

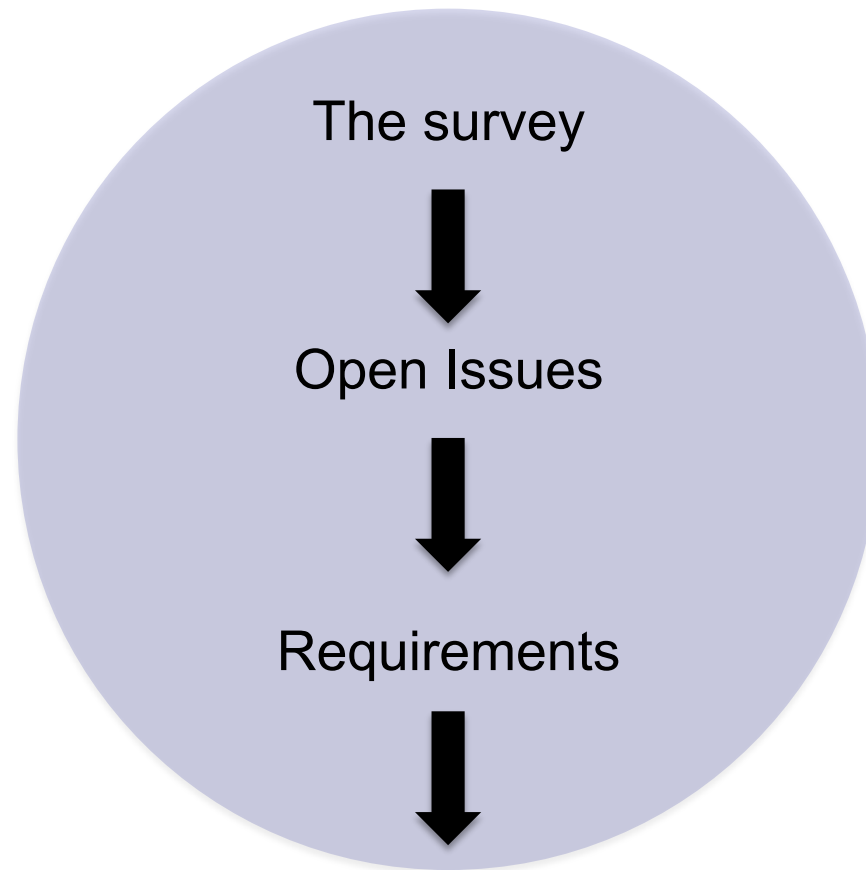
# A Survey and a Proposed Approach on Robust Tree Overlays

