Dynamic Intelligent Aggregation Services

Love and Strife in large-scale decentralized systems

Evangelos Pournaras BSc MSc

PhD Candidate

Faculty of Technology, Policy & Management

Systems Engineering Section

PhD Research

Distributed Computing

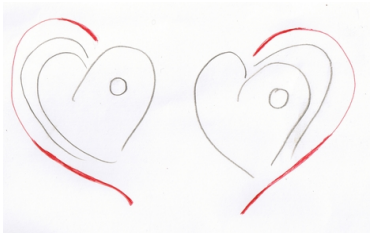
Multi-level Reconfigurable Self-organization
in **Overlay Services**

Dynamic Aggregation Services

Tree Self-organization Services

Demand-side Energy Management

Love and Strife



"These (elements) never cease changing place continually, now being all united by Love into one, now each borne apart by the hatred engendered of Strife, until they are brought together in the unity of the all, and become subject to it."

Continuous Information Change

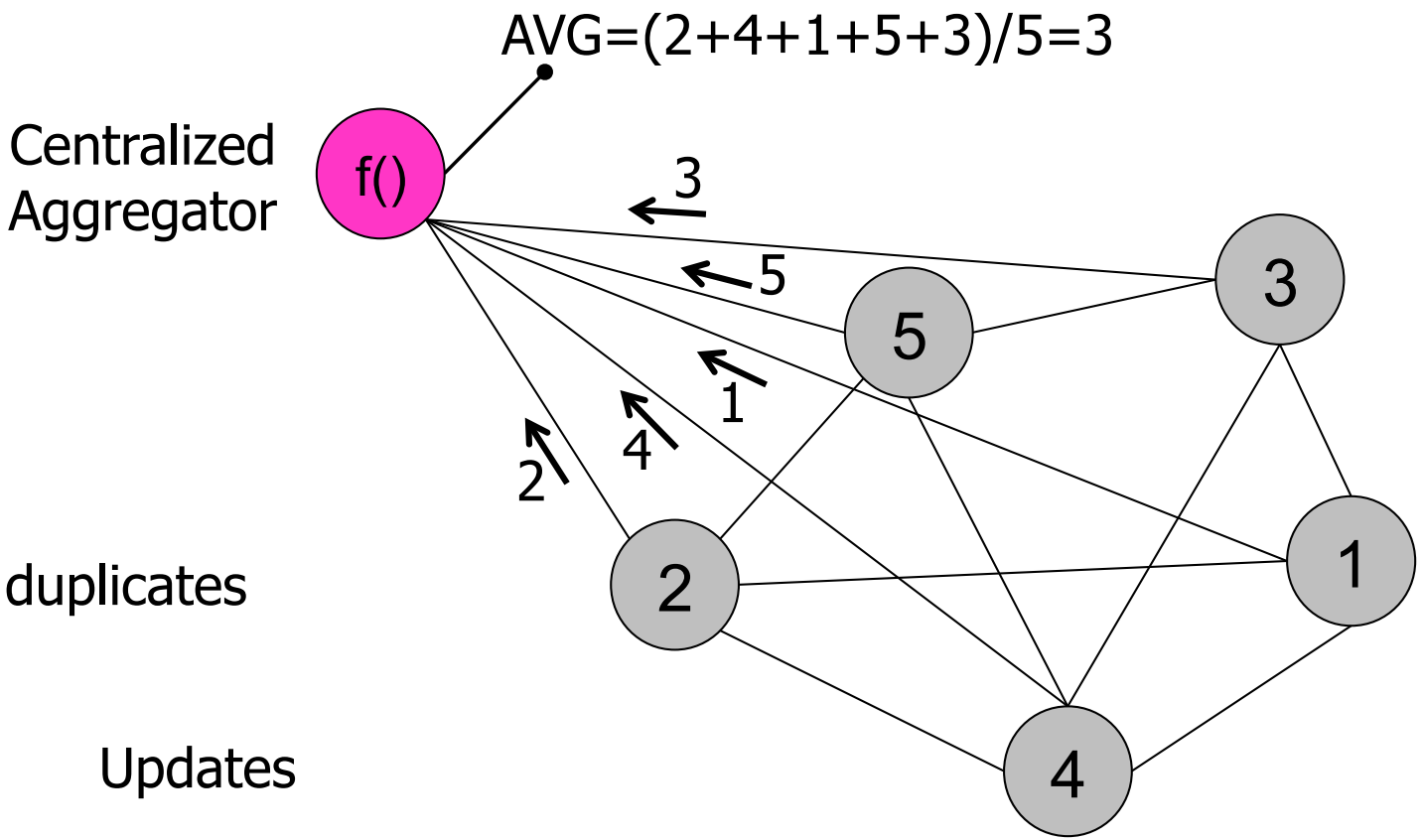
Large-scale Decentralized Systems

Accurate Information Aggregation

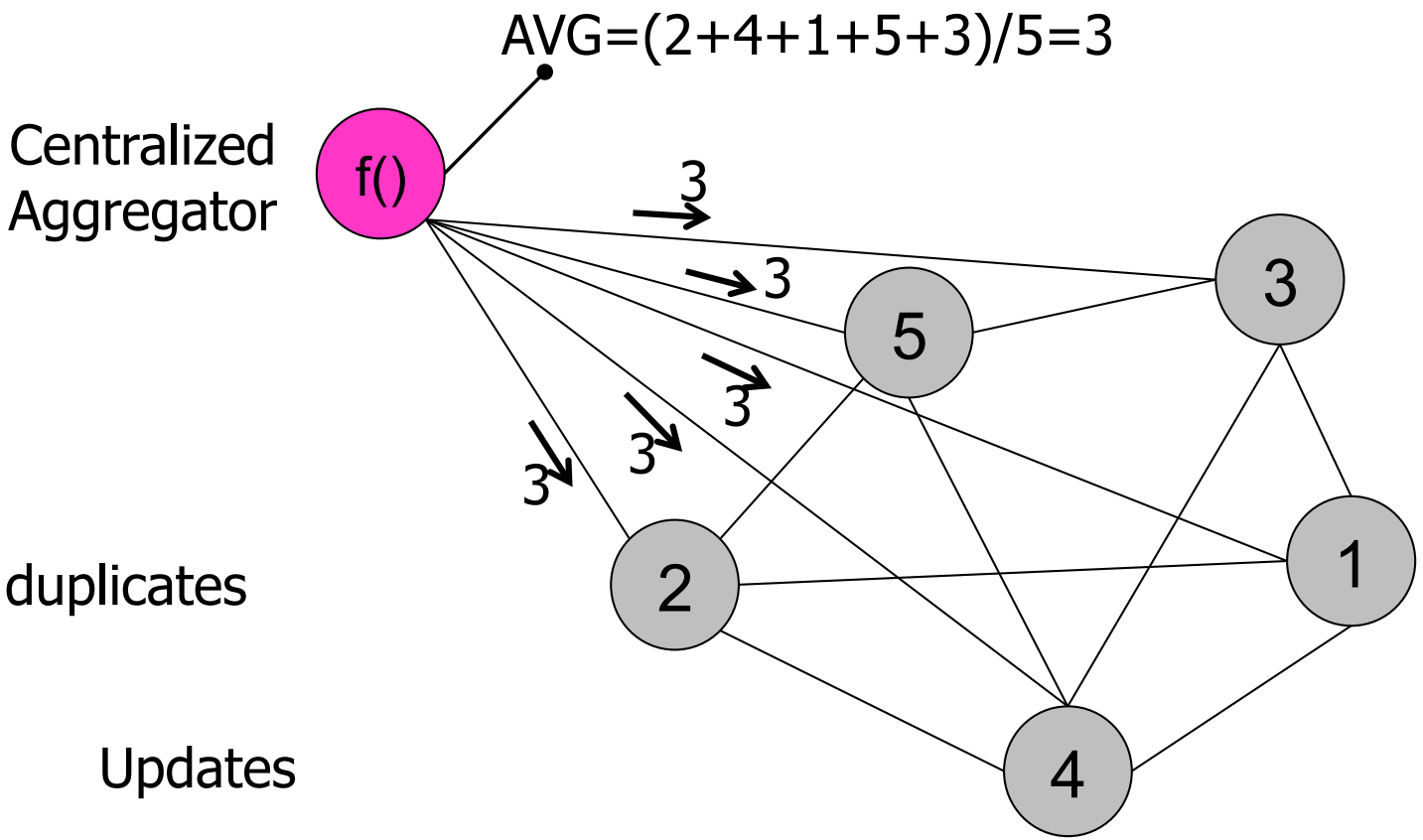
Local

Global

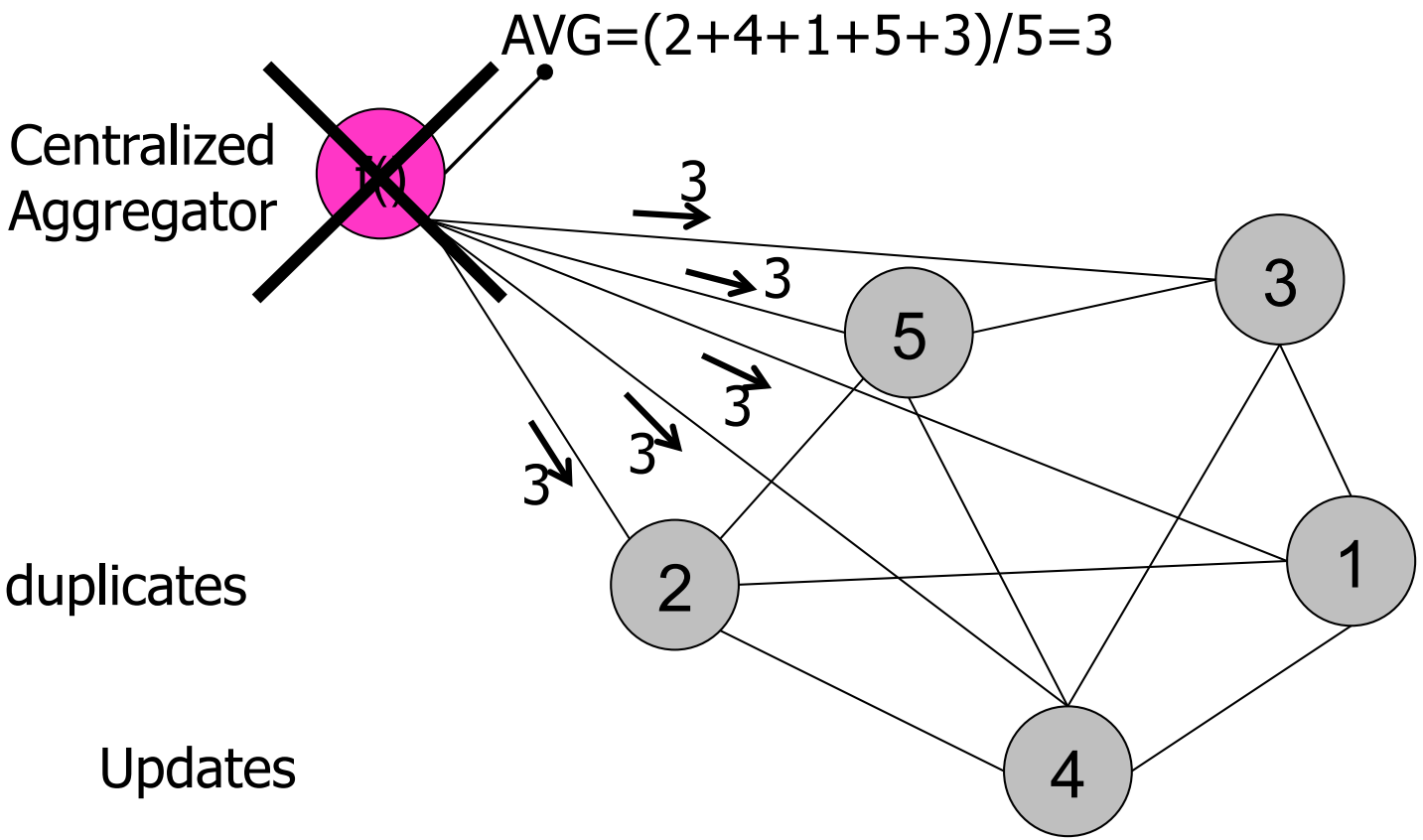
The Aggregation Problem



The Aggregation Problem



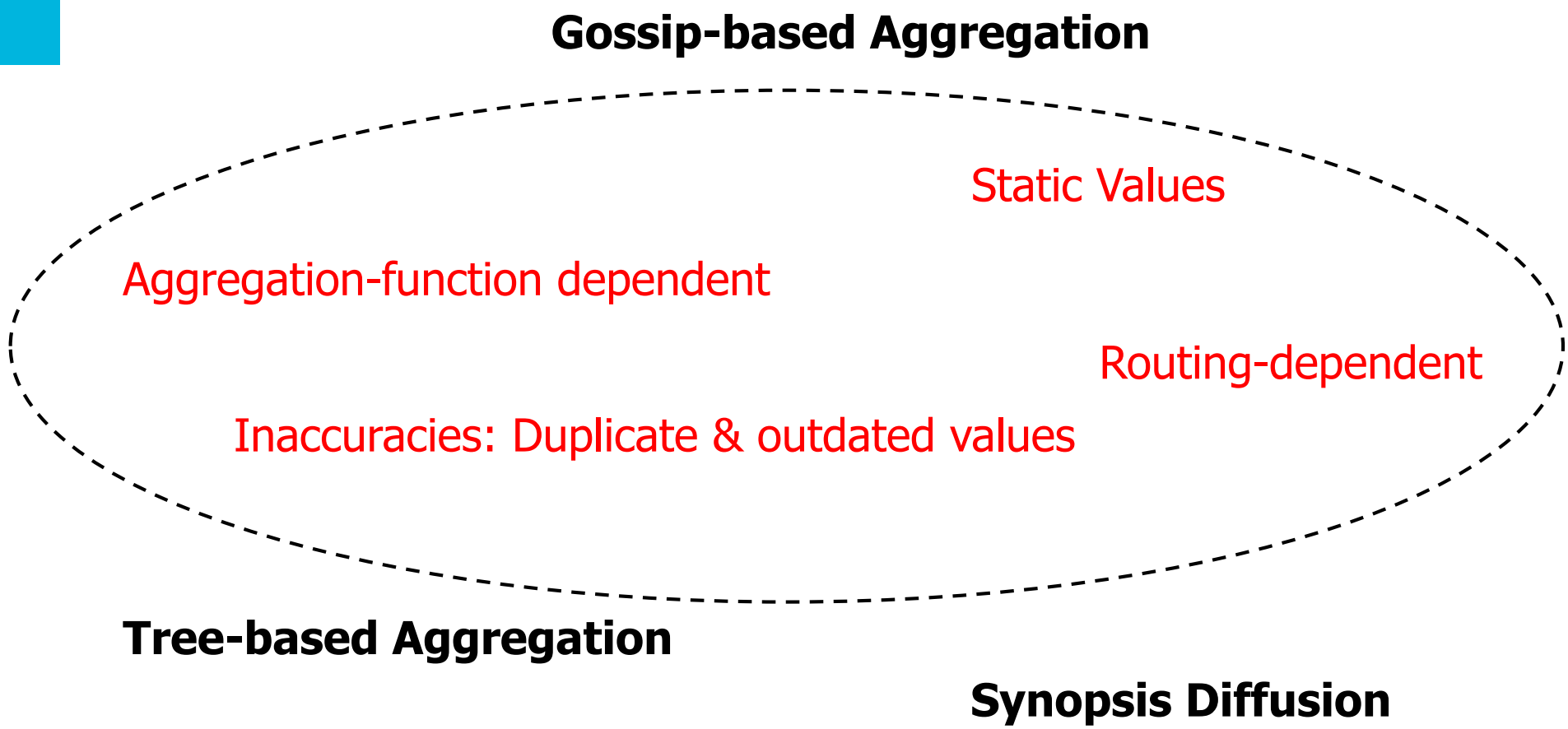
The Aggregation Problem



No duplicates

Updates

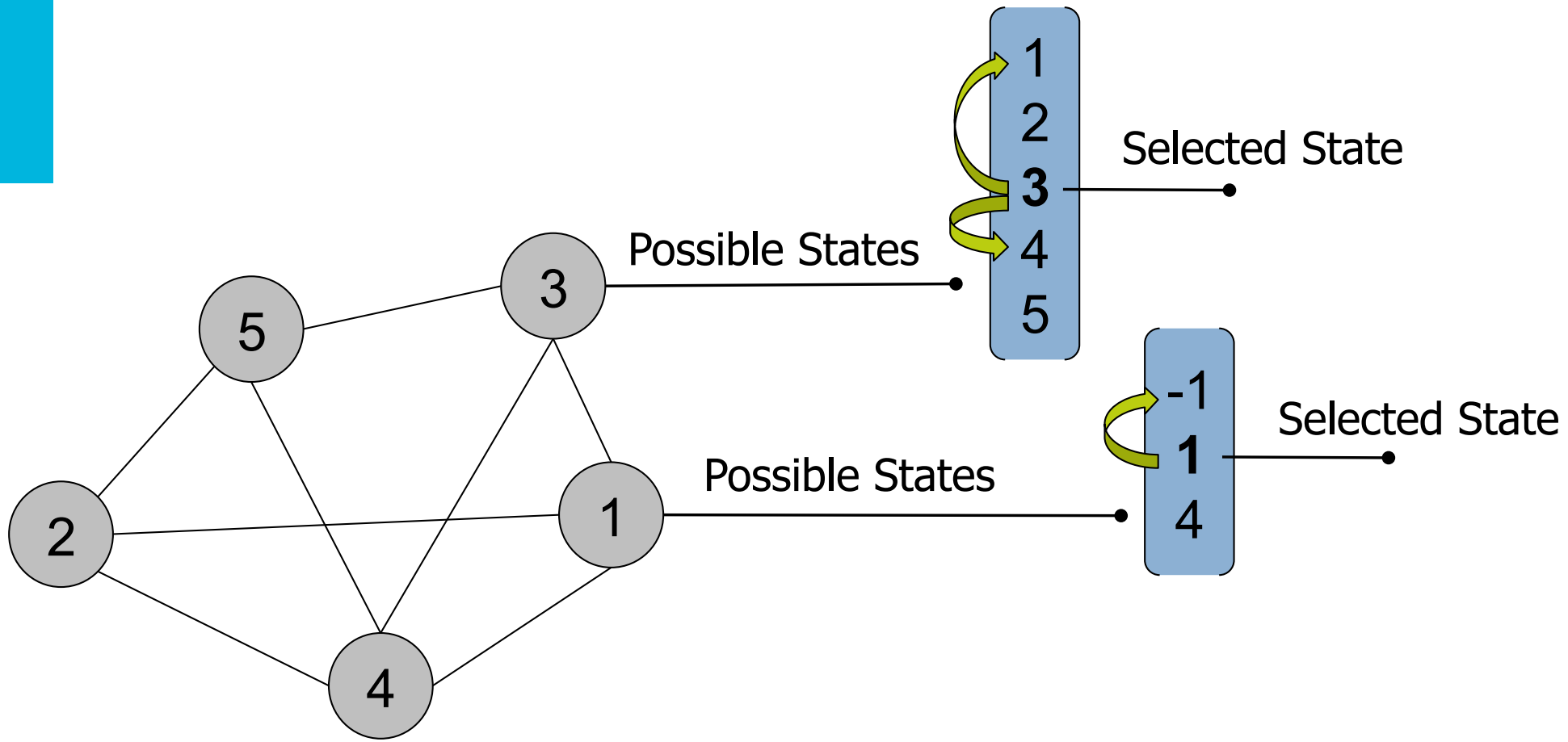
Decentralized Aggregation



Decentralized Aggregation (Cont.)

