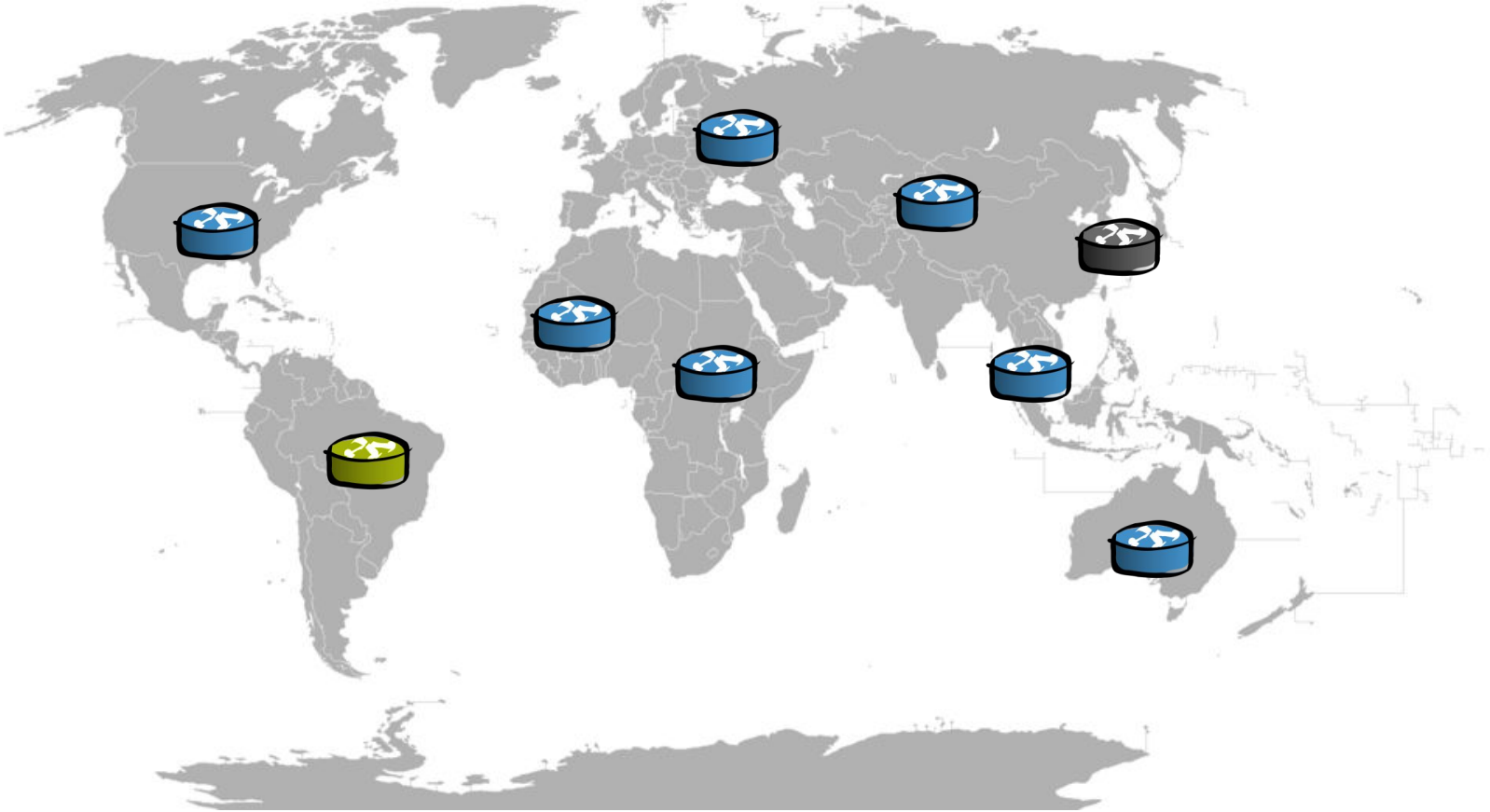
Uncovering BGP Action Communities and Community Squatters in the Wild

Brivaldo A. da Silva Jr Adriano B. de Carvalho Ítalo Cunha

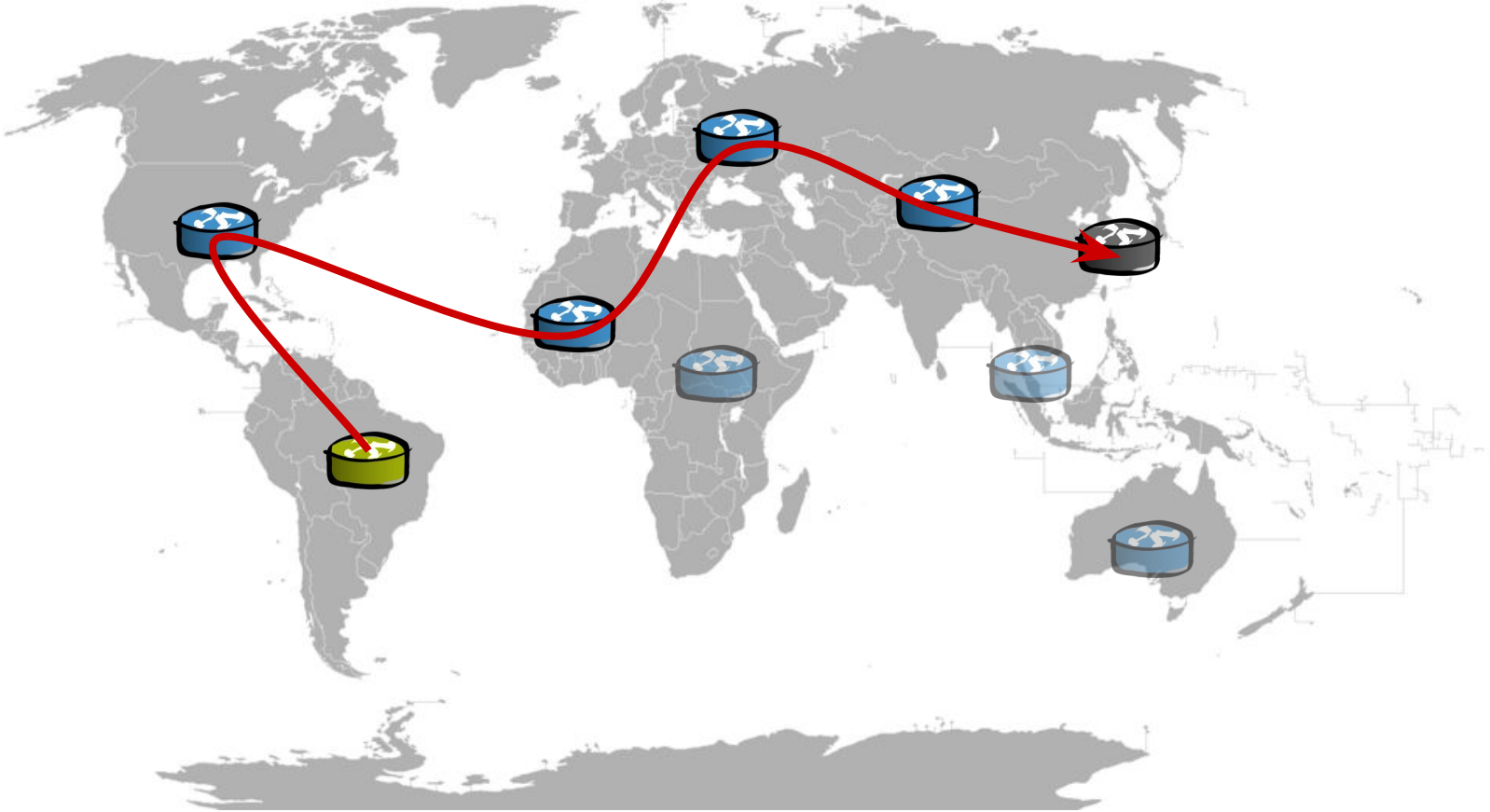
Timur Friedman Ethan Katz-Bassett Ronaldo A. Ferreira

SIGMETRICS 2025
Stony Brook, NY, USA

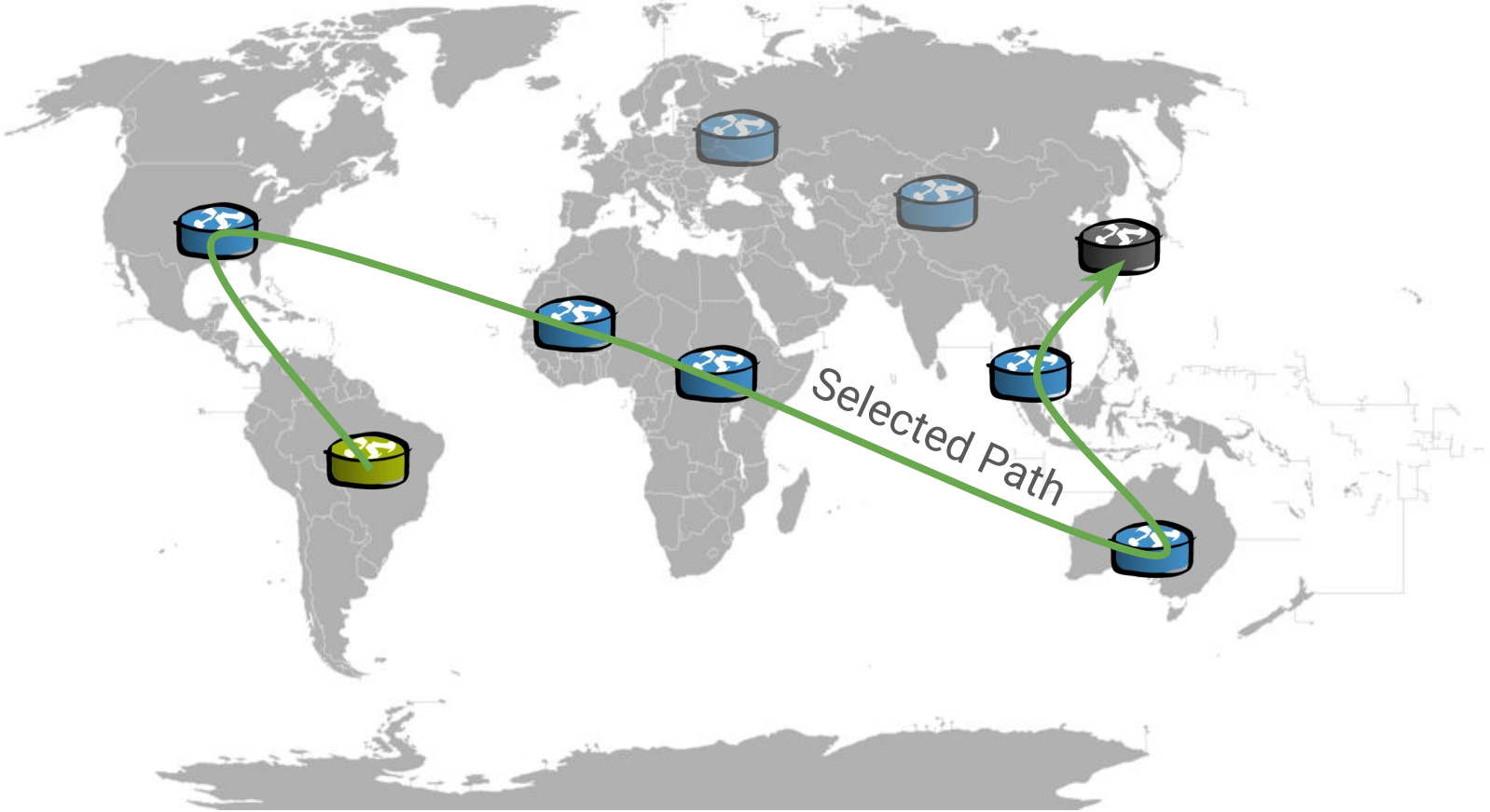
Motivating Example



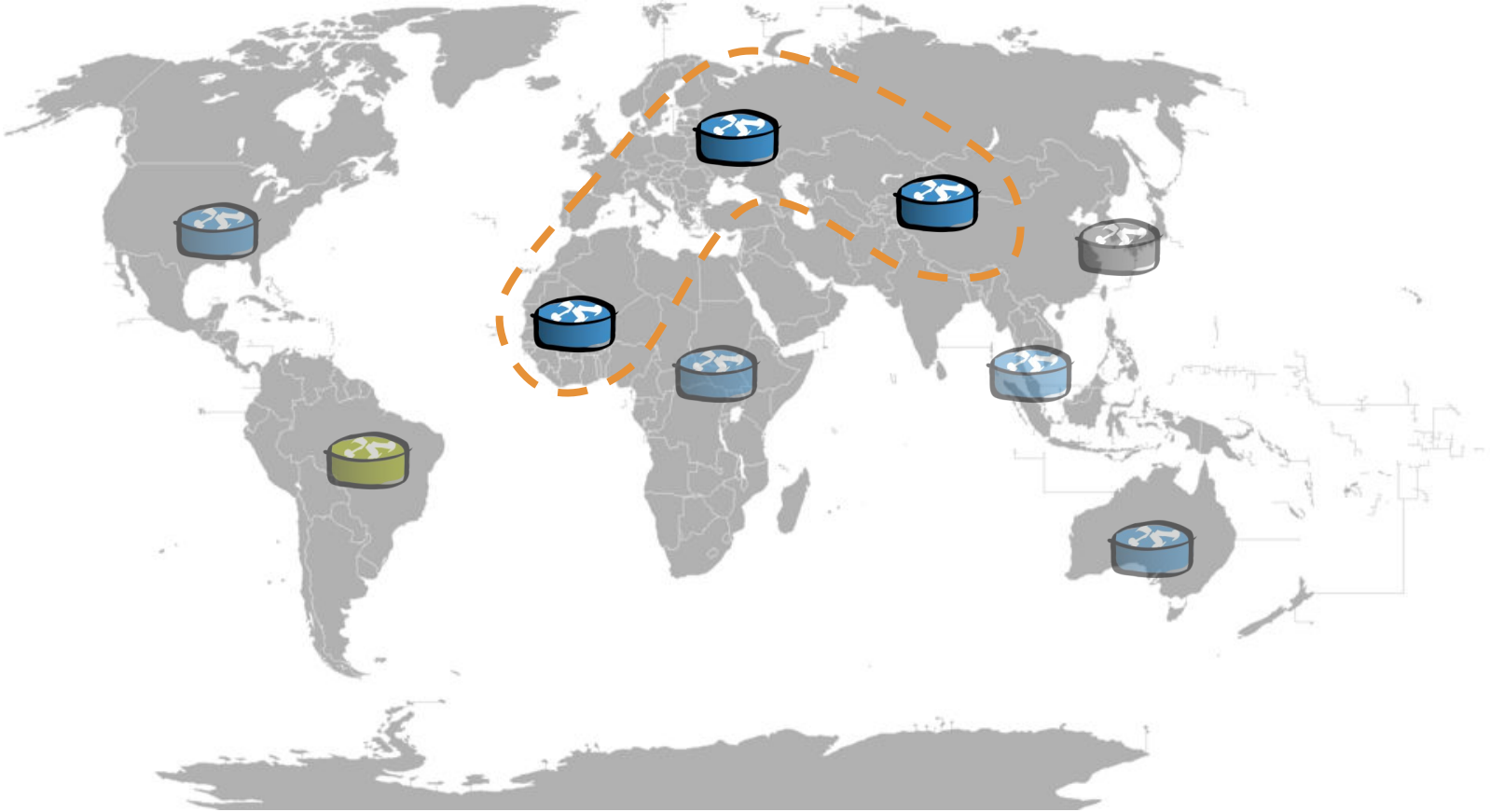
Motivating Example



Motivating Example



Motivating Example



What they have right now to help?

```
TABLE_DUMP2|1733097601|B|187.16.222.221|263009|5.157.139.0/24|263009 13786 2914 60068 212238|IGP|187.16.222.221|0|
TABLE_DUMP2|1733097601|B|187.16.222.168|262791|5.157.139.0/24|262791 53013 6762 60068 212238|IGP|187.16.222.168|0|
TABLE_DUMP2|1733097601|B|187.16.220.216|52873|5.157.139.0/24|52873 12956 60068 212238|IGP|187.16.220.216|0|0|12956
TABLE_DUMP2|1733097601|B|187.16.222.156|264479|5.157.139.0/24|264479 7195 174 60068 212238|IGP|187.16.222.156|0|0|
TABLE_DUMP2|1733097601|B|187.16.217.161|13786|5.157.139.0/24|13786 2914 60068 212238|IGP|187.16.217.161|0|108||NAG
TABLE_DUMP2|1733097601|B|187.16.221.193|37468|5.157.139.0/24|37468 1299 60068 212238|IGP|187.16.221.193|0|0|1299:3
0 37468:14100 37468:14101 37468:37468|NAG||
TABLE_DUMP2|1733097601|B|187.16.220.229|49544|5.157.139.0/24|49544 2914 60068 212238|IGP|187.16.220.229|0|10|2914:
AG||
```