**Evangelos Pournaras**

**Vrije Universiteit Amsterdam**

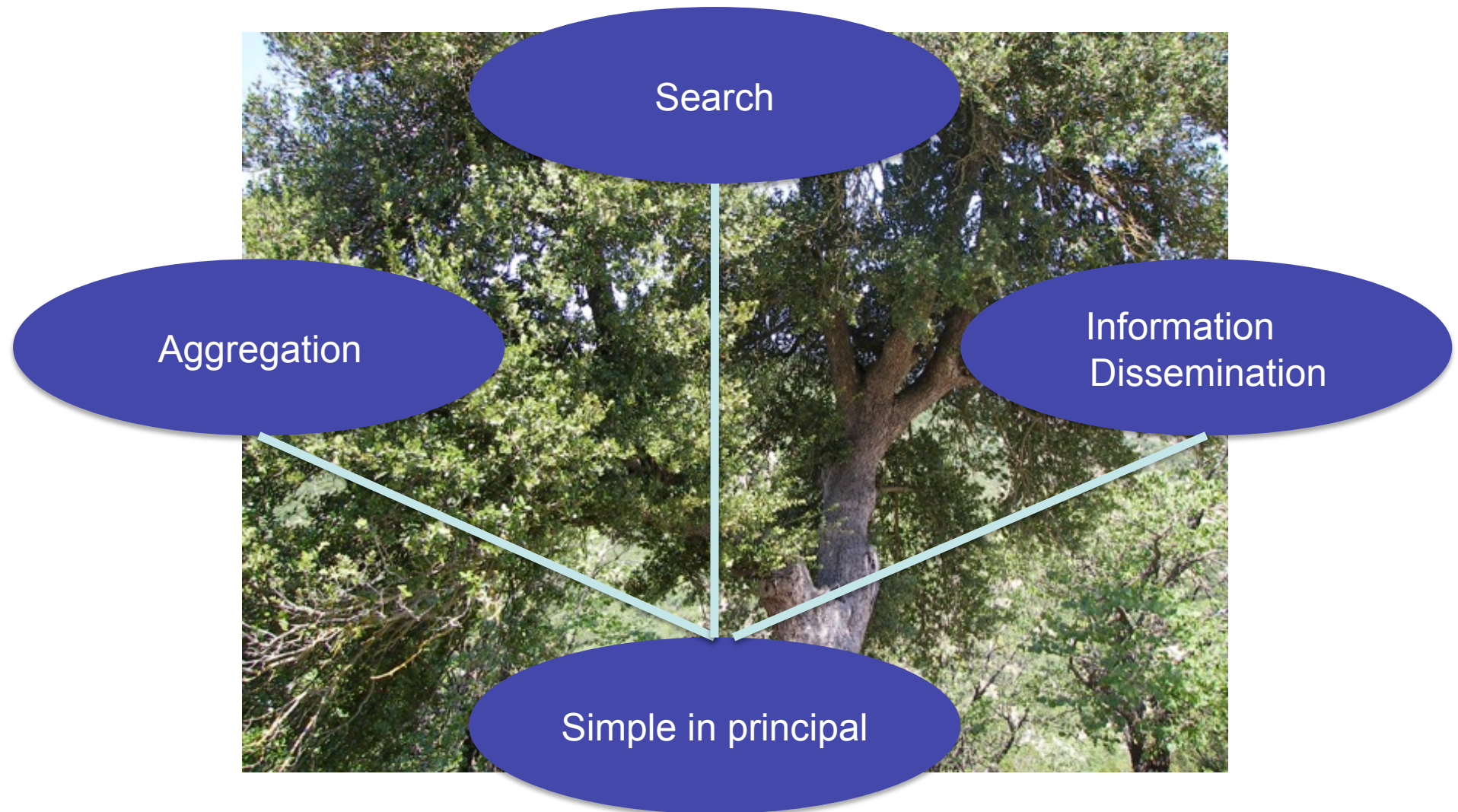
June 2009

# Focus

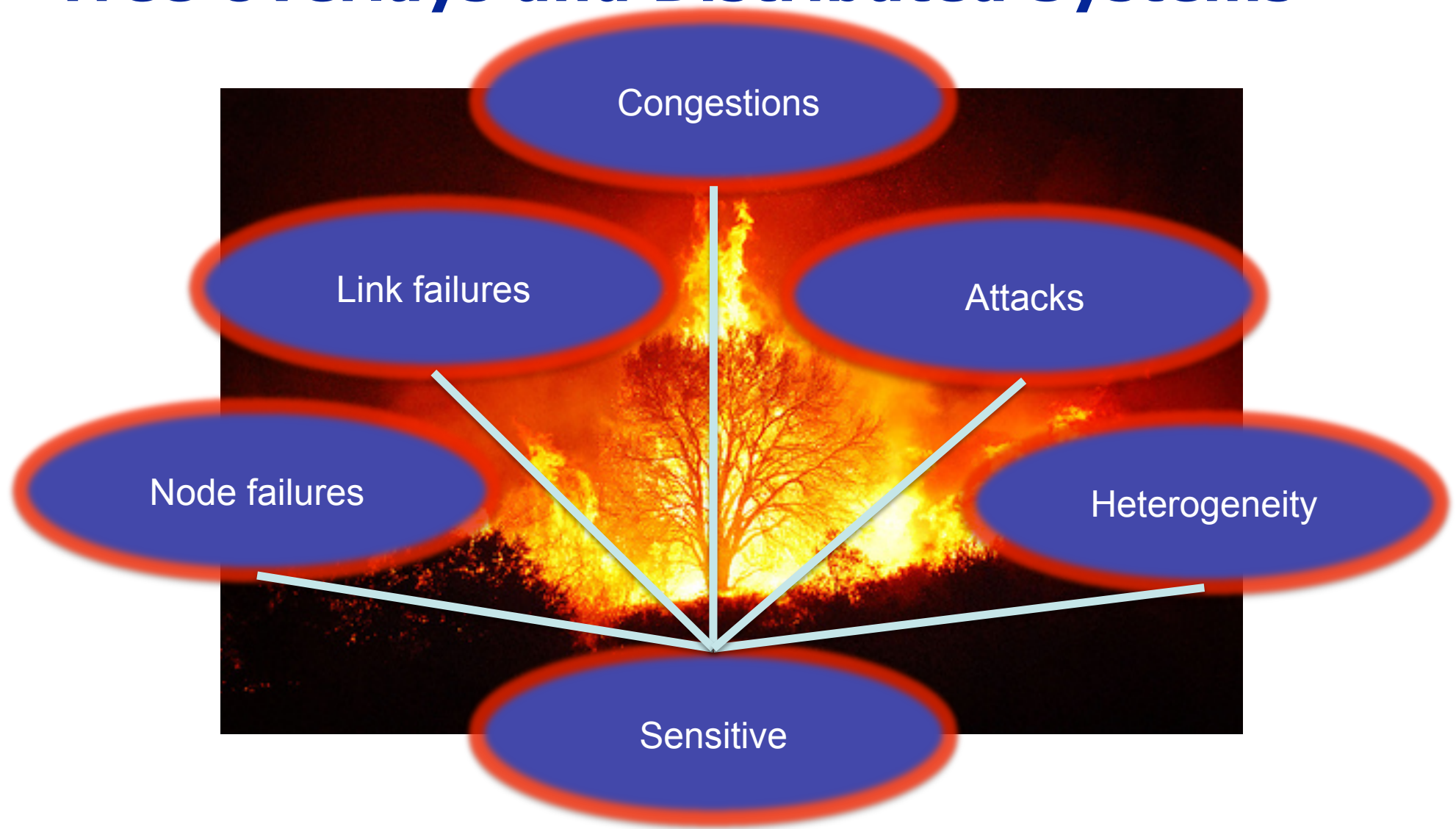


Proposed Approach: AETOS, the Adaptive Epidemic Tree Overlay Service

# Tree Structures

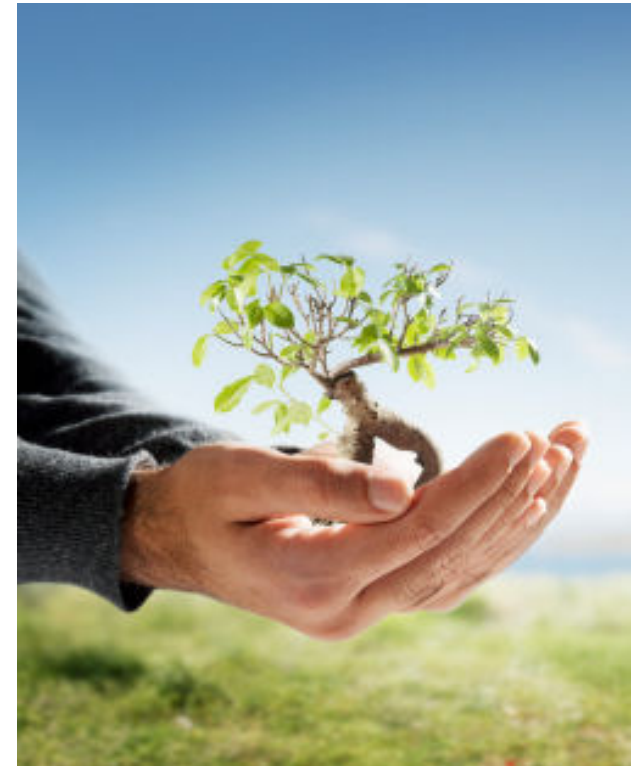


# Tree Overlays and Distributed Systems



# Central Approaches

- Unscalable and not an option for every type of application
- Disregarded in this survey



# General Approach for Reliable Trees

Application Optimization

Fault-tolerance

Self-organization

Maintenance



# Related Surveys

- Focus on one application scenario, e.g. Application Level Multicast
- Classify and examine methods according to the cross-link, in-tree or multiple-tree redundancy
- Tree vs. mesh based systems



- S. Birrer and F. E. Bustamante. A Comparison of Resilient Overlay Multicast Approaches. IEEE Journal on Selected Areas in Communications, 25(9):1695–1705, 2007.
- Z. Li and Y. Shin. Survey of Overlay Multicast Technology. June 2002.
- K.-H. Vik, C. Griwodz, and P. Halvorsen. Constructing low- latency overlay networks: Tree vs. mesh algorithms. In LCN, pages 36–43. IEEE, 2008.
- Y. Liu, Y. Guo, and C. Liang. A survey on peer-to-peer video streaming systems. Peer-to-Peer Networking and Applications, 1(1):18–28, 2008.

# A Survey on Robust Tree Overlays



# Aspects of Investigation

- Application type
- Performance metrics
- Complementary overlay
- Build and maintenance
- Decentralization level
- Proactiveness vs. Reactiveness



# Application Type

Crucial in most of the self-organization approaches