DIAS – Dynamic Intelligent Aggregation Service

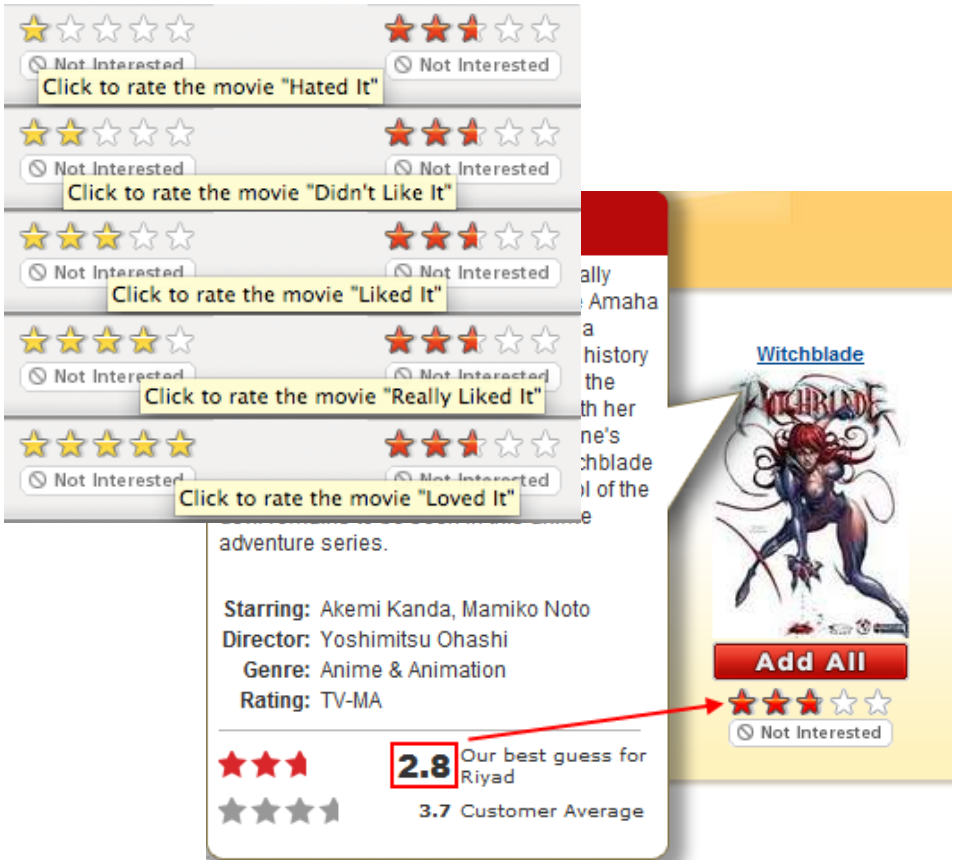
Modeling of Dynamics



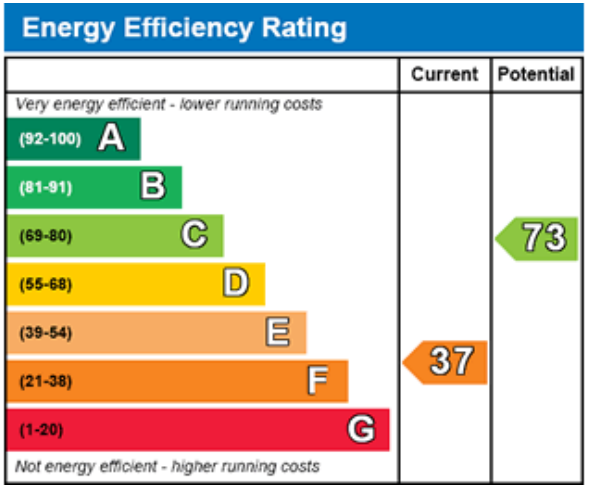
# of Possible Aggregates=# of possible states$^{\#\ \text{of Nodes}}$=3^{10}=59049!