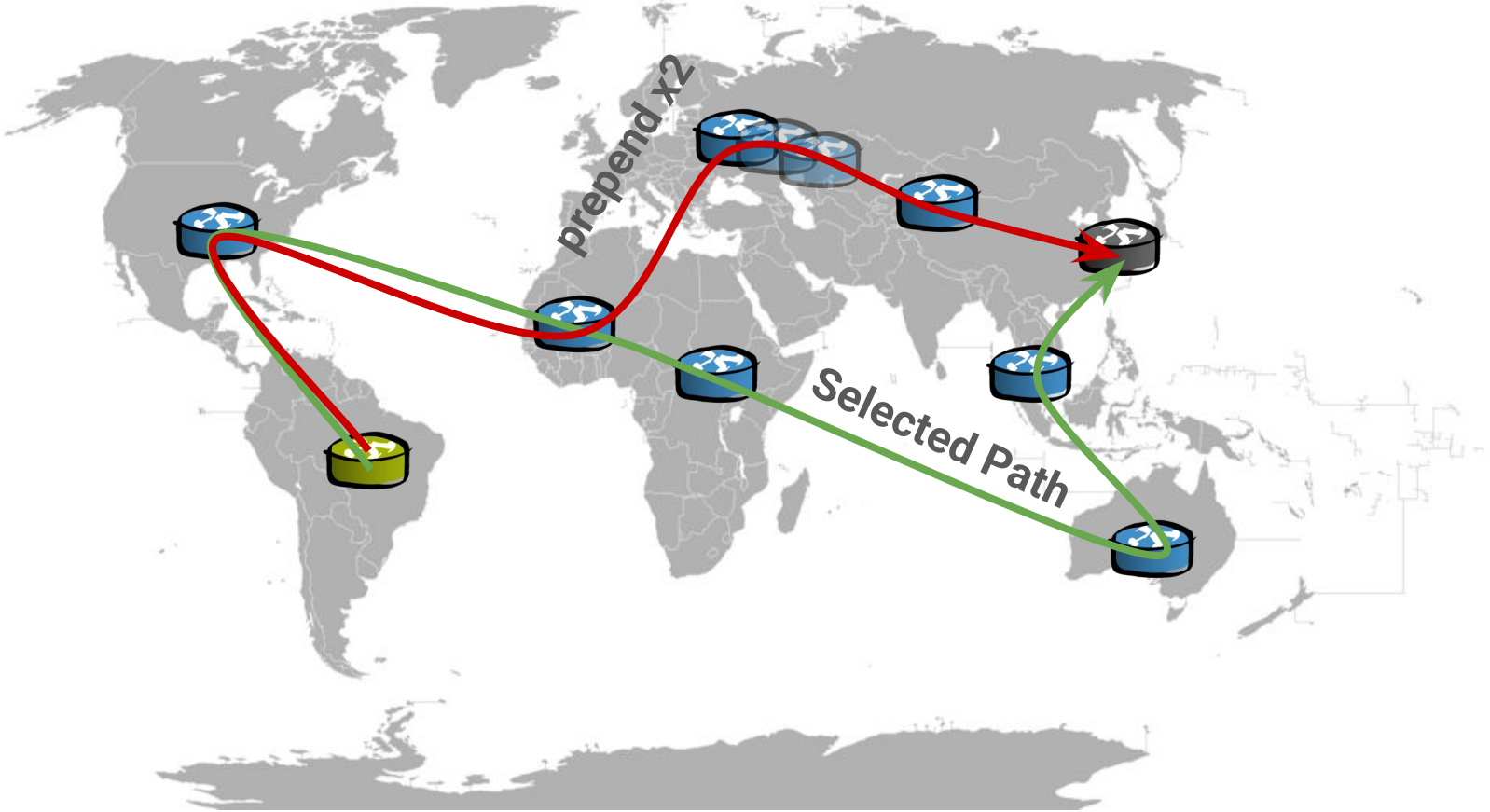
What they have right now to help?

```
TABLE_DUMP2|1733097601|B|187.16.222.221|263009|5.157.139.0/24|263009 13786 2914 60068 212238|IGP|187.16.222.221|0|
TABLE_DUMP2|1733097601|B|187.16.222.168|262791|5.157.139.0/24|262791 53013 6762 60068 212238|IGP|187.16.222.168|0|
TABLE_DUMP2|1733097601|B|187.16.220.216|52873|5.157.139.0/24|52873 12956 60068 212238|IGP|187.16.220.216|0|0|12956
TABLE_DUMP2|1733097601|B|187.16.222.156|264479|5.157.139.0/24|264479 7195 174 60068 212238|IGP|187.16.222.156|0|0|
TABLE_DUMP2|1733097601|B|187.16.217.161|13786|5.157.139.0/24|13786 2914 60068 212238|IGP|187.16.217.161|0|108||NAG
TABLE_DUMP2|1733097601|B|187.16.221.193|37468|5.157.139.0/24|37468 1299 60068 212238|IGP|187.16.221.193|0|0|1299:3
0 37468:14100 37468:14101 37468:37468 NAG||
TABLE_DUMP2|1733097601|B|187.16.220.229|49544|5.157.139.0/24|49544 2914 60068 212238|IGP|187.16.220.229|0|10|2914:
AG||
```

37468:37468

Observation:
Operators need additional information to understand Internet routing

Motivating Example

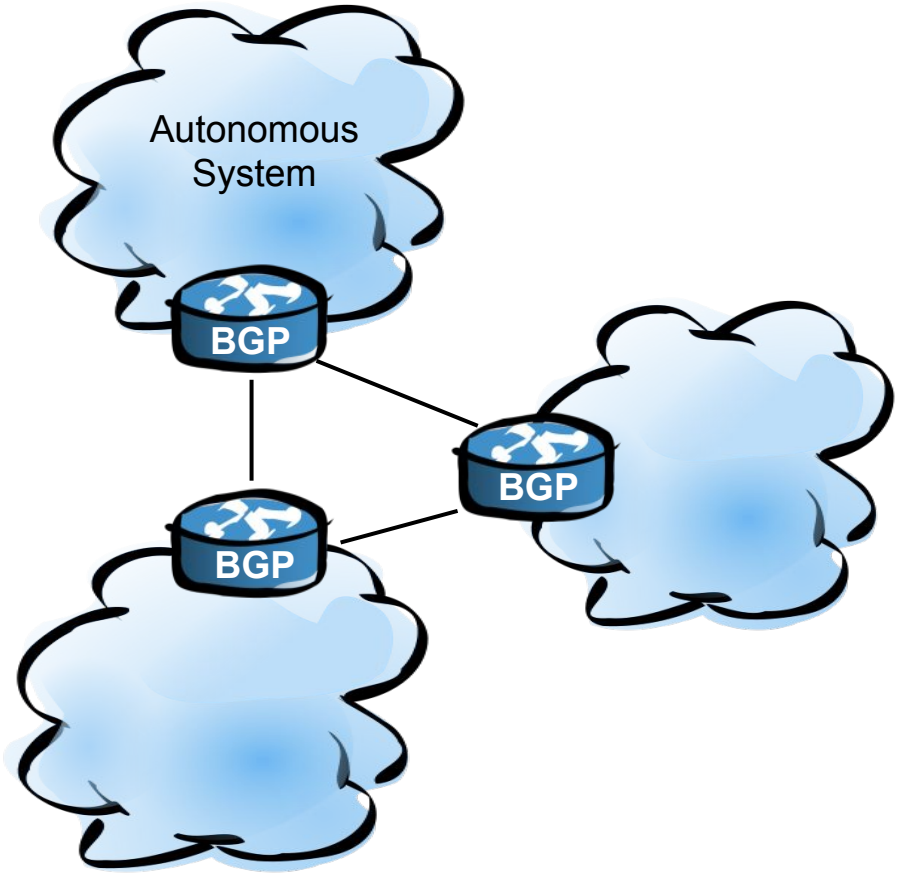


Problem Statement

How can we determine the semantics of BGP communities from routes observed by the route collectors?

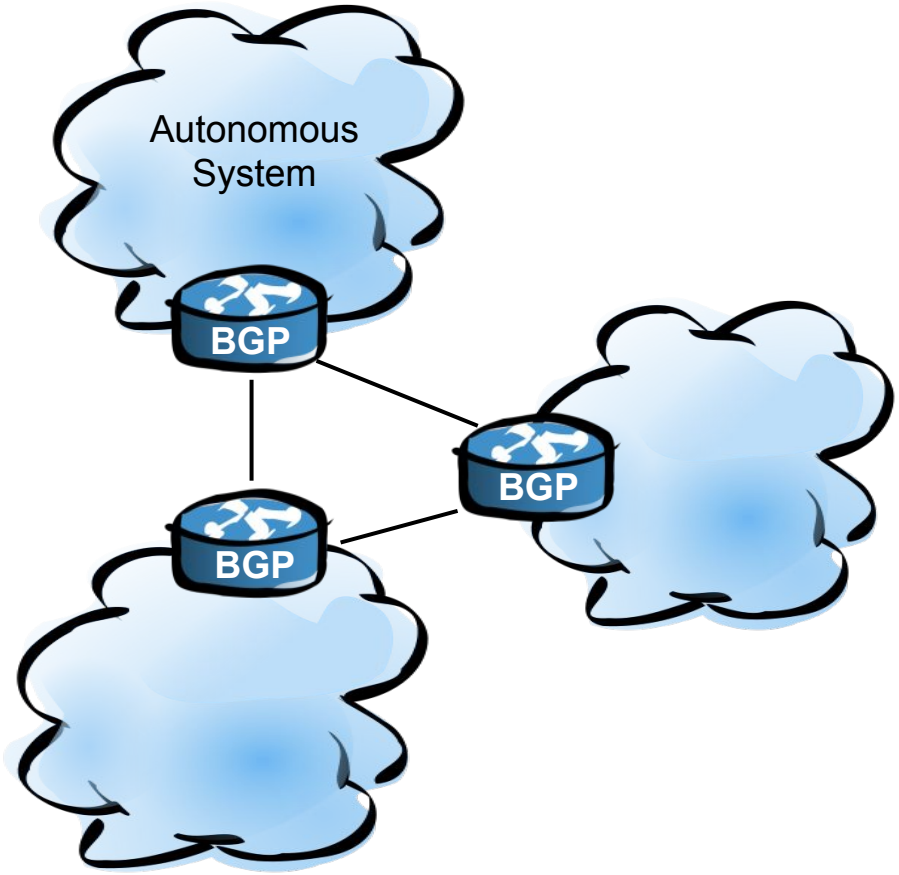
Internet Evolution

- BGP: the routing protocol used to exchange Internet routes and reachability information.



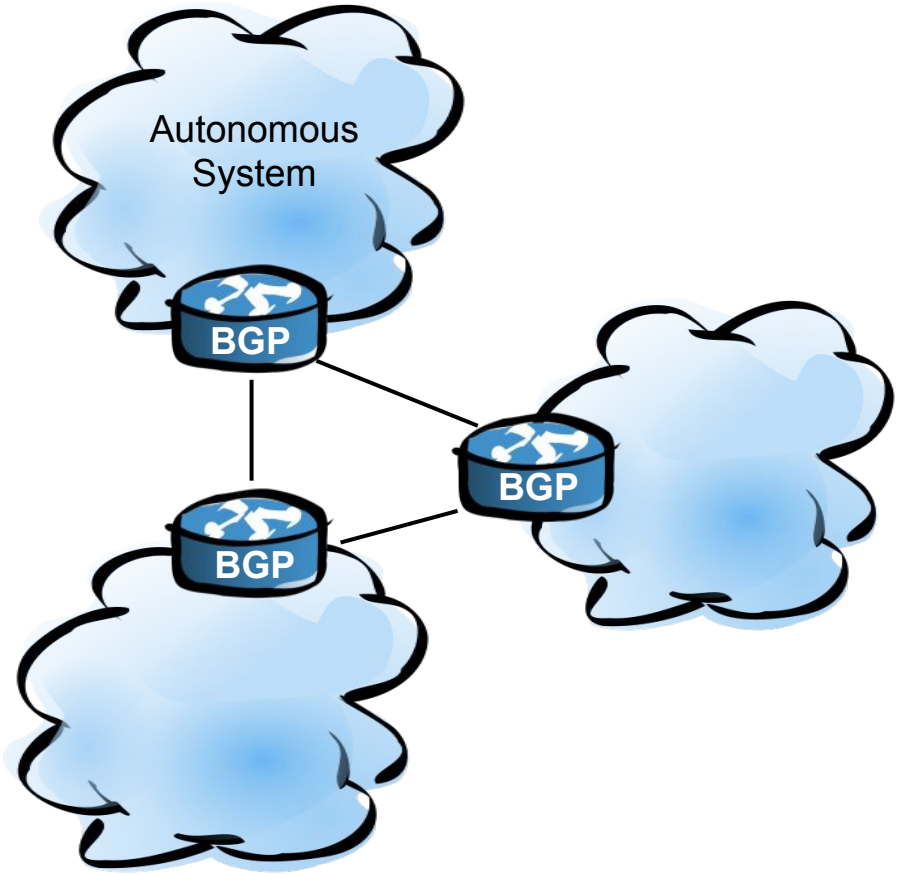
Internet Evolution

- BGP: the routing protocol used to exchange Internet routes and reachability information.
- BGP is old but allows complex policies supports stability and reliability goals.



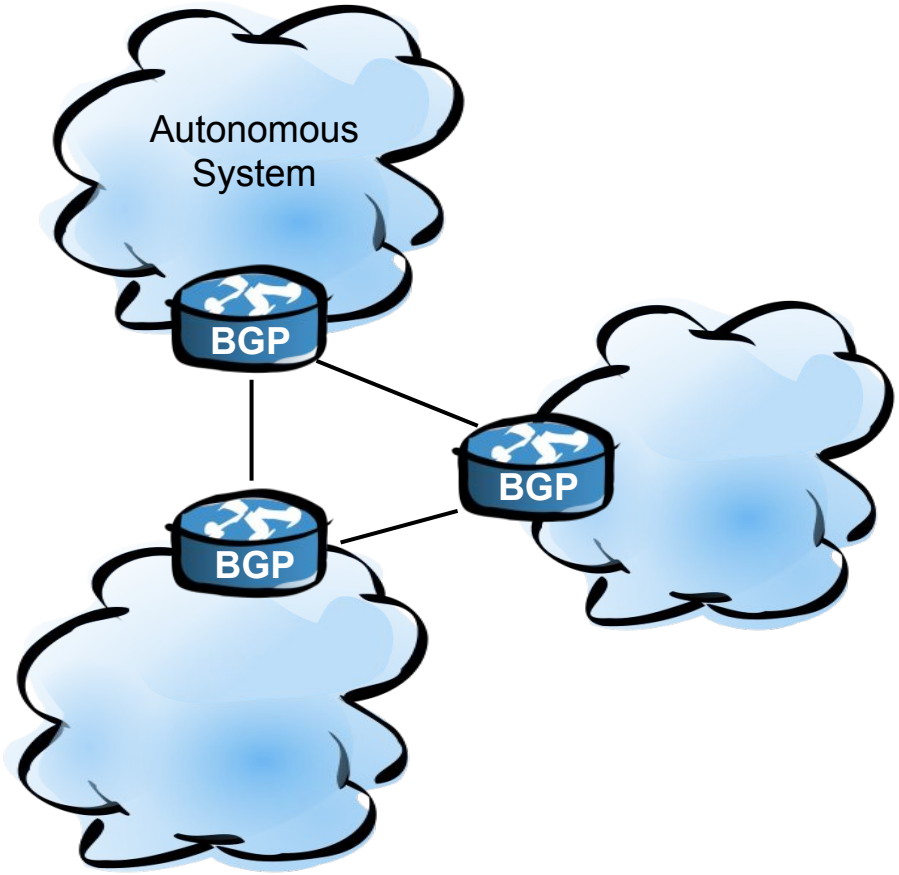
Internet Evolution

- BGP: the routing protocol used to exchange Internet routes and reachability information.
- BGP is old but allows complex policies supports stability and reliability goals.
- Operators explore all available options to improve the “user experience”.



