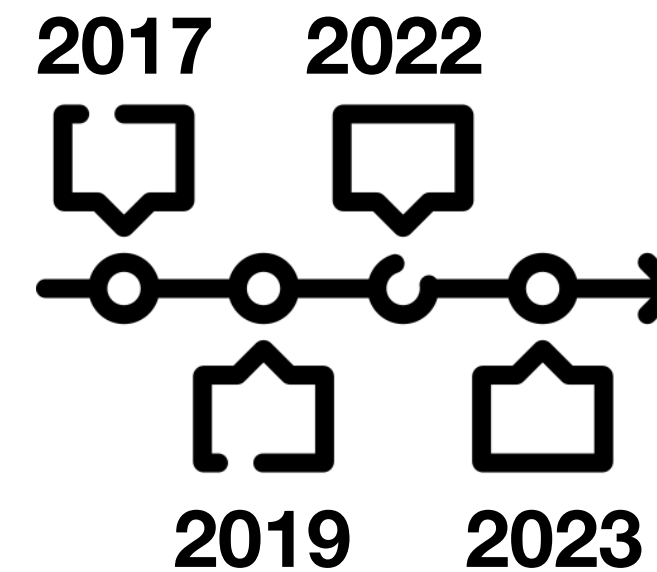
***DFOH* is robust against adversarial inputs**

DFOH: A System to Detect Forged-Origin Hijacks

<https://dfoh.uclouvain.be>



DFOH runs in a commodity server



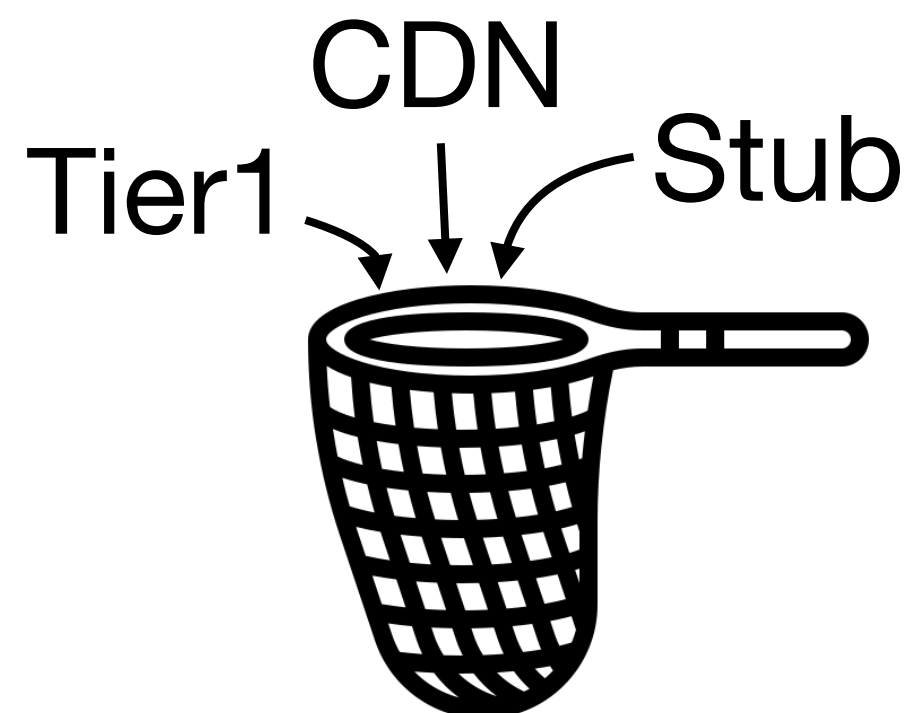
DFOH detects past hijacks



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