Applications

Recommender Systems

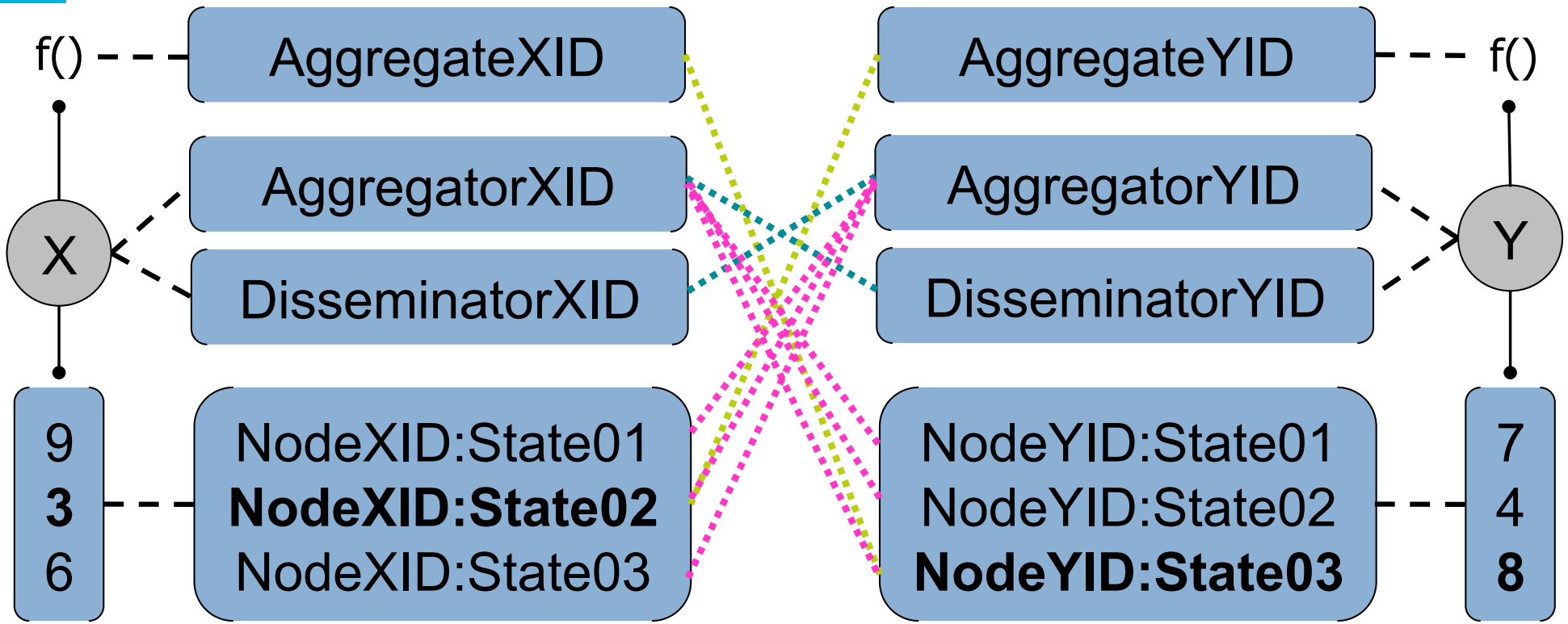


Smart Energy Systems



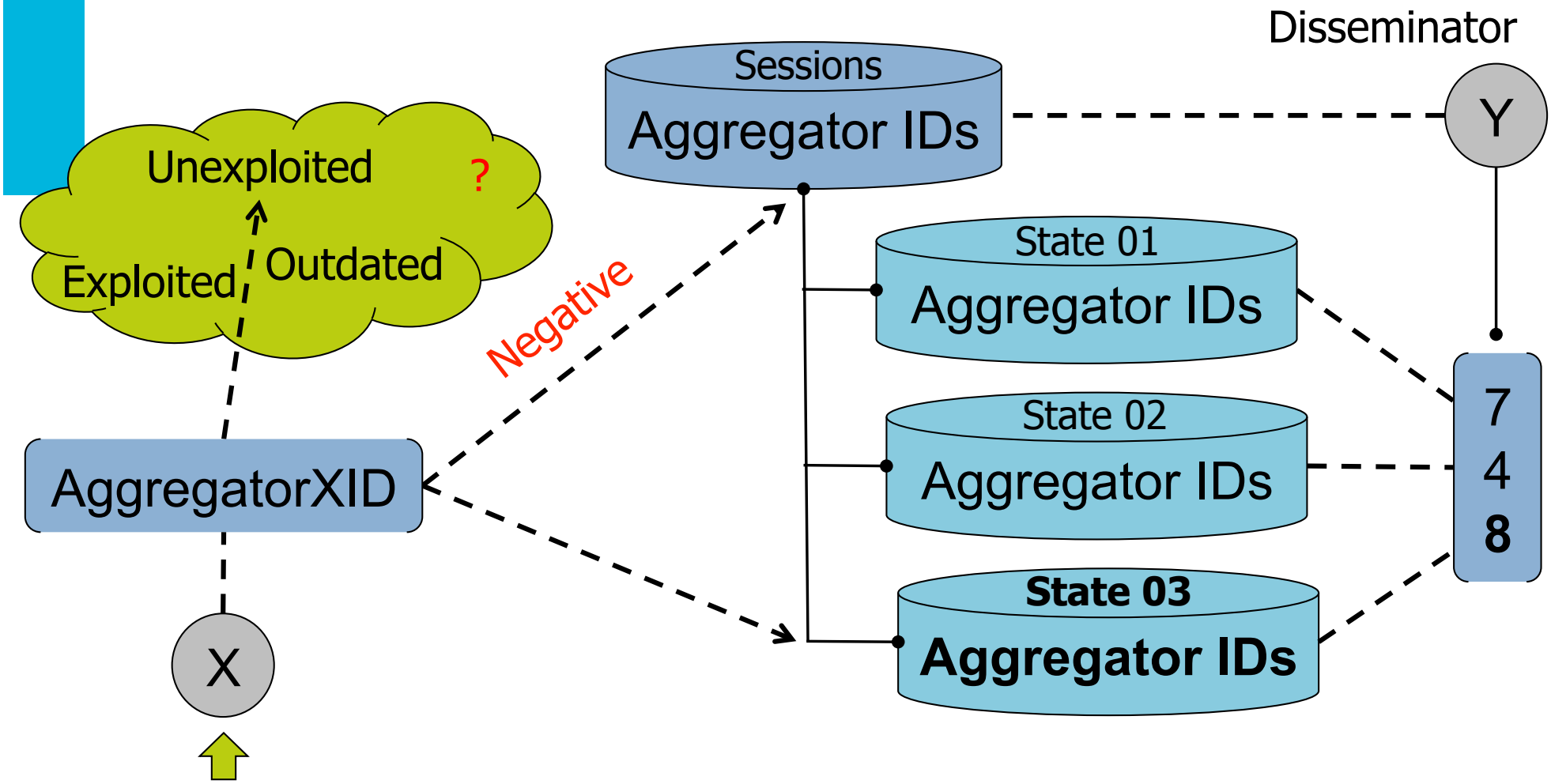
Mutual Aggregation Memberships

Node: Aggregator and Disseminator

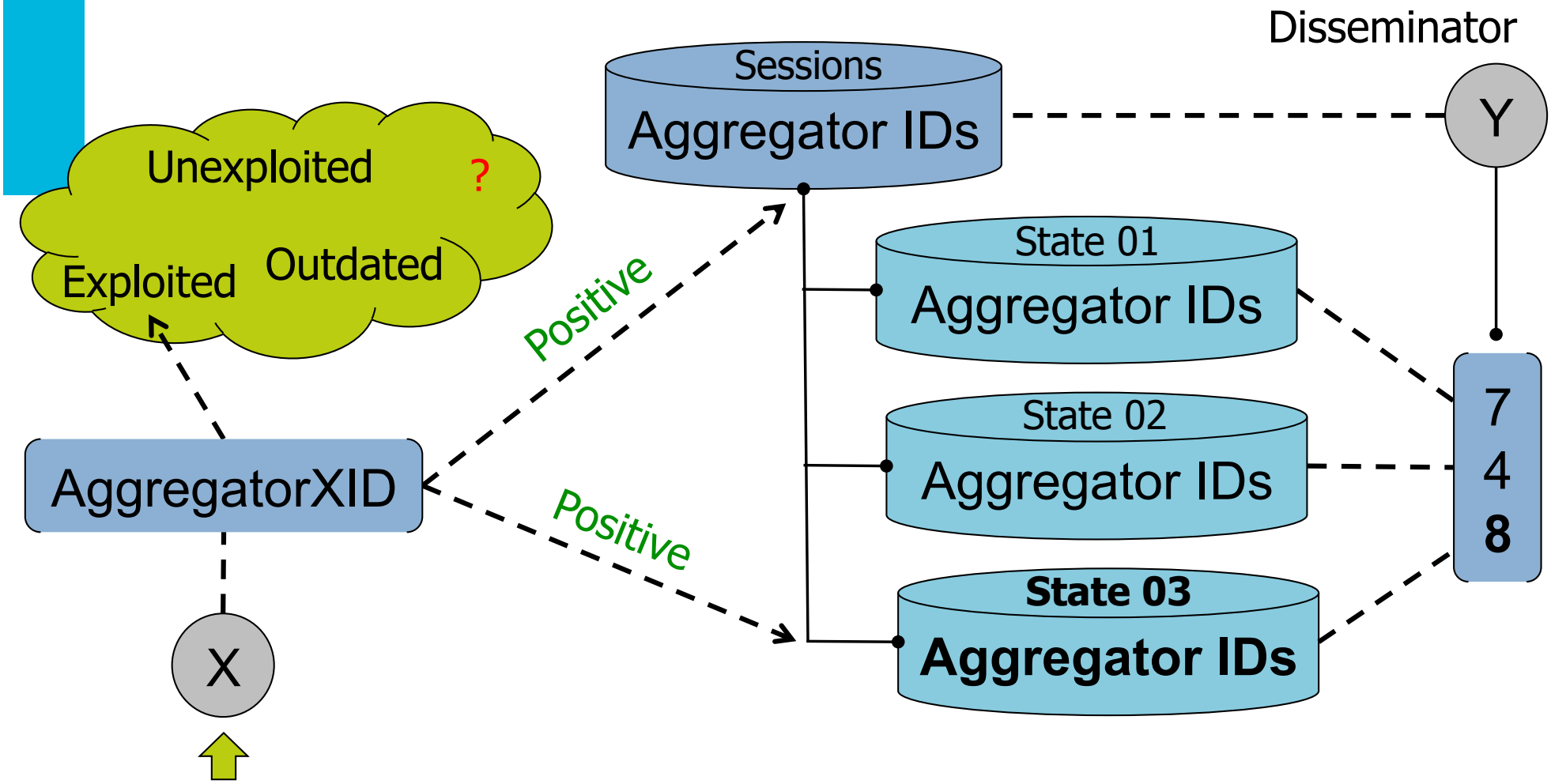


Aggregation Session

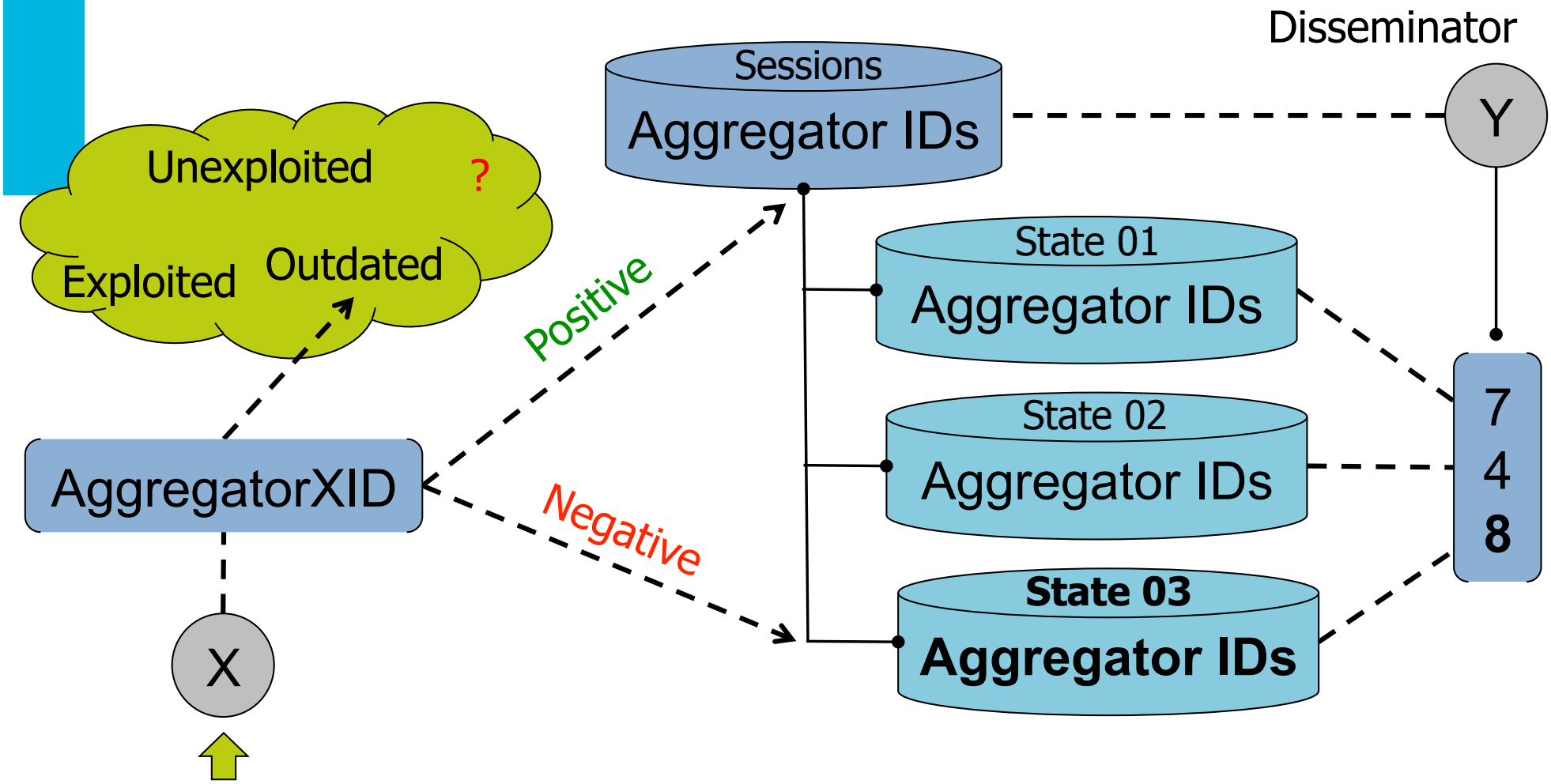
Classification



Classification



Classification



Aggregation Memberships

Can we **explicitly** store aggregation memberships in a decentralized system?