Internet Evolution

- BGP: the routing protocol used to exchange Internet routes and reachability information.
- BGP is old but allows complex policies supports stability and reliability goals.
- Operators explore all available options to improve the “user experience”.
- Network operators have growing reliance on BGP communities to enrich policies.



The problem is

- Community semantics are not standardized

The problem is

- ❏ Community semantics are not standardized
- ❏ Hard to know the semantics of a specific community

Is the documentation reliable and up to date?

What exactly does this community do?

The problem is

- ❏ Community semantics are not standardized
- ❏ Hard to know the semantics of a specific community
- ❏ Hard to know the community for some specific goal

Is the documentation reliable and up to date?

What exactly does this community do?

Where is the documentation related to that community?

The problem is

- ❏ Community semantics are not standardized
- ❏ Hard to know the semantics of a specific community
- ❏ Hard to know the community for some specific goal
- ❏ We have two types of communities

Is the documentation reliable and up to date?

What exactly does this community do?

Is this an information or action community?

Where is the documentation related to that community?

Inferring BGP Action Communities

Inference of Action Communities

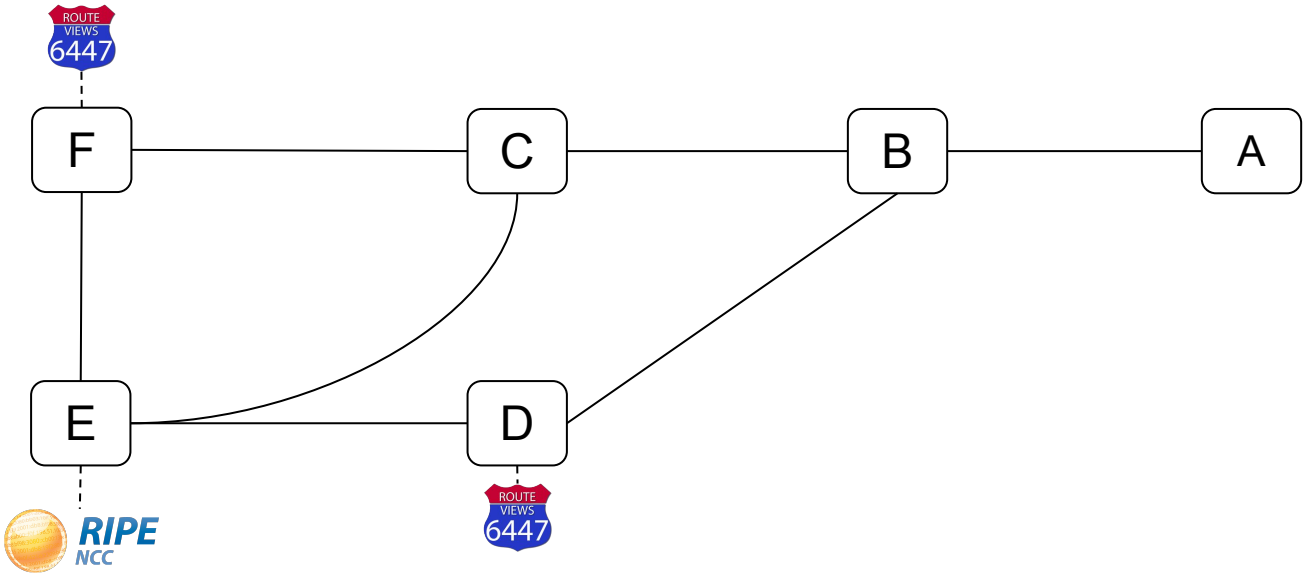
- Our inference algorithm quantifies how often a community co-occurs with its controlling AS.
  - Frequent co-occurrences: information community
  - Few co-occurrences, possibly noise: action community

Inference of Action Communities

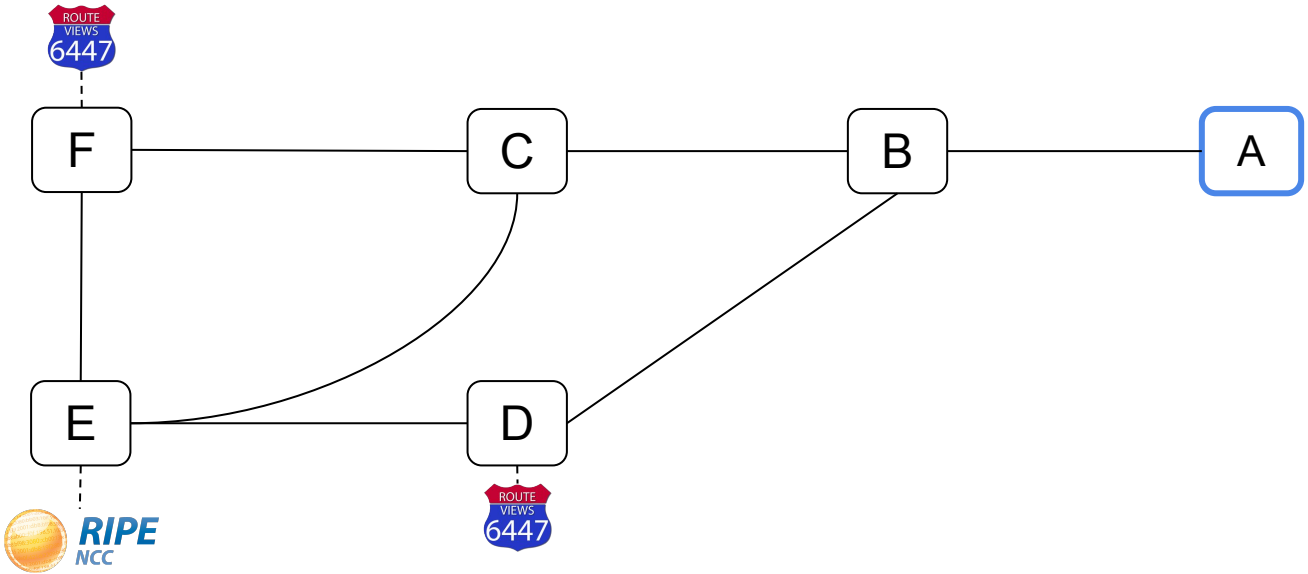
- Our inference algorithm quantifies how often a community co-occurs with its controlling AS.
  - Frequent co-occurrences: information community
  - Few co-occurrences, possibly noise: action community

- Also, action communities should be used sparingly, while information communities should appears in all paths.

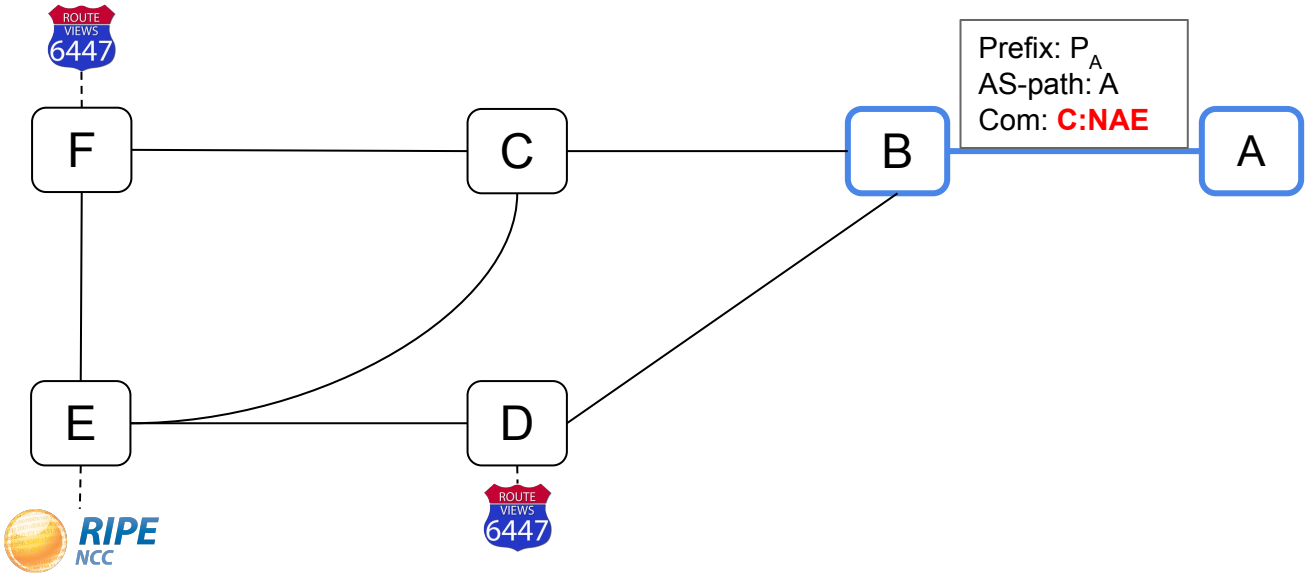
Communities Appearing **without** Controlling AS



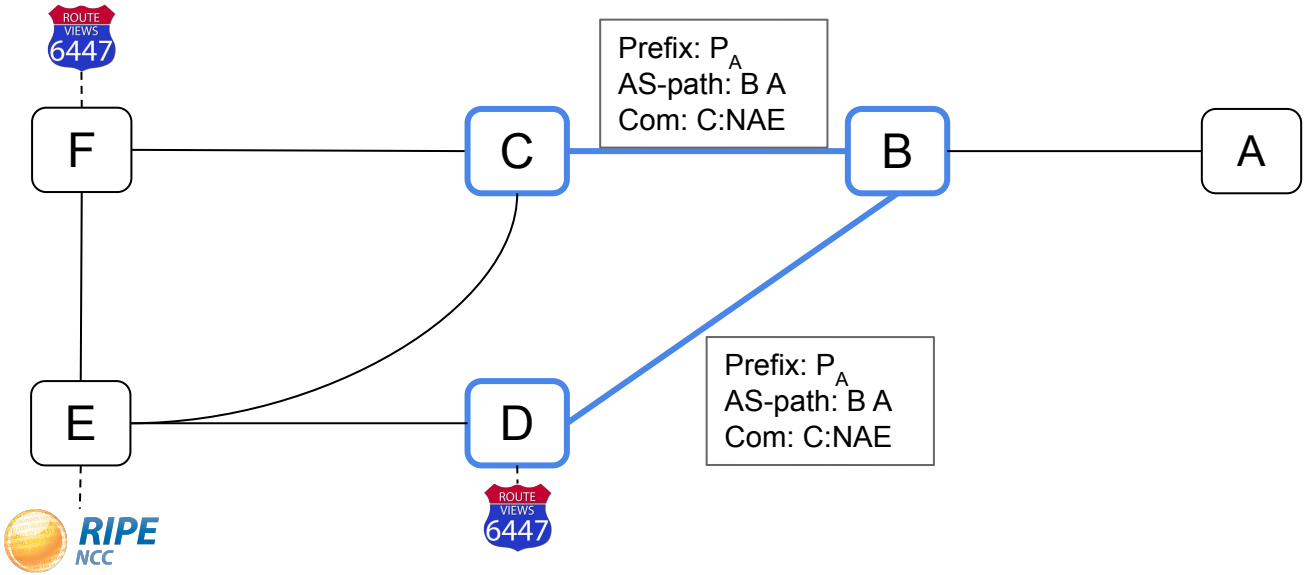
Communities Appearing without Controlling AS



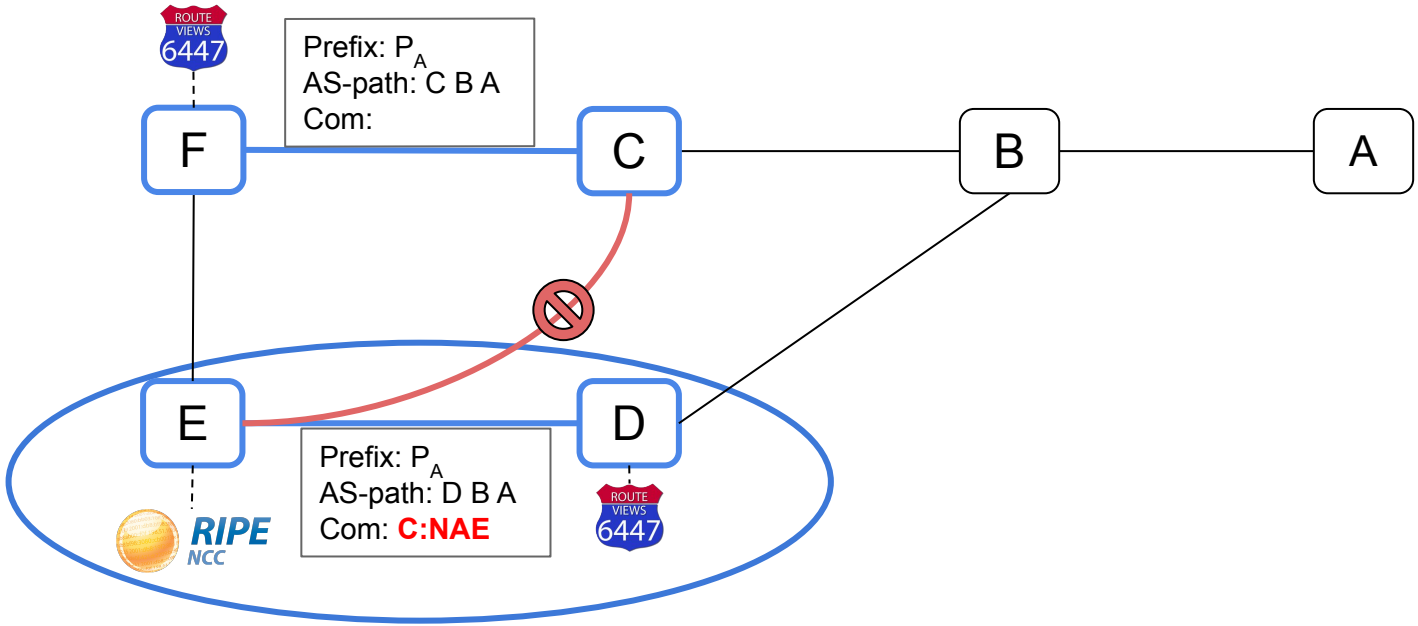
Communities Appearing without Controlling AS



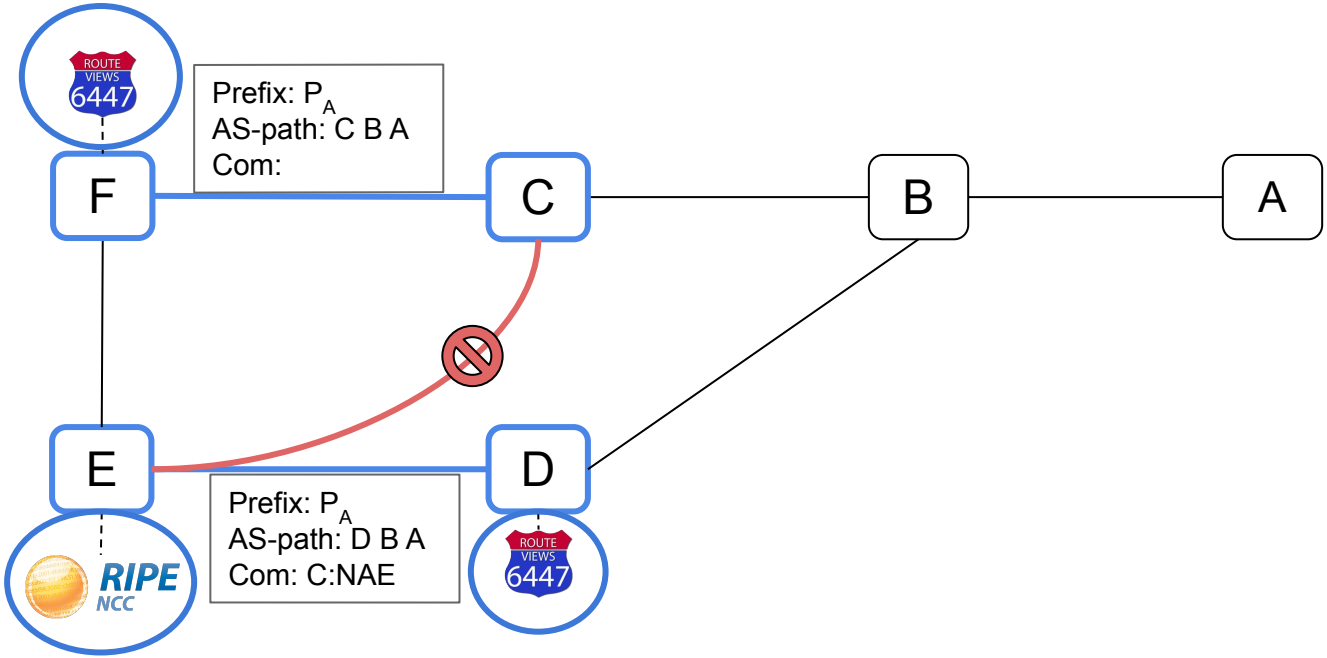
Communities Appearing without Controlling AS



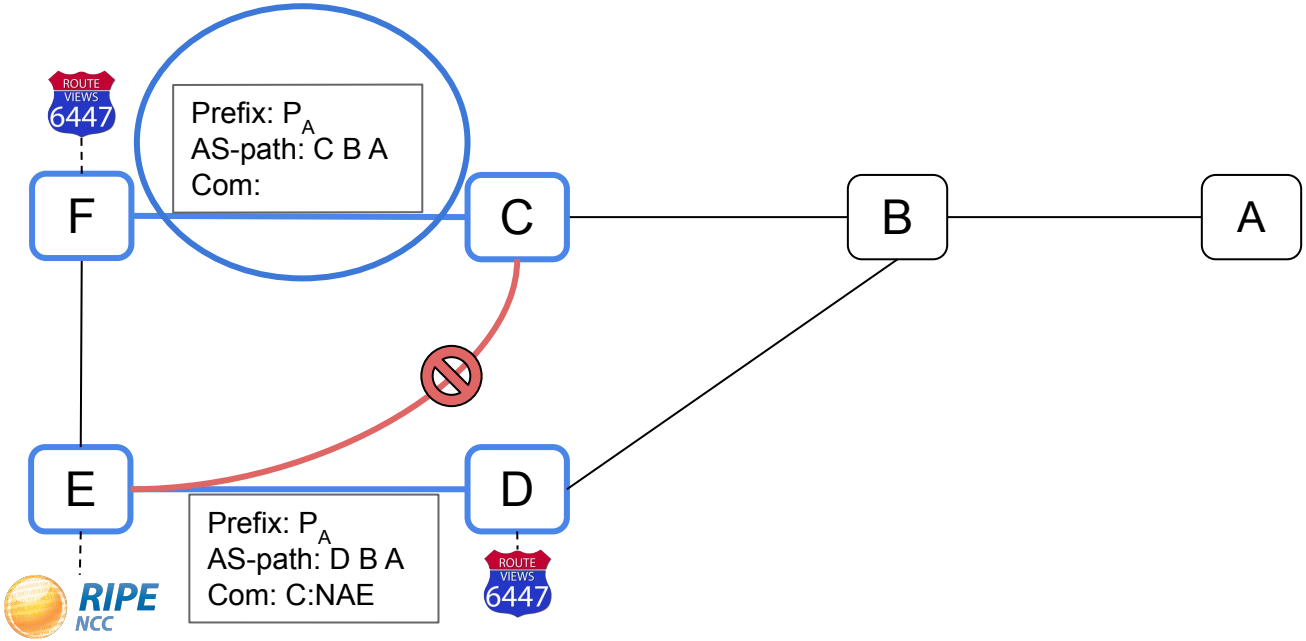
Communities Appearing without Controlling AS



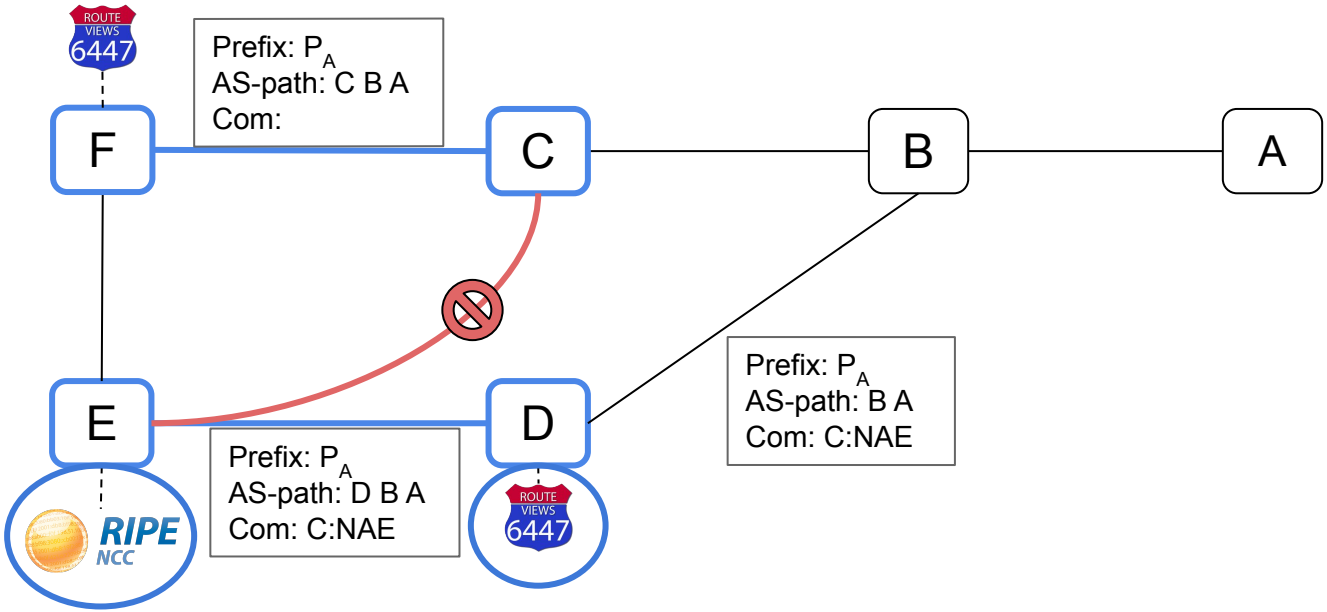
Communities Appearing without Controlling AS



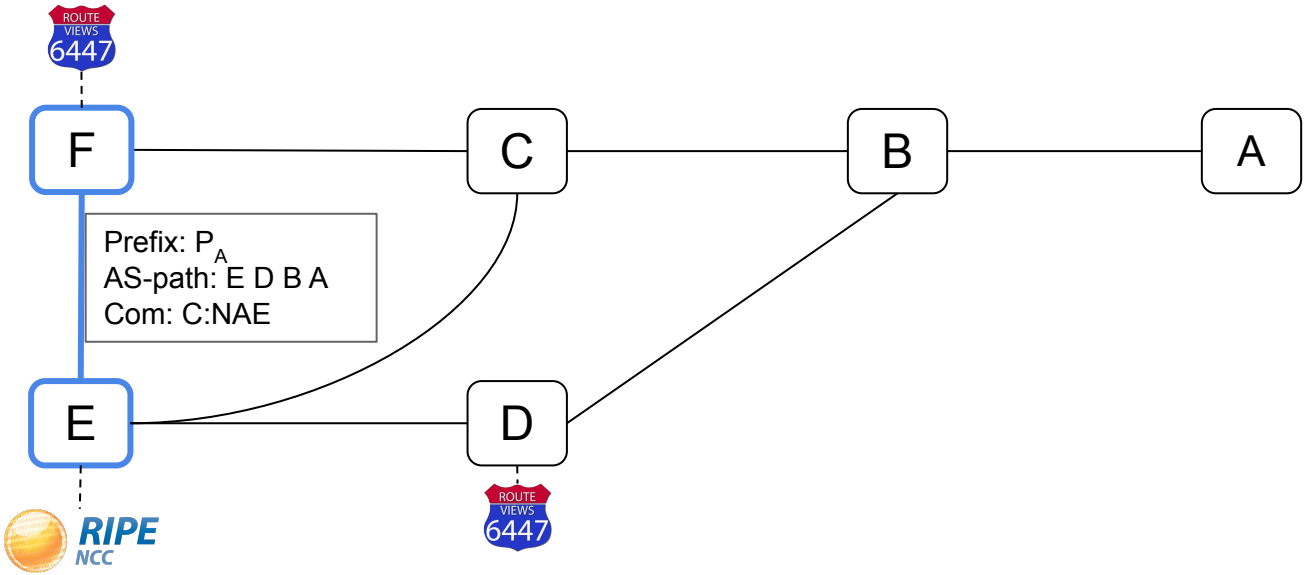
Communities Appearing without Controlling AS



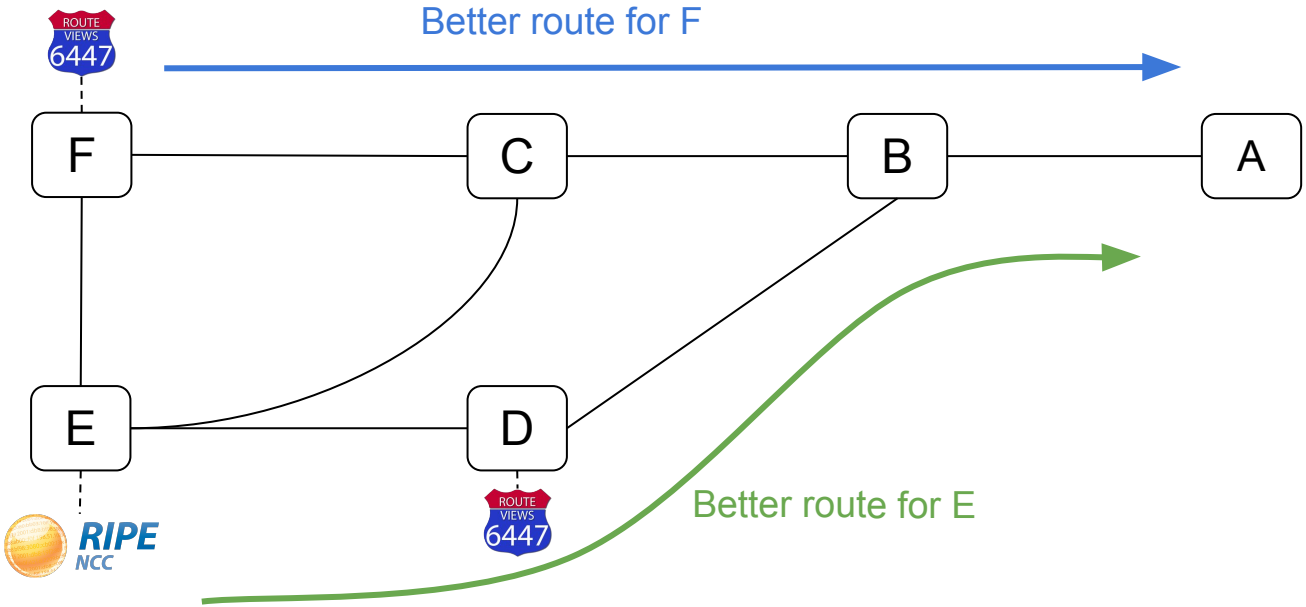
Communities Appearing without Controlling AS



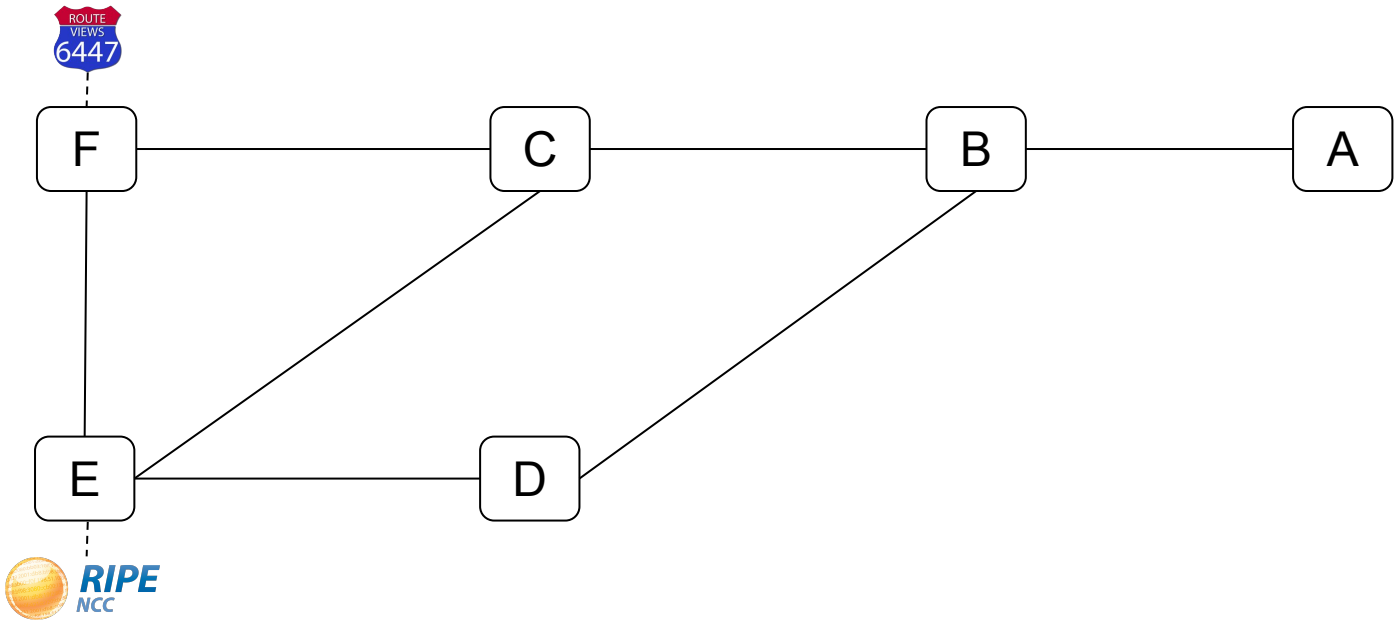
Communities Appearing without Controlling AS



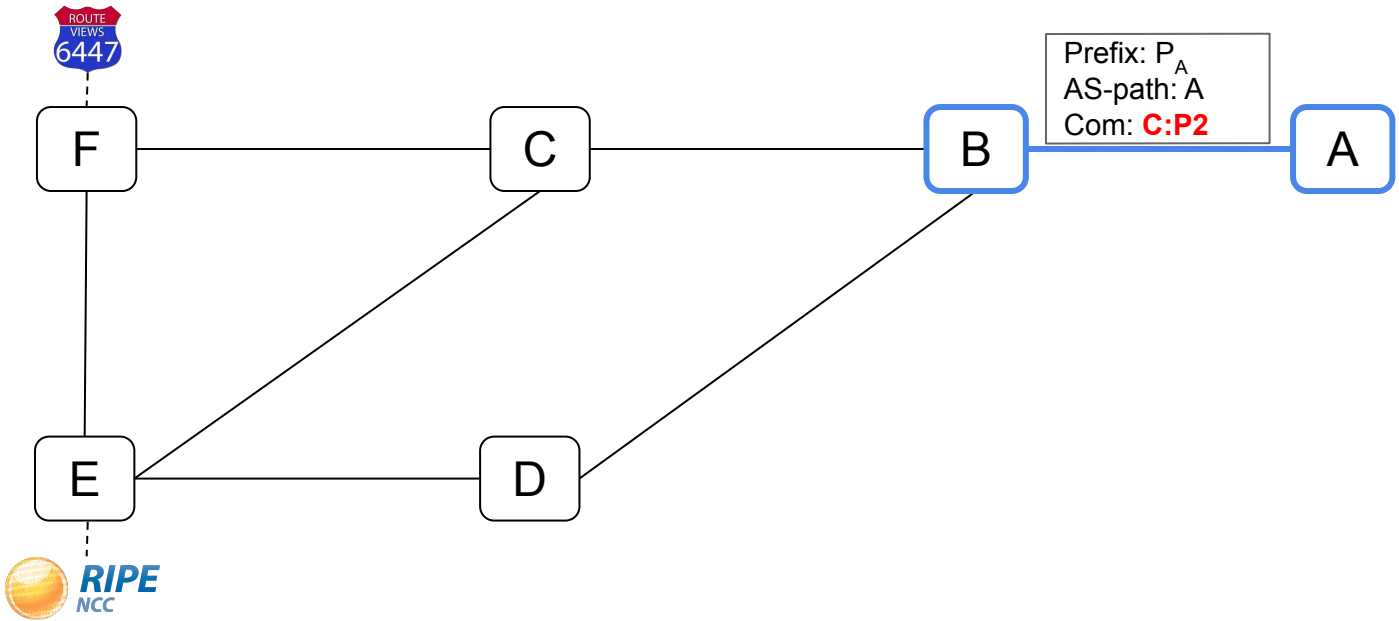
Communities Appearing without Controlling AS