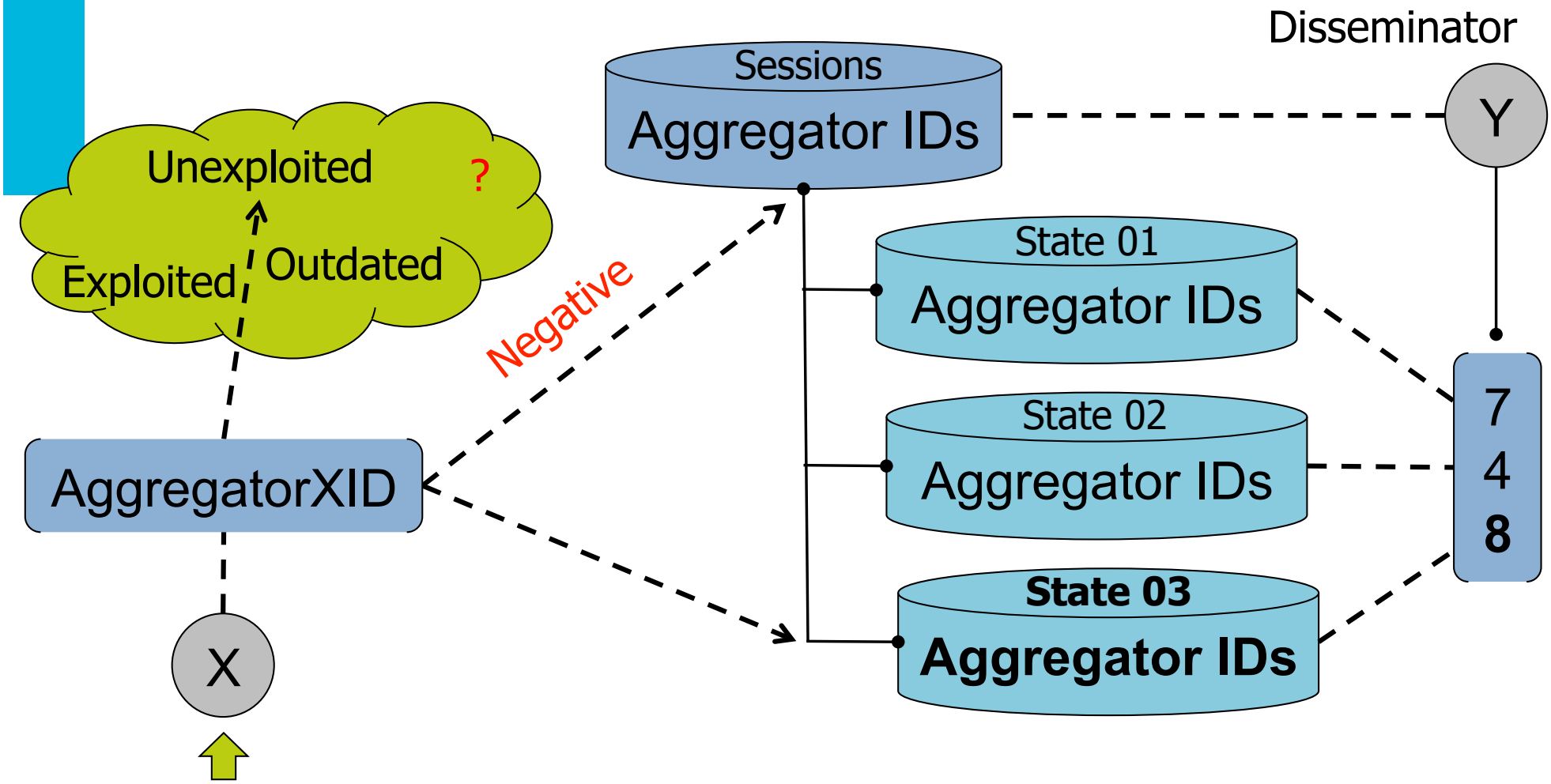
Can we overcome this problem?