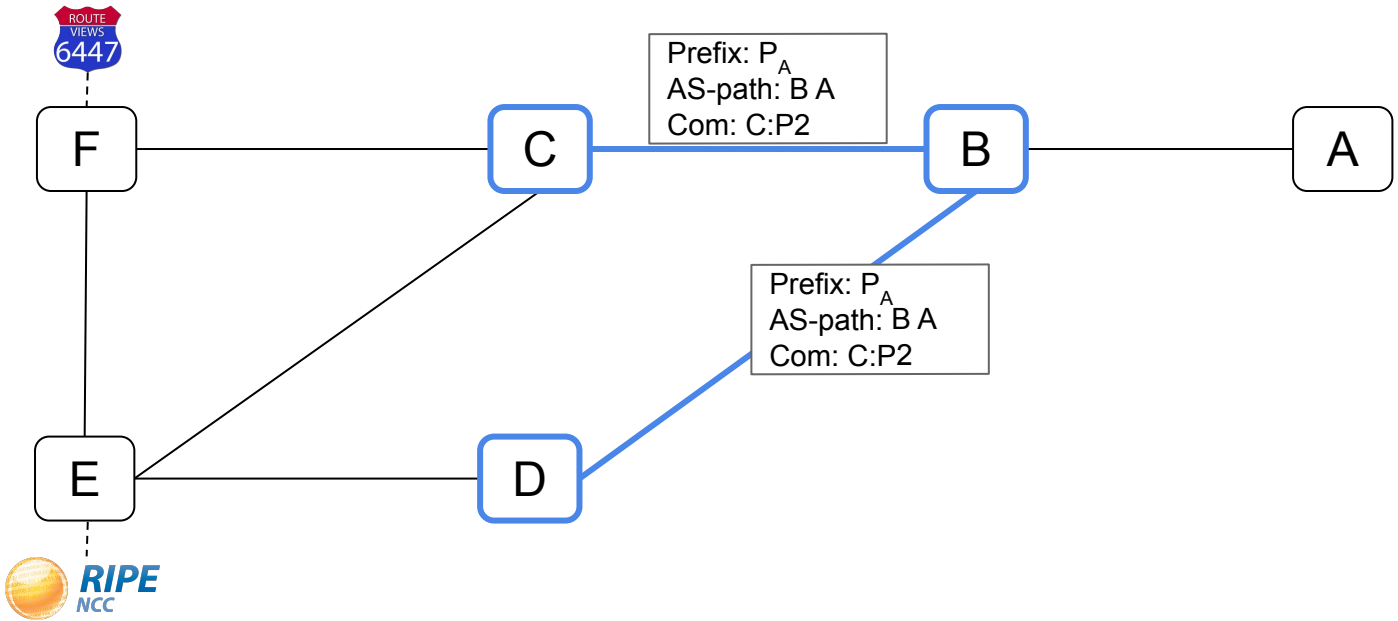
Communities Appearing **with** Controlling AS



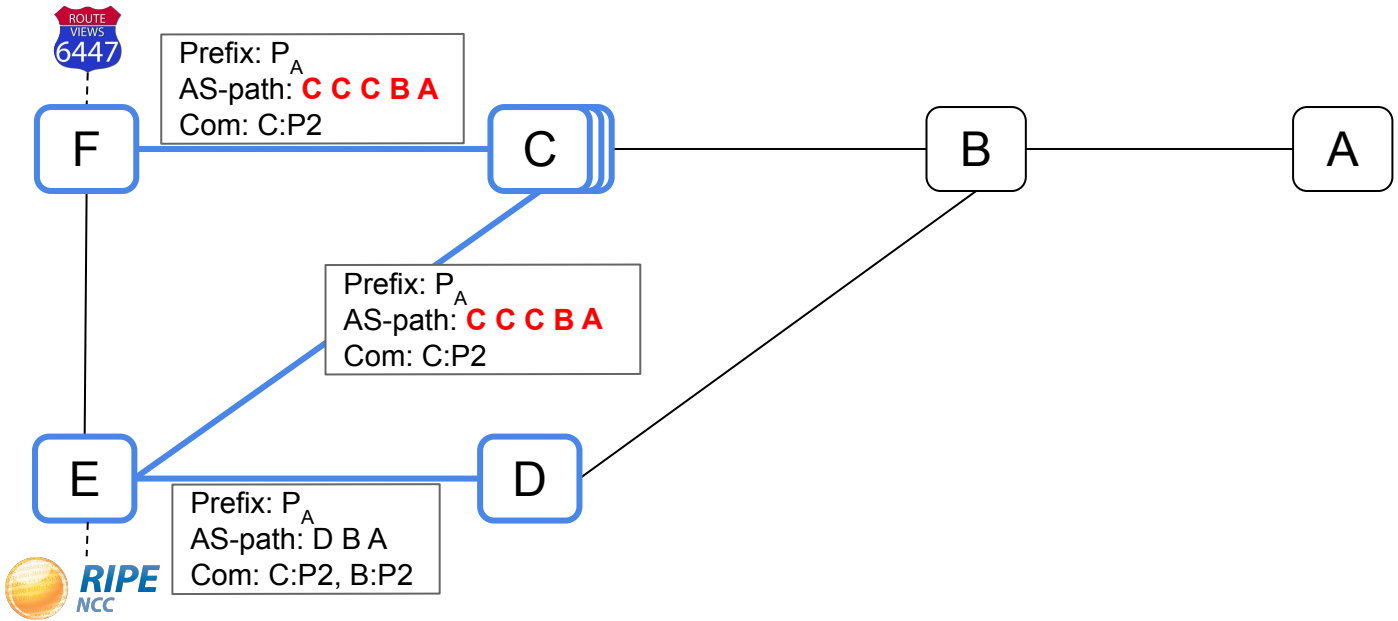
Communities Appearing with Controlling AS



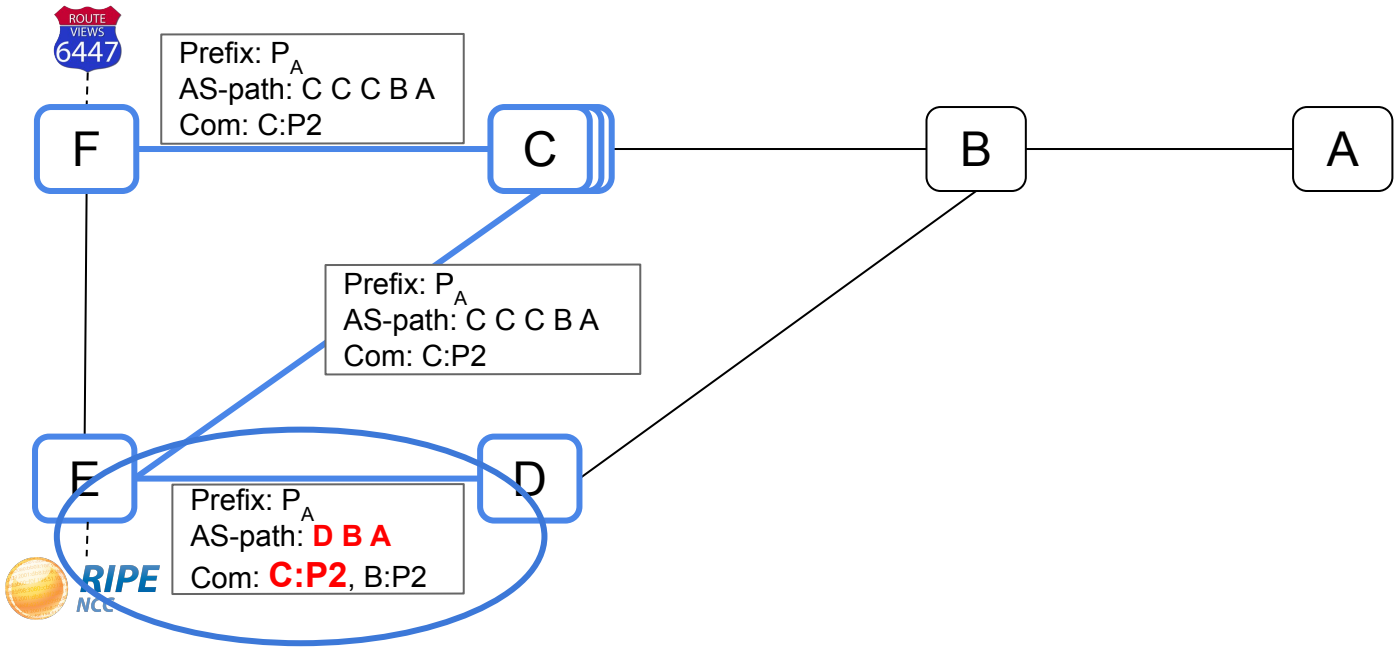
Communities Appearing with Controlling AS



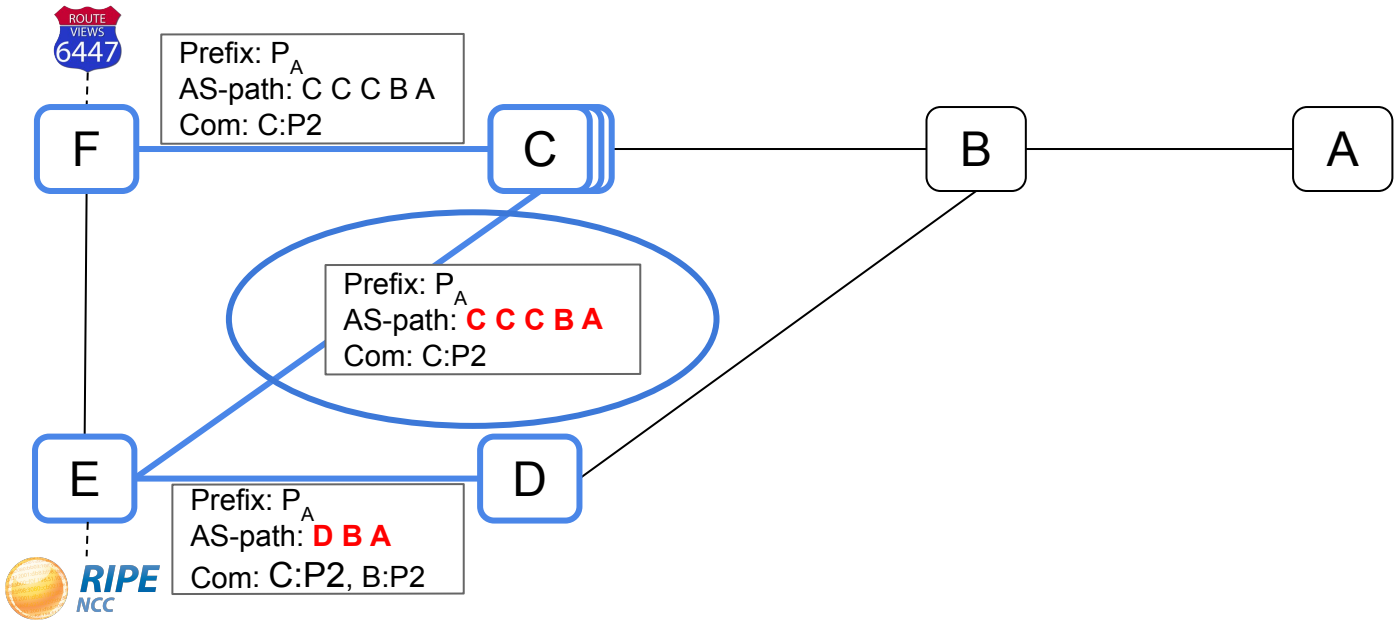
Communities Appearing with Controlling AS



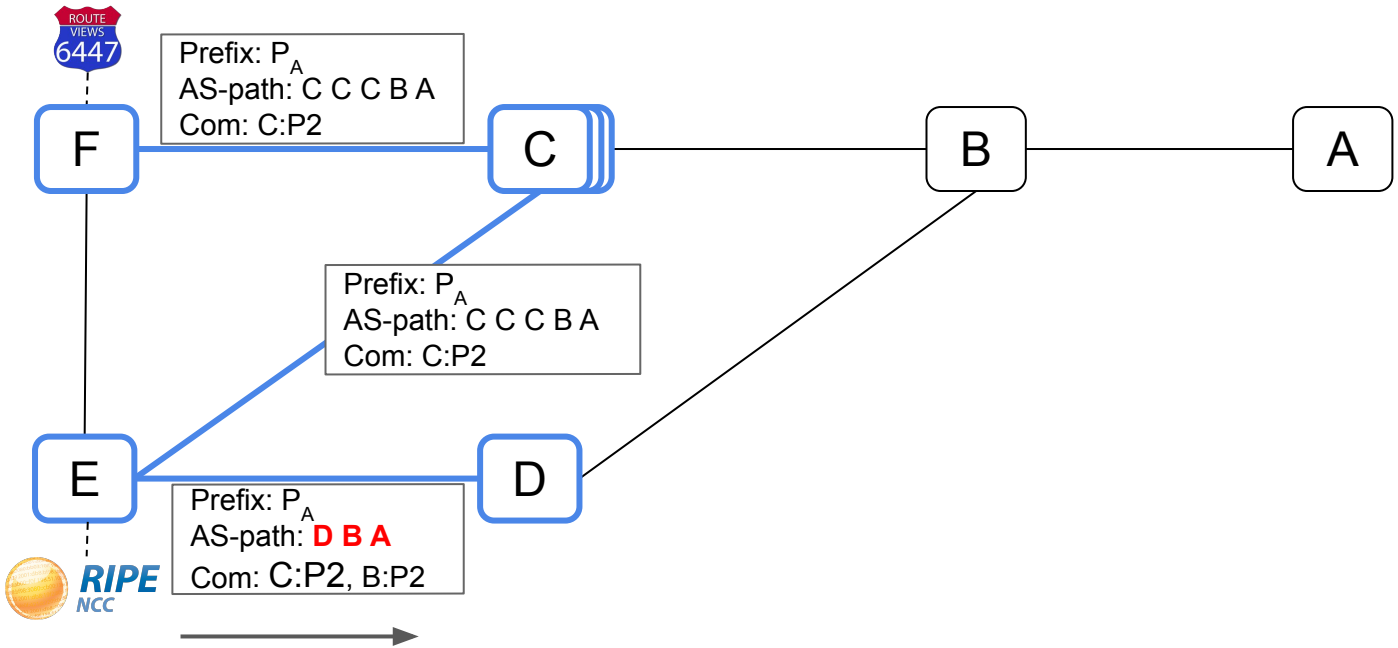
Communities Appearing with Controlling AS



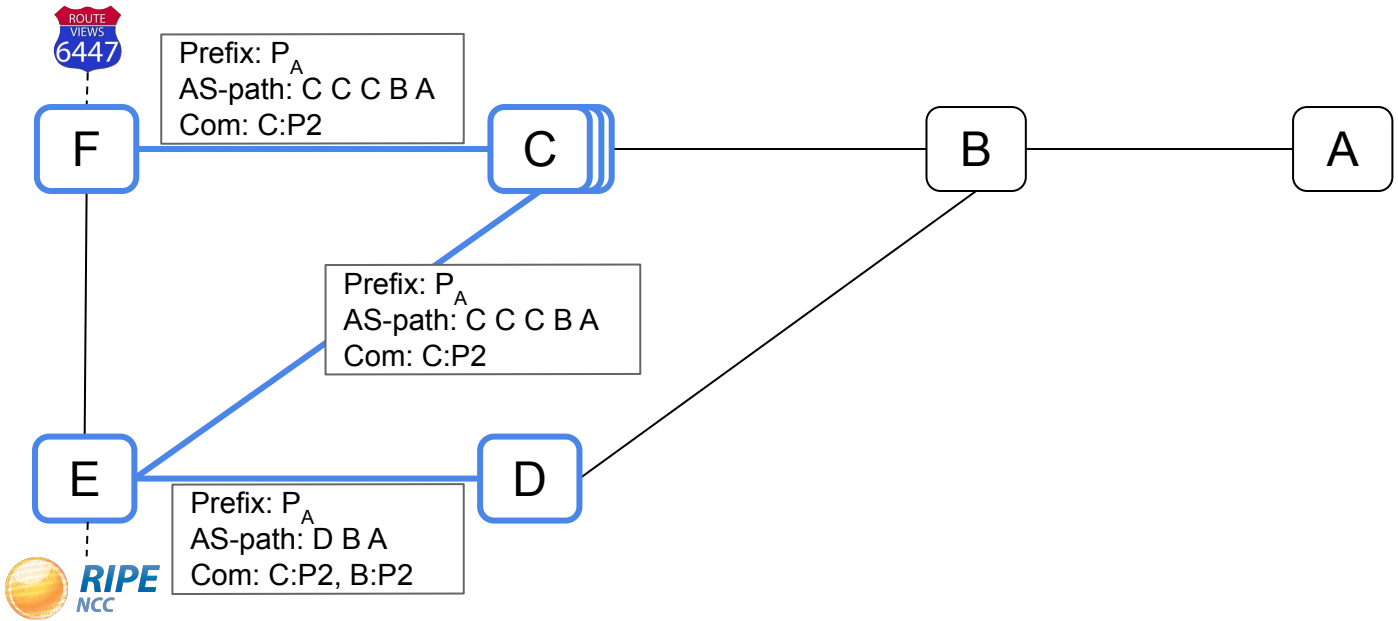
Communities Appearing with Controlling AS



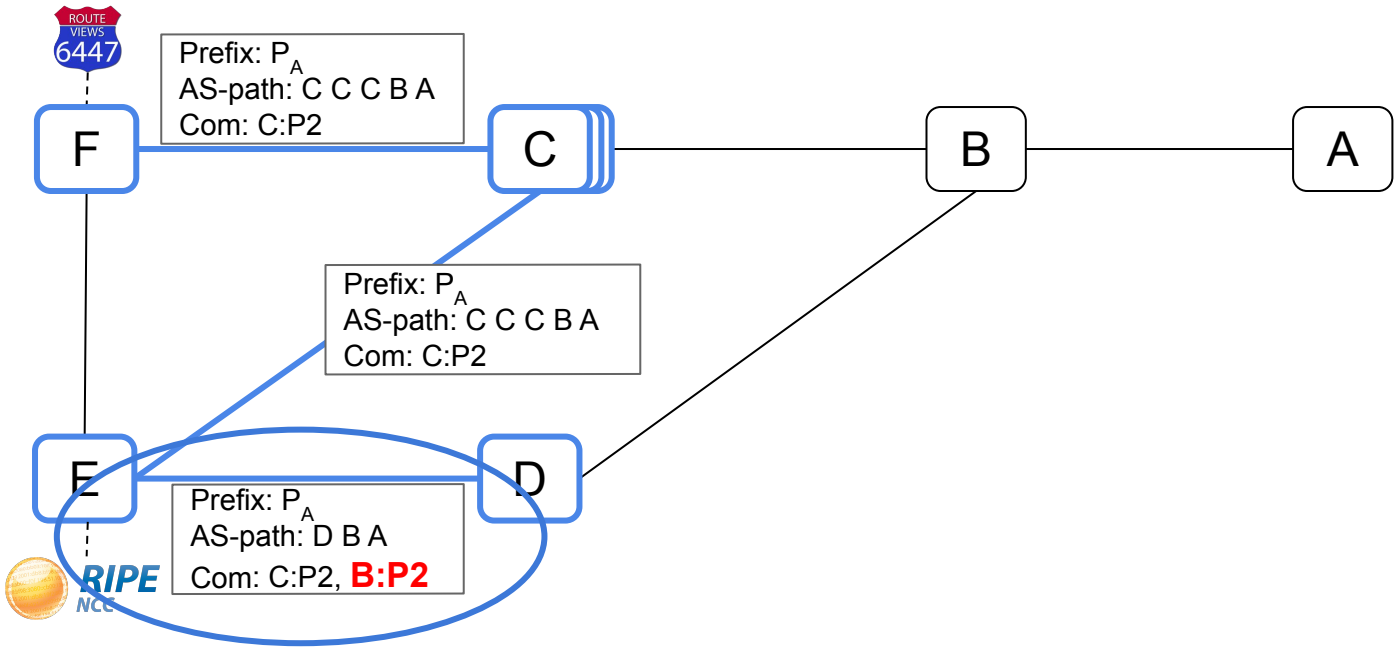
Communities Appearing with Controlling AS



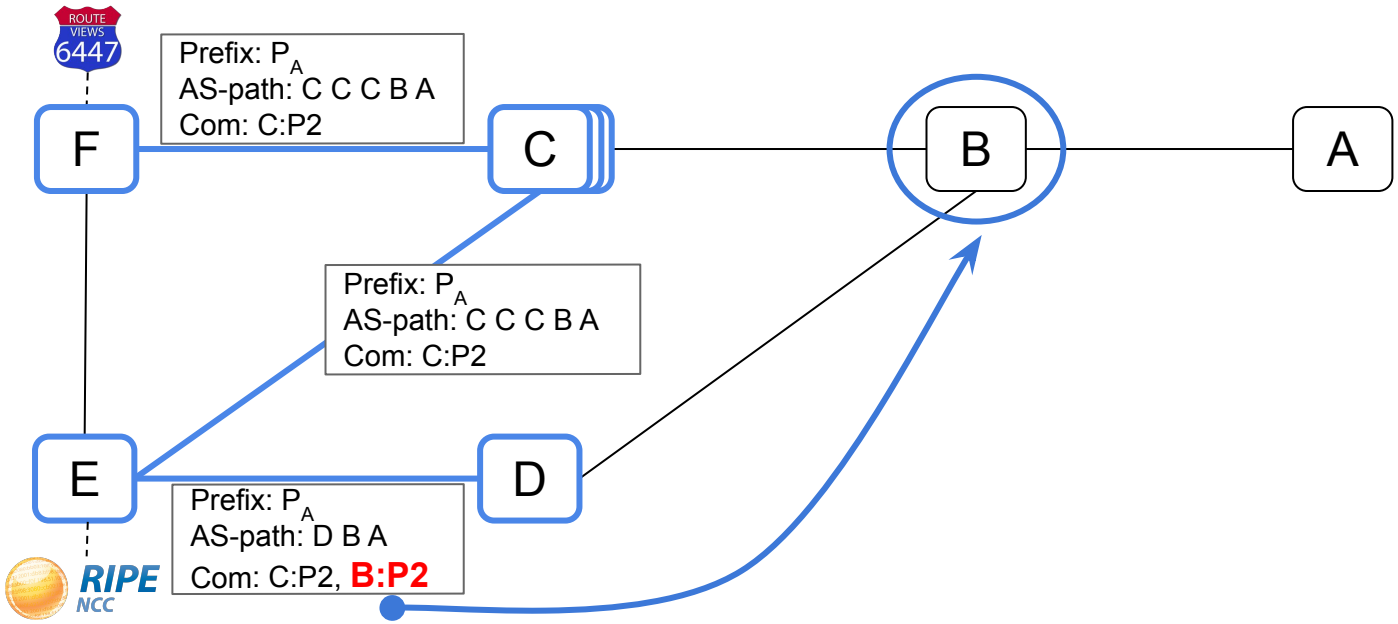
Communities Appearing with Controlling AS



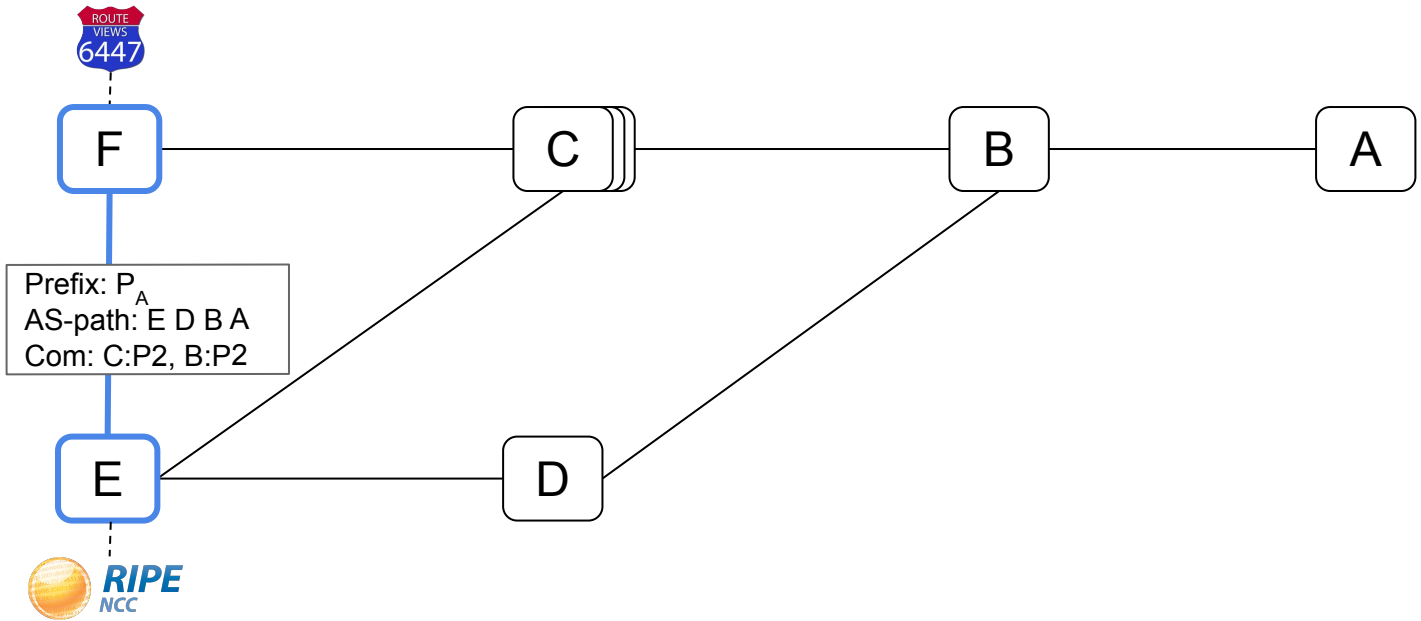
Communities Appearing with Controlling AS



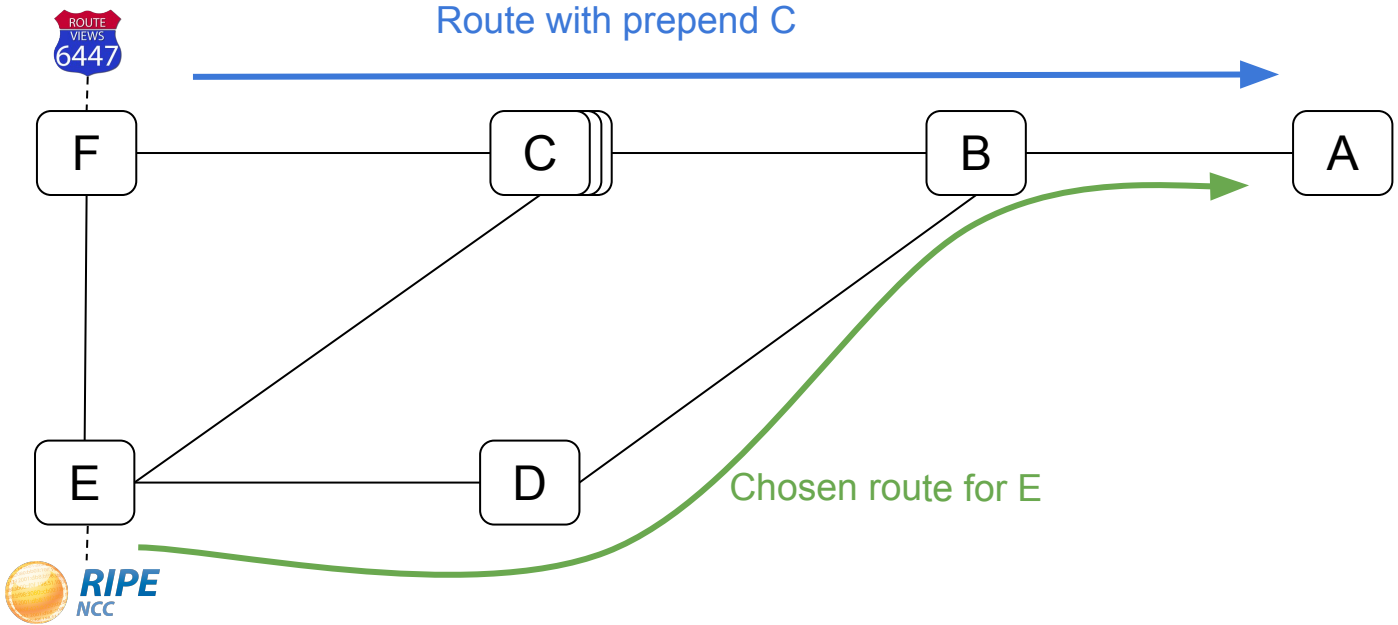
Communities Appearing with Controlling AS



Communities Appearing with Controlling AS



Communities Appearing with Controlling AS



AS Squatters

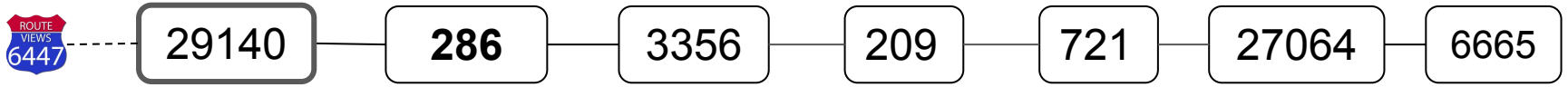
AS Squatting Relations

- During the evaluation we saw that some communities appears constantly in the announcements (more likely information communities). But without its controlling ASes.

- We conjecture this occurs because ASes borrow (squat) information communities from other ASes.

- For example, AS286 (KPN) apparently uses information communities from AS3257 (GTT), which makes sense as these networks have merged.

AS-path: 29140 **286** 3356 209 721 27064 6665
Com: **3257:8794**



AS Squatting Relations

- Our goal was to identify every AS X that *systematically* tags routes with BGP information communities controlled by another AS Y.

AS Squatting Relations

- Our goal was to identify every AS X that *systematically* tags routes with BGP information communities controlled by another AS Y.

- The challenge lies in differentiating between:
  - an AS X squatting AS Y's information communities;
  - an AS X simply using AS Y's action communities.

AS Squatting Relations

- Our goal was to identify every AS X that *systematically* tags routes with BGP information communities controlled by another AS Y.

- The challenge lies in differentiating between:
  - an AS X squatting AS Y's information communities;
  - an AS X simply using AS Y's action communities.

- We also have special cases:
  - Typo Error (AS15985 tagging communities starting with 15895:XXX).
  - Integer overflow (4-byte ASNs do not fit in BGP communities).

Additional Challenges

- An operator may also define non-standard BGP communities, where the first 2 bytes are set to a value different than the controlling AS's number.
- AS9002 (RETN) uses community X:65533 as an action community that asks "prepend AS9002 three times when exporting the route to AS X."
- Our algorithm would correctly infer the action communities but associate them with incorrect controlling ASes.

Uncovering Missing Action Communities

Capturing Low-Visibility Communities

- Our inference algorithm requires a minimum number of announcements carrying a community to classify it as an action community with high confidence.

Capturing Low-Visibility Communities

- Our inference algorithm requires a minimum number of announcements carrying a community to classify it as an action community with high confidence.

- However, route collectors do not provide complete coverage of the Internet.