Yes, with the Love and Strife of **bloom filters**

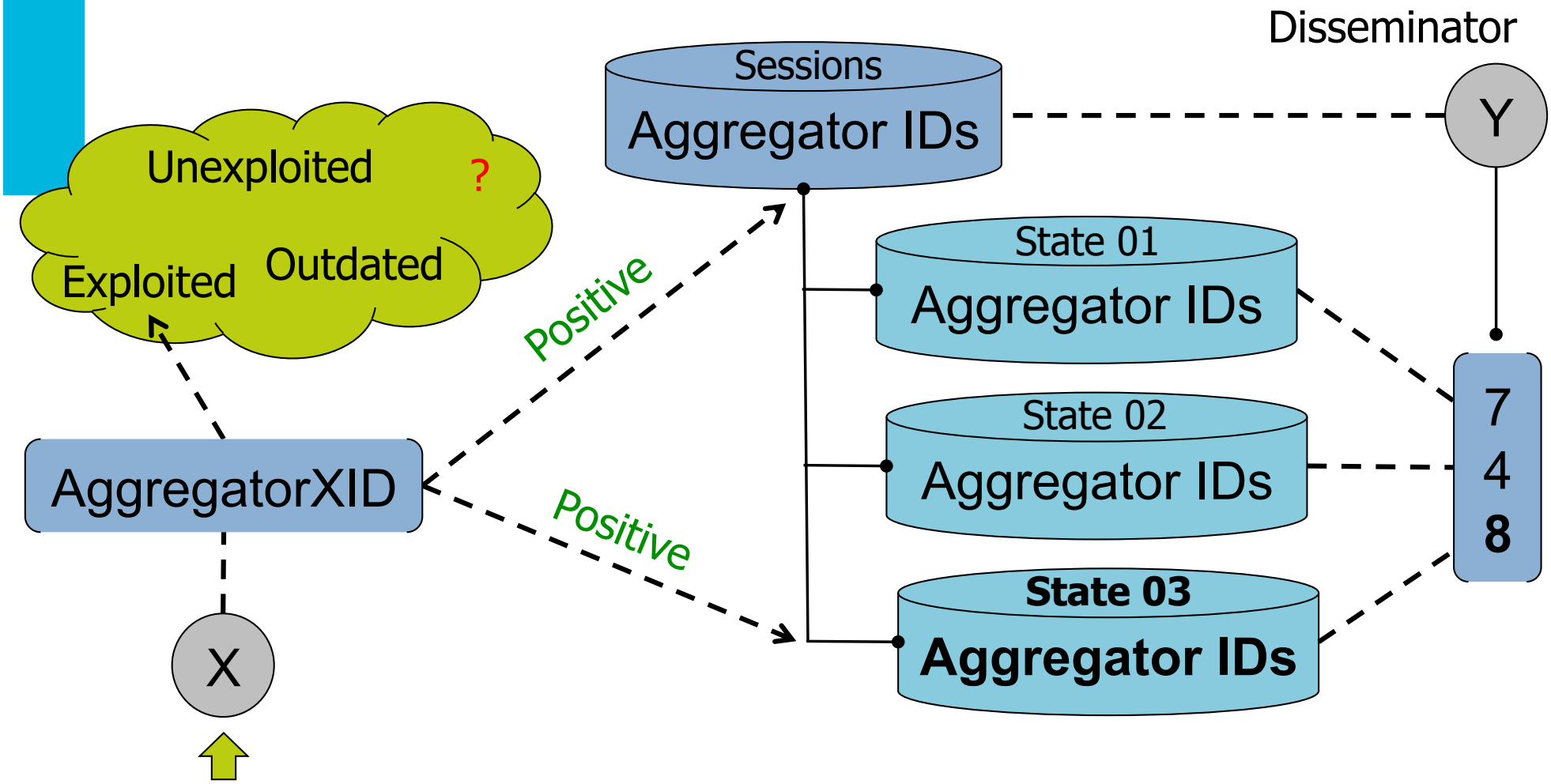
Probabilistic data structure

Large space savings at a cost of false positives

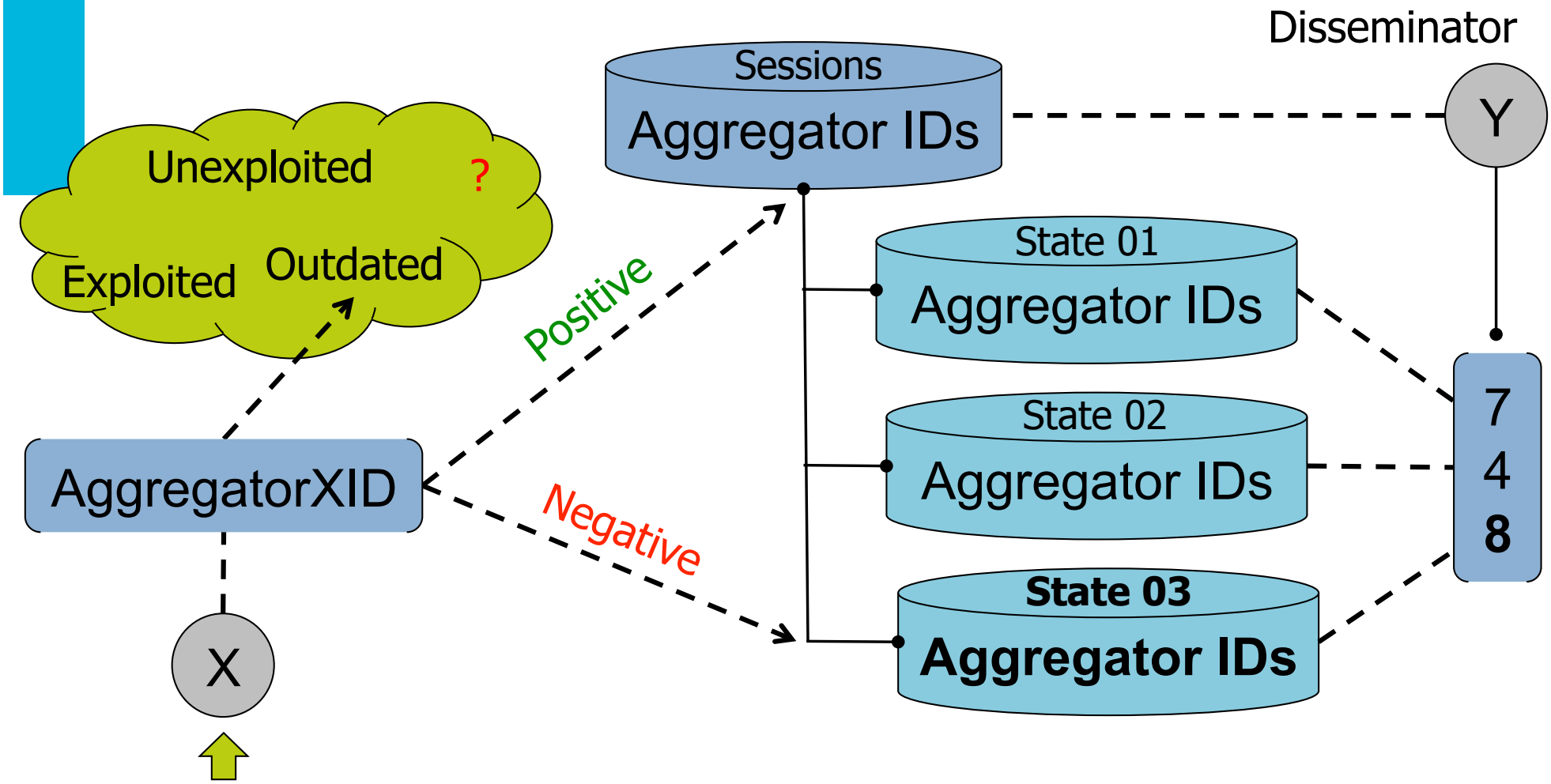
Classification



Classification



Classification



Bloom Filters

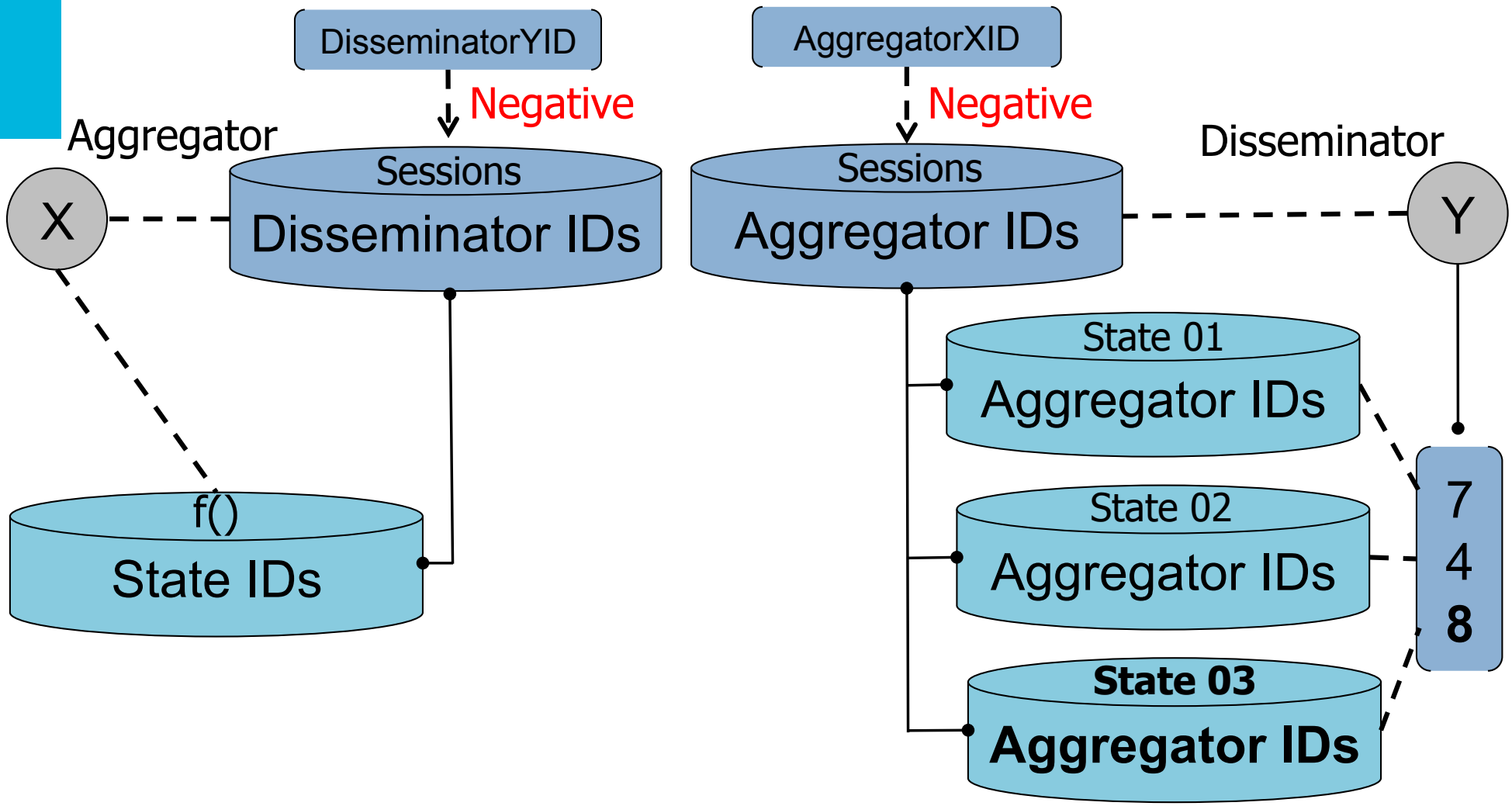
Inconsistent aggregation sessions are possible!

Goal: Minimization of aggregation inaccuracies due to false positives

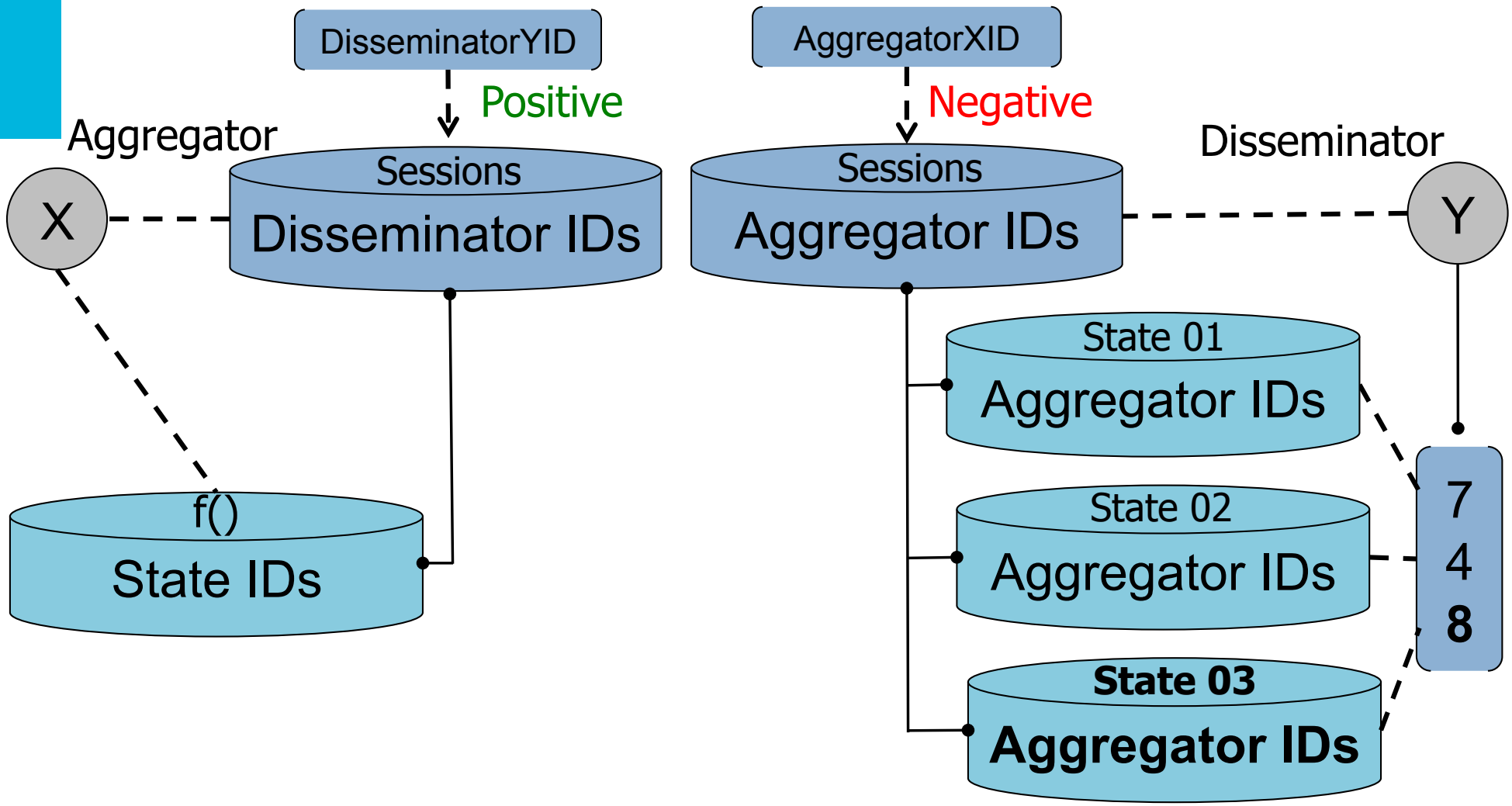
How?

Mutual membership checks!

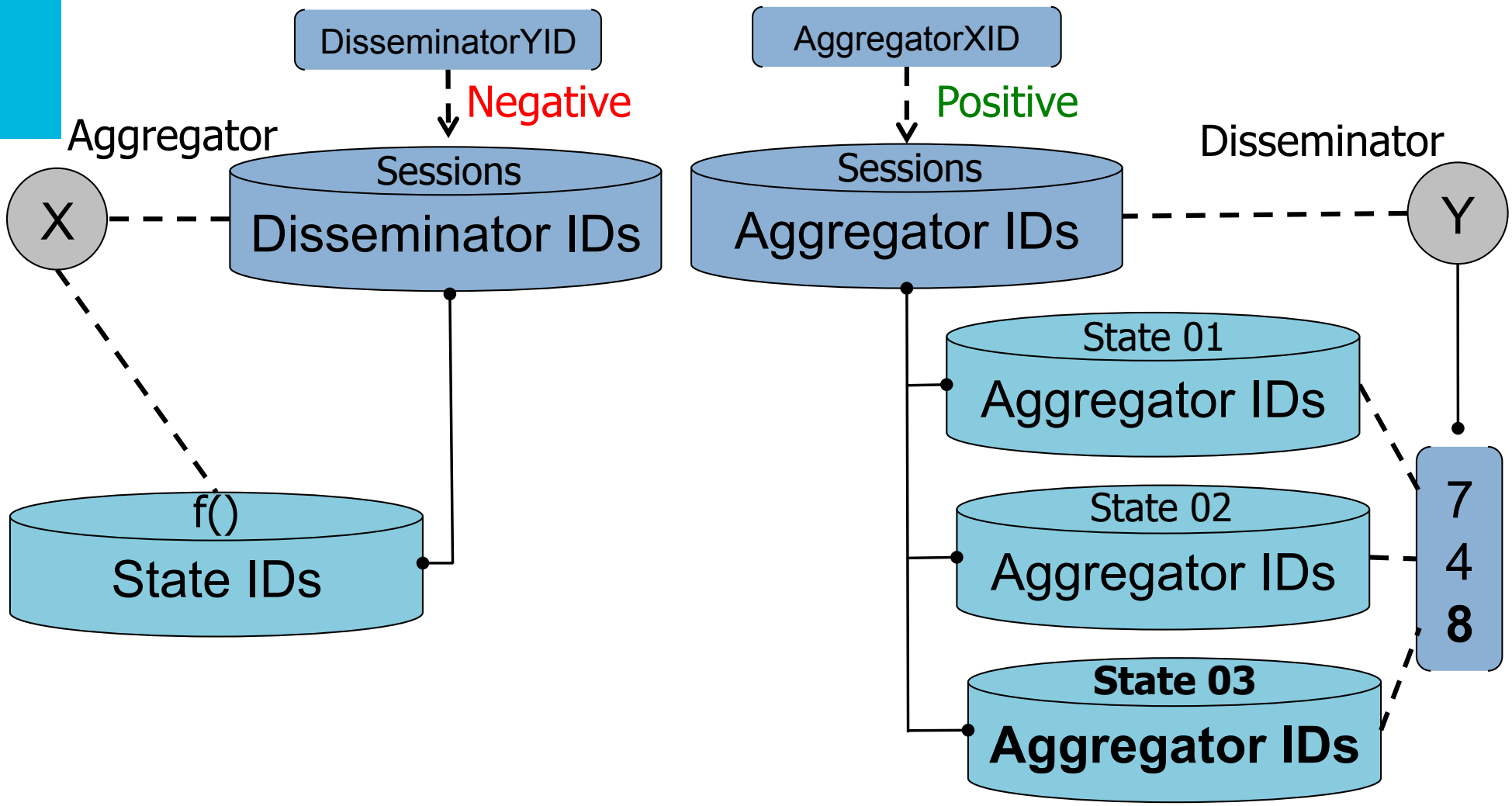
Mutual Membership Checks



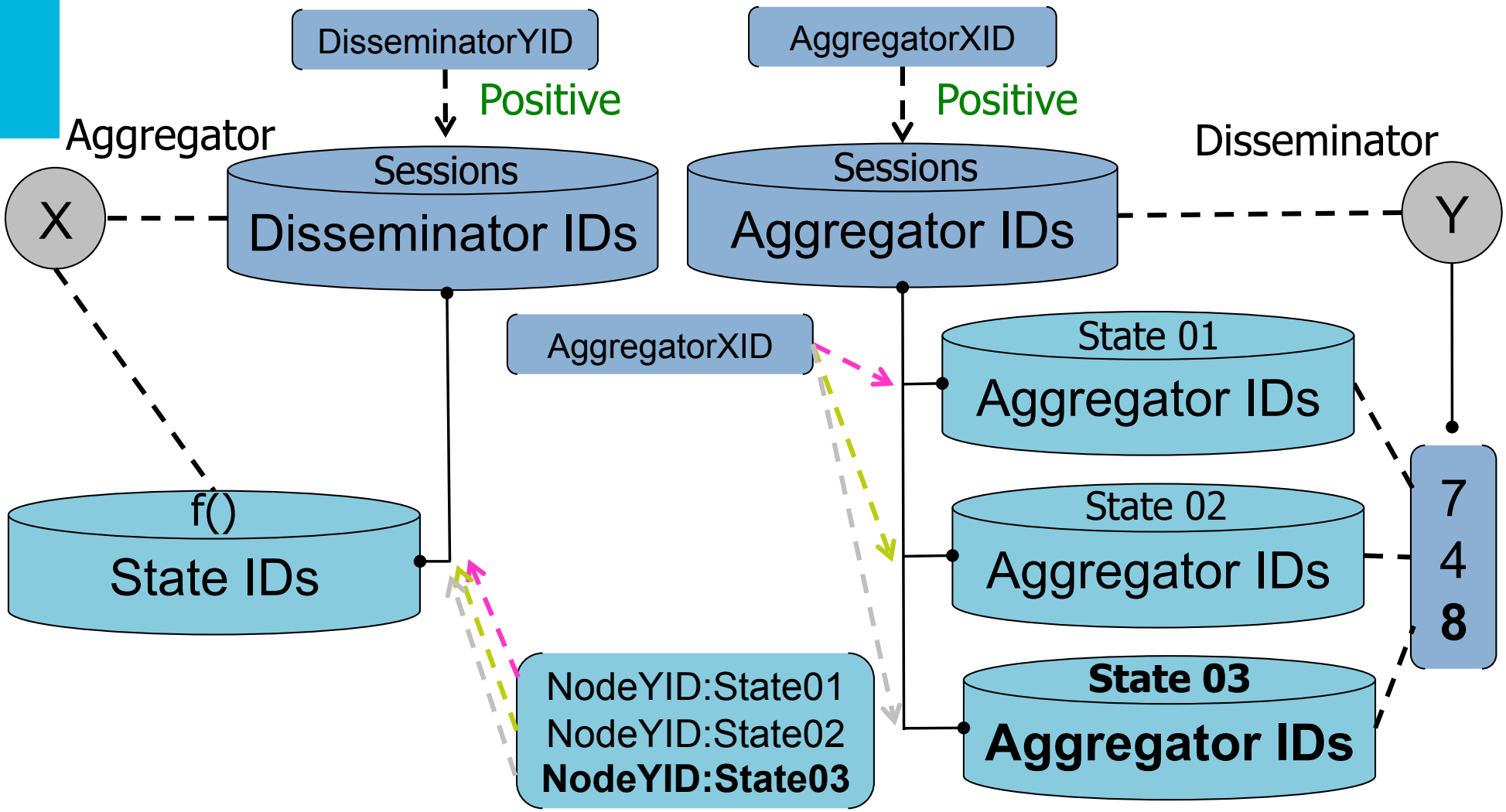
Mutual Membership Checks



Mutual Membership Checks



Mutual Membership Checks



Evaluation

DIAS prototype in Protopeer

XSiena Bloom Filter implementation

Node sampling: Gossiping (Peer Sampling Service)