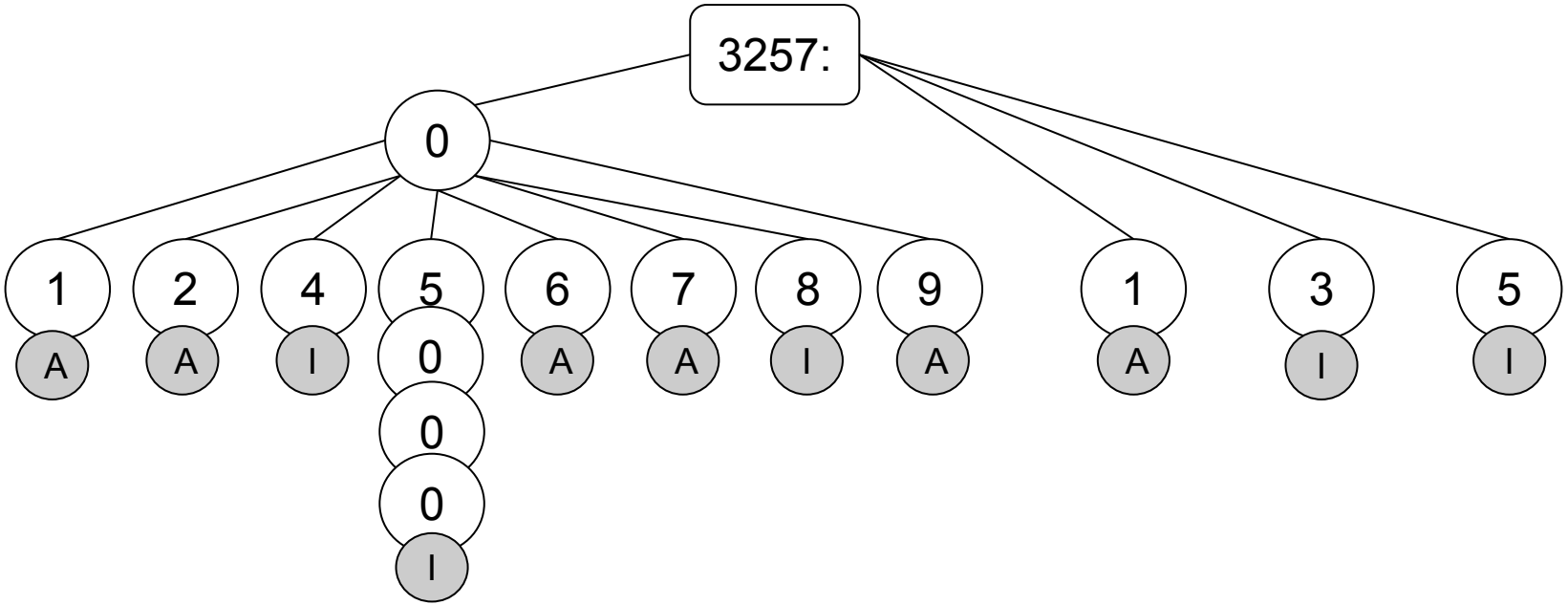
Capturing Low-Visibility Communities

- Our inference algorithm requires a minimum number of announcements carrying a community to classify it as an action community with high confidence.

- However, route collectors do not provide complete coverage of the Internet.

- To circumvent this limitation, we use the communities we infer with high confidence to build a **prefix tree**. Then, we classify other communities with low visibility.

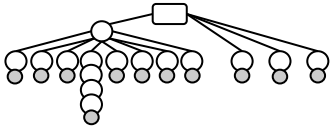
Prefix-Tree representation of communities from an AS



Prefix-Tree in Action

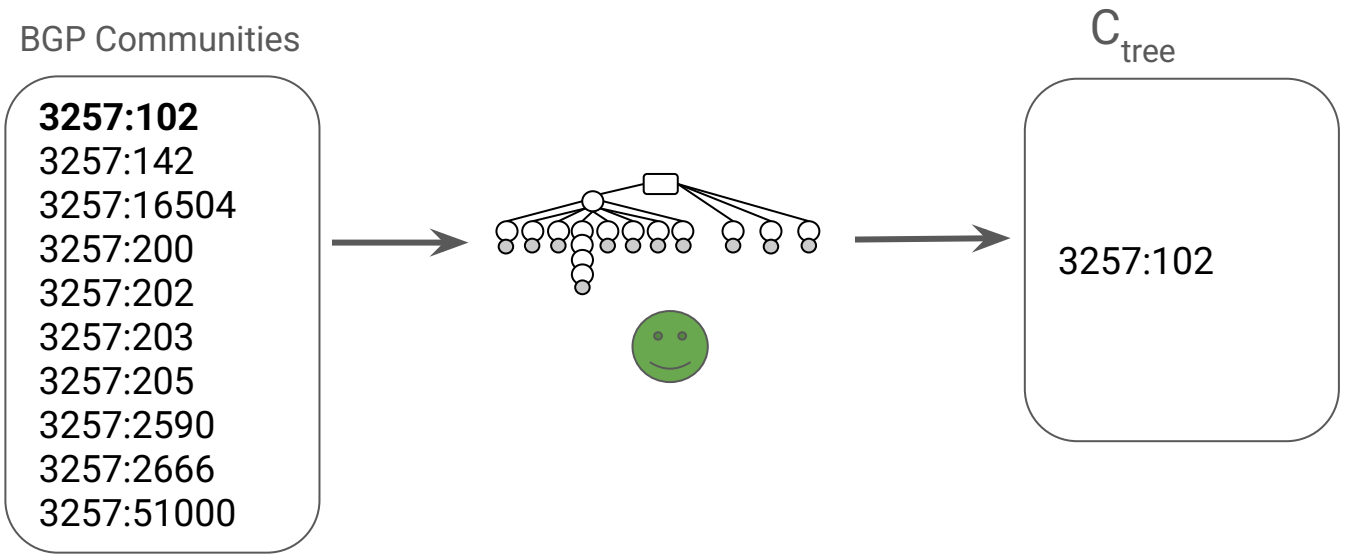
BGP Communities

3257:102
3257:142
3257:16504
3257:200
3257:202
3257:203
3257:205
3257:2590
3257:2666
3257:51000