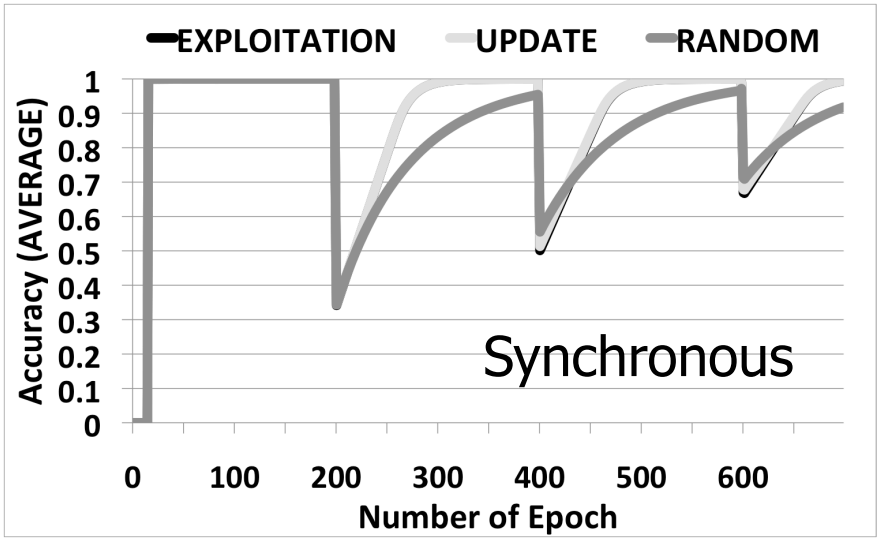
Large-scale network: 1500 nodes

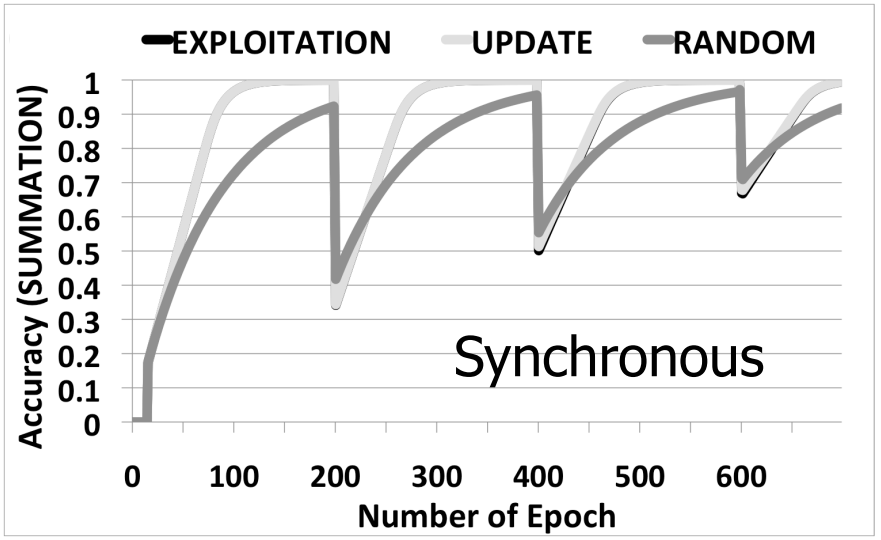
5 possible states/node

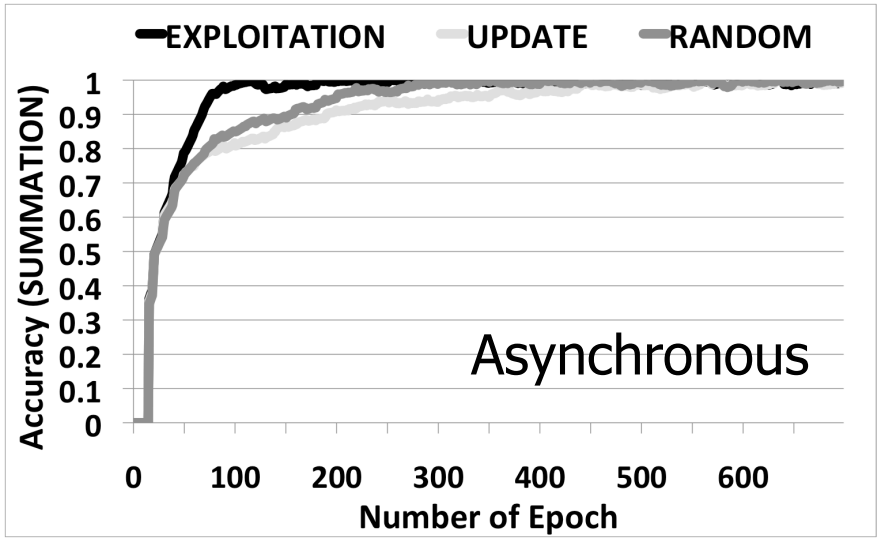
Synchronous/asynchronous state changes

Aggregation strategies: EXPOITATION, UPDATE, RANDOM

Evaluation


EXPLOITATION
UPDATE
RANDOM
Accuracy (AVERAGE)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
600
Number of Epoch
Synchronous


EXPLOITATION
UPDATE
RANDOM
Accuracy (SUMMATION)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
600
Number of Epoch
Synchronous

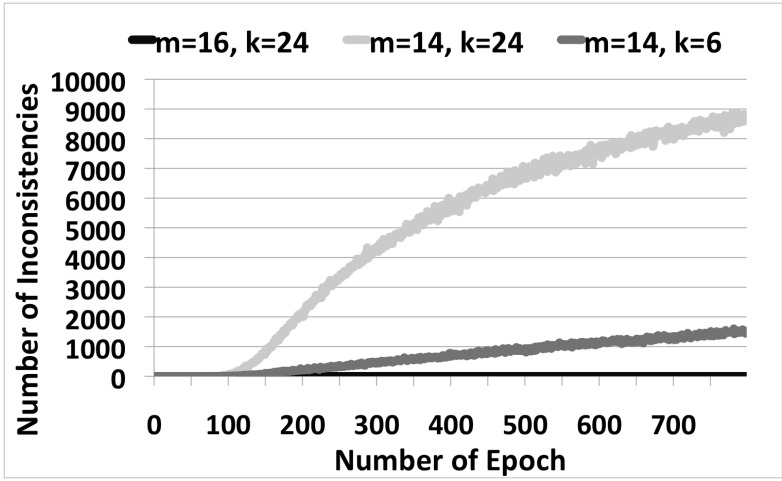

EXPLOITATION
UPDATE
RANDOM
Accuracy (SUMMATION)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
600
Number of Epoch
Asynchronous

Evaluation (cont.)

No influence in accuracy under false positive

Why?

Inconsistencies are detected by **mutual membership checks**



Conclusions

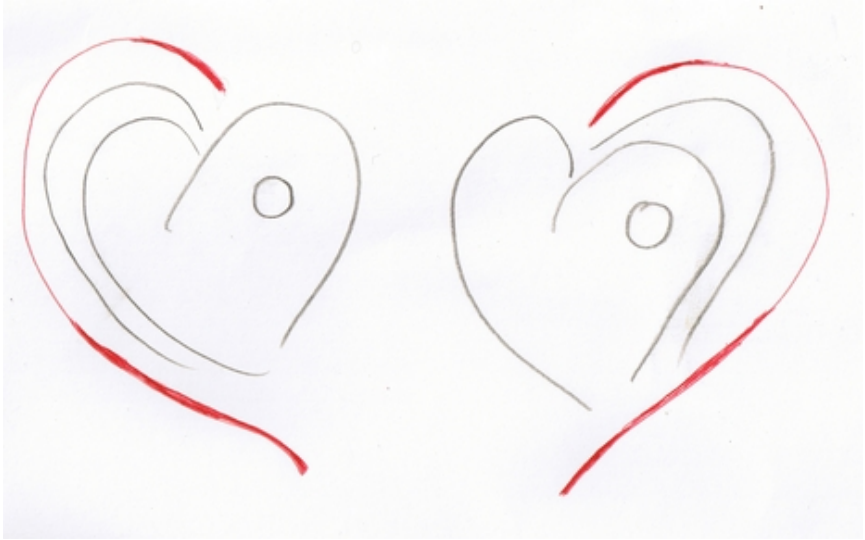
Generic-Multiple aggregation functions

Dynamically changing values

Inaccuracies are minimized: Detection of duplicate and outdated values

Mutual memberships checks: False positive tolerance

Questions?



More information

www.evangelospournaras.com

e.pournaras@tudelft.nl