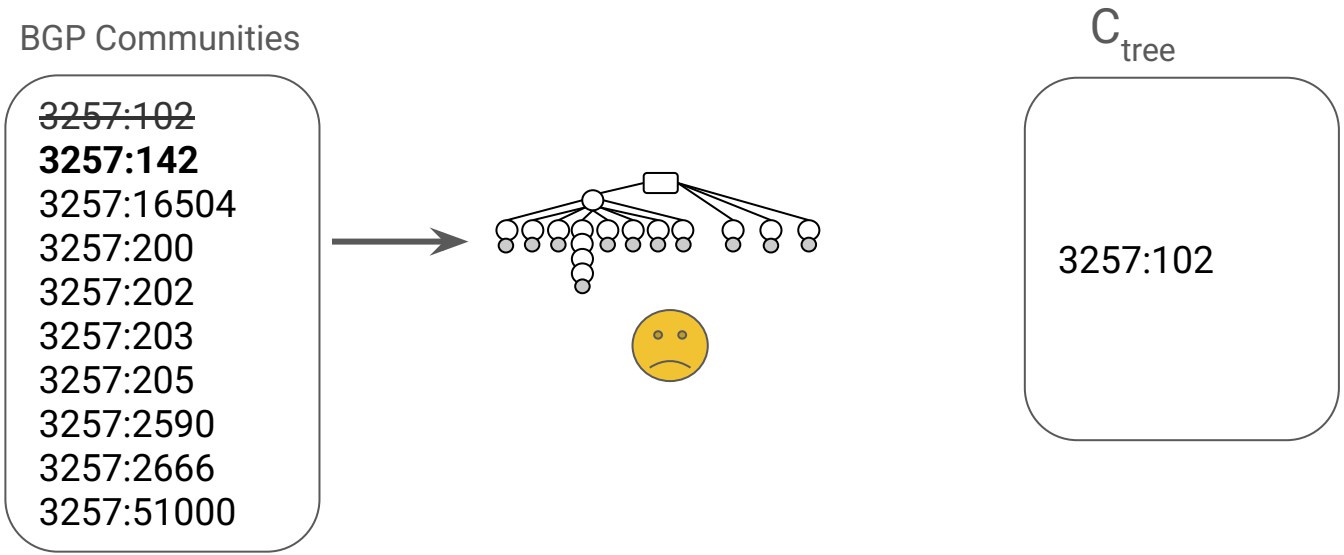


C_{tree}

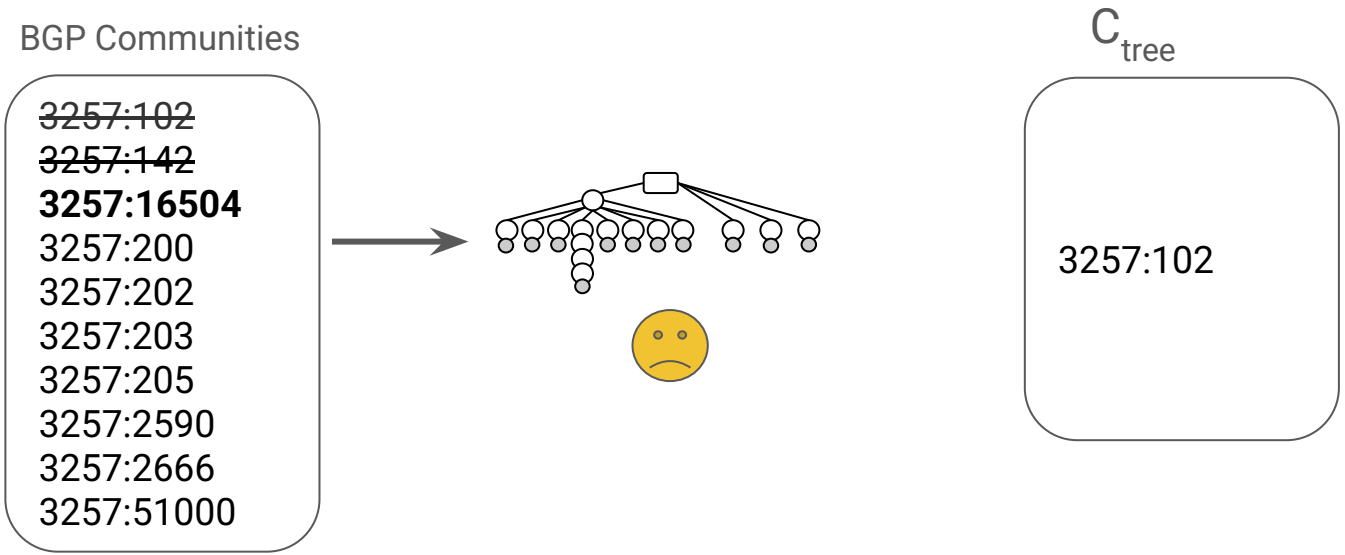
Prefix-Tree in Action



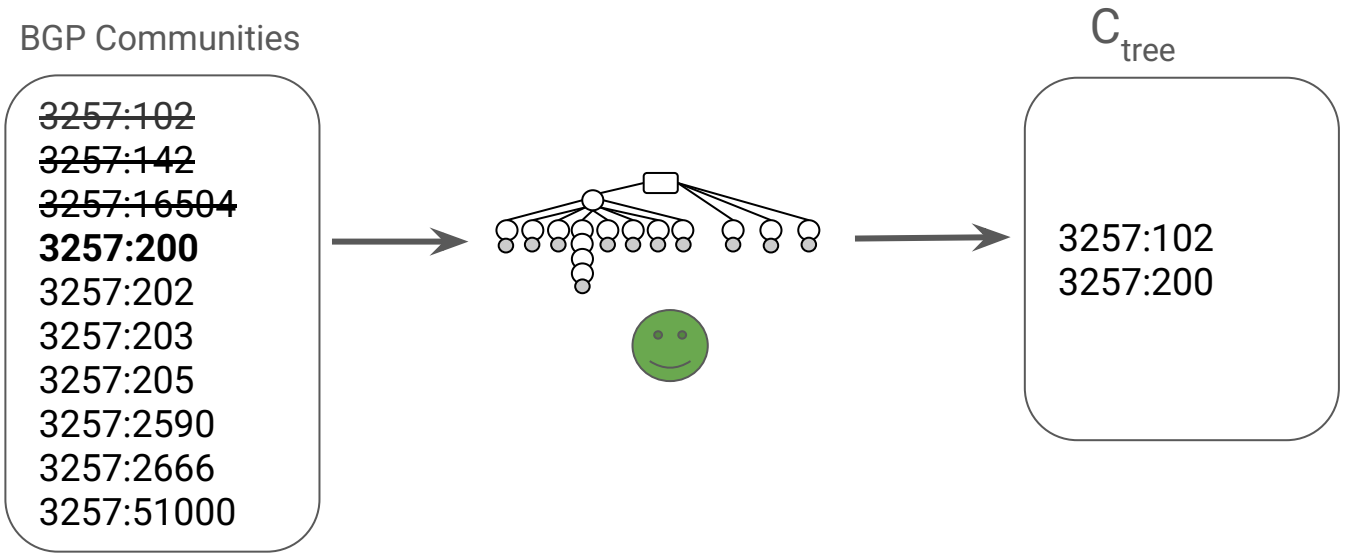
Prefix-Tree in Action



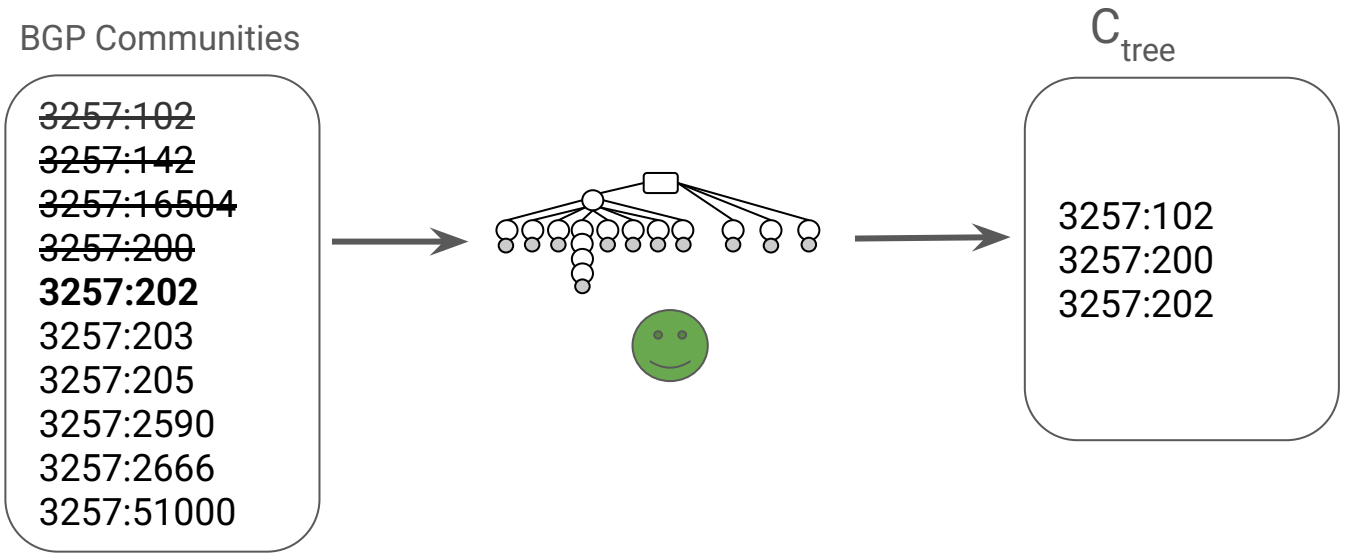
Prefix-Tree in Action



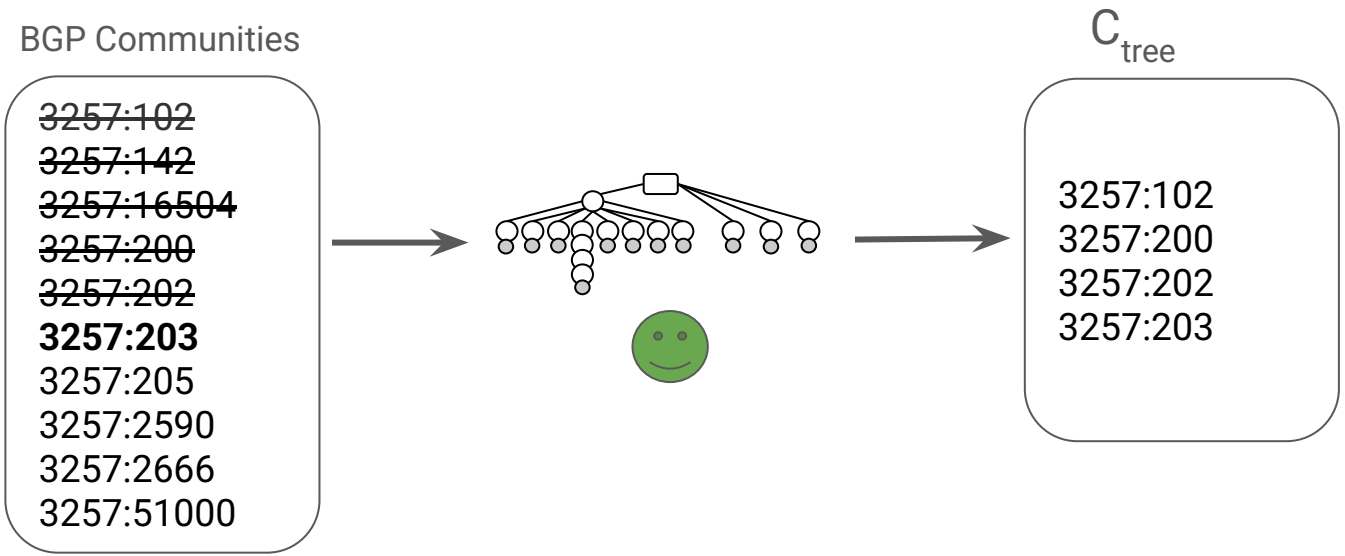
Prefix-Tree in Action



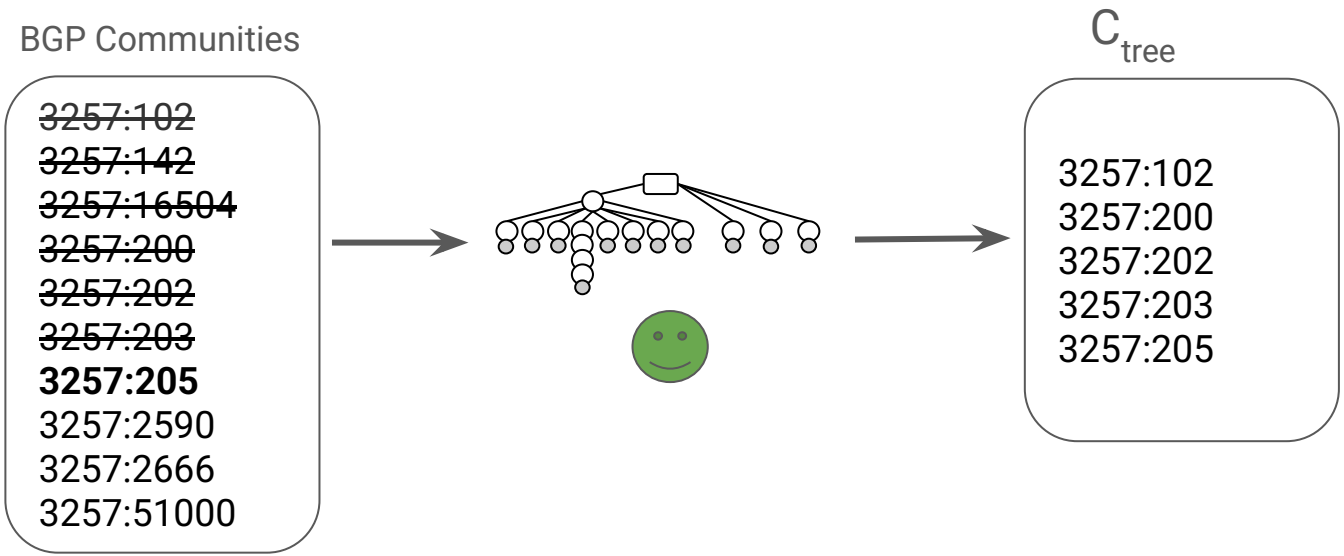
Prefix-Tree in Action



Prefix-Tree in Action



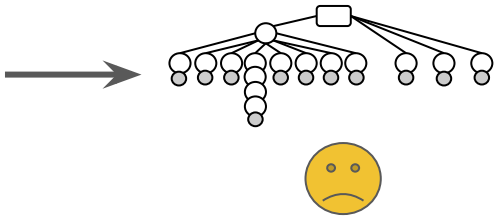
Prefix-Tree in Action



Prefix-Tree in Action

BGP Communities

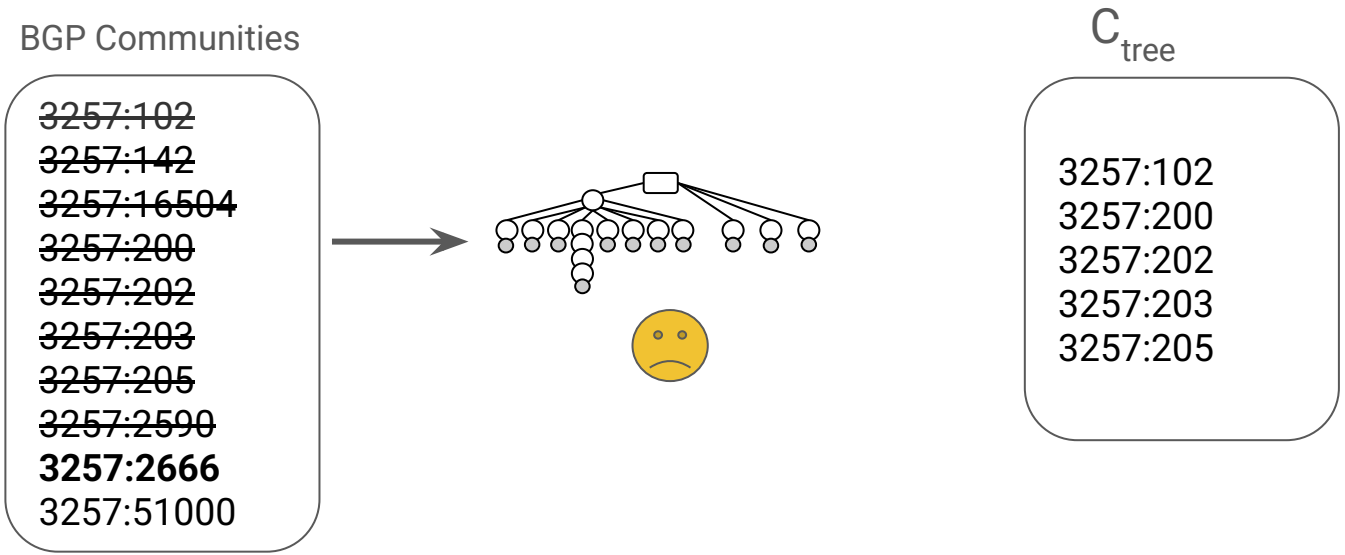
~~3257:102~~
~~3257:142~~
~~3257:16504~~
~~3257:200~~
~~3257:202~~
~~3257:203~~
~~3257:205~~
3257:2590
3257:2666
3257:51000



C_{tree}

3257:102
3257:200
3257:202
3257:203
3257:205

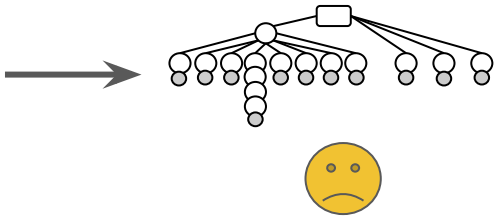
Prefix-Tree in Action



Prefix-Tree in Action

BGP Communities

~~3257:102~~
~~3257:142~~
~~3257:16504~~
~~3257:200~~
~~3257:202~~
~~3257:203~~
~~3257:205~~
~~3257:2590~~
~~3257:2666~~
3257:51000



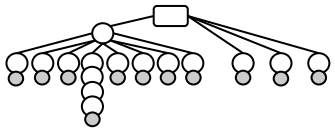
C_{tree}

3257:102
3257:200
3257:202
3257:203
3257:205

Prefix-Tree in Action

BGP Communities

~~3257:102~~
~~3257:142~~
~~3257:16504~~
~~3257:200~~
~~3257:202~~
~~3257:203~~
~~3257:205~~
~~3257:2590~~
~~3257:2666~~
~~3257:51000~~



Finished

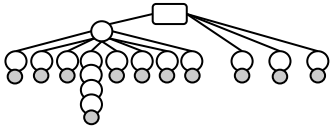
C_{tree}

3257:102
3257:200
3257:202
3257:203
3257:205

Prefix-Tree in Action

BGP Communities

~~3257:102~~
~~3257:142~~
~~3257:16504~~
~~3257:200~~
~~3257:202~~
~~3257:203~~
~~3257:205~~
~~3257:2590~~
~~3257:2666~~
~~3257:51000~~



C_{tree}

3257:102
3257:200
3257:202
3257:203
3257:205

an extra community
directly from the dumps.

Results and Dataset

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend			
ASN	Num	Prec	Rec	Unk
1299	131	0.98	0.43	36
2914	61	1.0	0.93	22
3257	36	0.88	0.38	19
3491	62	0.94	0.25	12
3549	34	1.0	0.57	22
5511	18	0.86	0.35	4
6461	43	1.0	0.64	4
6663	2	1.0	1.0	1

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend			
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36
2914	61	1.0	0.93	22	61	1.0	0.93	22
3257	36	0.88	0.38	19	36	0.88	0.38	19
3491	62	0.94	0.25	12	69	0.95	0.28	13
3549	34	1.0	0.57	22	34	1.0	0.57	22
5511	18	0.86	0.35	4	18	0.86	0.35	4
6461	43	1.0	0.64	4	43	1.0	0.64	4
6663	2	1.0	1.0	1	6	1.0	1.0	5

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

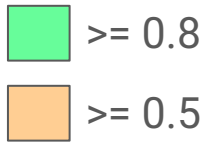
	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree			
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree				GT ∩ BGP		
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Act	Info	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80	218	98	138
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23	42	81	30
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23	39	844	26
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50	16	139	60
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31	21	22	109
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7	35	51	253
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4	61	289	54
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5	1	0	23



Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree				GT ∩ BGP		
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Act	Info	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80	218	98	138
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23	42	81	30
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23	39	844	26
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50	16	139	60
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31	21	22	109
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7	35	51	253
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4	61	289	54
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5	1	0	23

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree				GT ∩ BGP		
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Act	Info	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80	218	98	138
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23	42	81	30
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23	39	844	26
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50	16	139	60
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31	21	22	109
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7	35	51	253
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4	61	289	54
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5	1	0	23

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree				GT ∩ BGP		
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Act	Info	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80	218	98	138
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23	42	81	30
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23	39	844	26
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50	16	139	60
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31	21	22	109
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7	35	51	253
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4	61	289	54
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5	1	0	23

>= 0.8

>= 0.5

Inference for Selected ASes Tier 1 and 2

	Inf. without Prepend				Inf. with Prepend				Inf. Prepend with Tree				GT ∩ BGP		
ASN	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Num	Prec	Rec	Unk	Act	Info	Unk
1299	131	0.98	0.43	36	131	0.98	0.43	36	340	0.84	1.0	80	218	98	138
2914	61	1.0	0.93	22	61	1.0	0.93	22	67	0.95	1.0	23	42	81	30
3257	36	0.88	0.38	19	36	0.88	0.38	19	61	0.87	0.85	23	39	844	26
3491	62	0.94	0.25	12	69	0.95	0.28	13	252	0.94	0.99	50	16	139	60
3549	34	1.0	0.57	22	34	1.0	0.57	22	49	1.0	0.86	31	21	22	109
5511	18	0.86	0.35	4	18	0.86	0.35	4	39	0.94	0.88	7	35	51	253
6461	43	1.0	0.64	4	43	1.0	0.64	4	63	0.95	0.92	4	61	289	54
6663	2	1.0	1.0	1	6	1.0	1.0	5	6	1.0	1.0	5	1	0	23

>= 0.8

>= 0.5

Prefix-Tree vs Previous Work

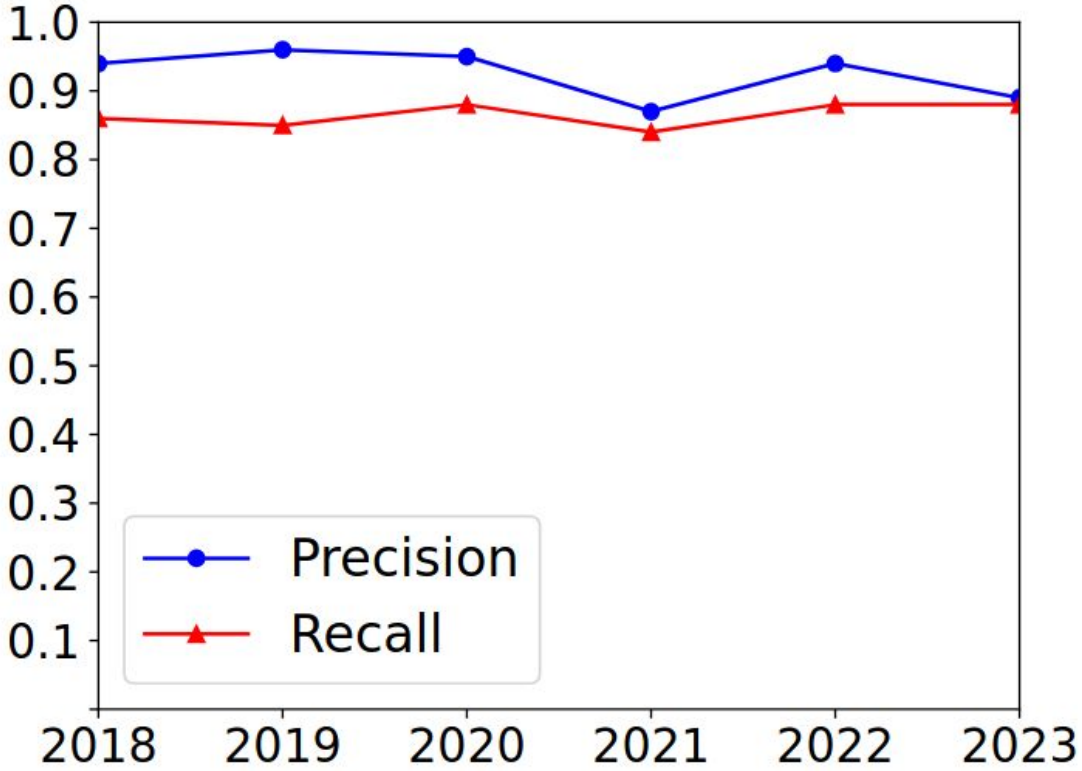
- Observing our strategy versus the previous work, we were able to have better results on Level3 (AS3549) and Next Layer Telecom (AS1764) because it intermix action and information communities violating the fixed size of blocks for community semantics.

Prefix-Tree vs Previous Work

- Observing our strategy versus the previous work, we were able to have better results on Level3 (AS3549) and Next Layer Telecom (AS1764) because it intermix action and information communities violating the fixed size of blocks for community semantics.

- For GTT (AS3257), our approach have better results because we compute the AS squatters (e.g., siblings), removing the noise from the relationship in the semantics inference.

Inference of Action Communities (Longitudinal)



Conclusion

Conclusion

- We explored the current state of the use of the BGP communities, showing how these communities significantly impact Internet operations.

Conclusion

- We explored the current state of the use of the BGP communities, showing how these communities significantly impact Internet operations.

- The results our work enable researchers and network operators to use our inferred datasets for multiple purposes, such as:
  - Identifying anomalies in traffic patterns;
  - Identifying hidden relationships between ASes;
  - Detecting traffic engineering from route announcements;
  - Improving the understanding of network dynamics.

Conclusion

- Our algorithms performed well in identifying action communities, achieving an average precision of 92.5% and an average recall of 86.5% from 2018 to 2023 for major Internet providers (Tier-1 and Tier-2 ASes).

Conclusion

- Our algorithms performed well in identifying action communities, achieving an average precision of 92.5% and an average recall of 86.5% from 2018 to 2023 for major Internet providers (Tier-1 and Tier-2 ASes).

- Additionally, our algorithm to uncover squatting relationships can complement techniques for validating AS-relationship inferences, tracking route changes, and inferring sibling ASes. Of the 54 pairs of squatting relationships we uncovered, five are sibling relationship not detected by previous works.

Conclusion

- Our algorithms performed well in identifying action communities, achieving an average precision of 92.5% and an average recall of 86.5% from 2018 to 2023 for major Internet providers (Tier-1 and Tier-2 ASes).

- Additionally, our algorithm to uncover squatting relationships can complement techniques for validating AS-relationship inferences, tracking route changes, and inferring sibling ASes. Of the 54 pairs of squatting relationships we uncovered, five are sibling relationship not detected by previous works.

- All our data and source code are public available on Github

Thank you!