One-to-many communication model, efficient information dissemination, low communication cost

## Application Level Multicast

- G. Tan, S. A. Jarvis, X. Chen, and D. P. Spooner. Performance Analysis and Improvement of Overlay Construction for Peer-to-Peer Live Streaming. *Simulation*, 82(2):93–106, 2006.
- J. Liu and M. Zhou. Tree-assisted gossiping for overlay video distribution. *Multimedia Tools Appl.*, 29(3):211–232, 2006.

## Publish-subscribe, grids, sensor networks

- P. Costa and D. Frey. Publish-Subscribe Tree Maintenance over a DHT. In *ICDCSW '05: Proceedings of the Fourth International Workshop on Distributed Event-Based Systems (DEBS) (ICDCSW'05)*, pages 414–420, Washington, DC, USA, 2005. IEEE Computer Society.
- A. J. Chakravarti, G. Baumgartner, and M. Lauria. The organic grid: self-organizing computation on a peer-to-peer network. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 35(3):373–384, 2005.
- D. England, B. Veeravalli, and J. B. Weissman. A Robust Spanning Tree Topology for Data Collection and Dissemination in Distributed Environments. *IEEE Transactions on Parallel and Distributed Systems*, 18(5):608–620, 2007.

## Distributed databases

- H. V. Jagadish, B. C. Ooi, K.-L. Tan, Q. H. Vu, and R. Zhang. Speeding up search in peer-to-peer networks with a multi-way tree structure. In *SIGMOD '06: Proceedings of the 2006*.
- M. Li, W.-c. Lee, and A. Sivasubramaniam. DPTree: A Balanced Tree Based Indexing Framework for Peer-to-Peer Systems. In *ICNP '06: Proceedings of the Proceedings of the 2006 IEEE International Conference on Network Protocols*, pages 12–21, Washington, DC, USA, 2006. IEEE Computer Society



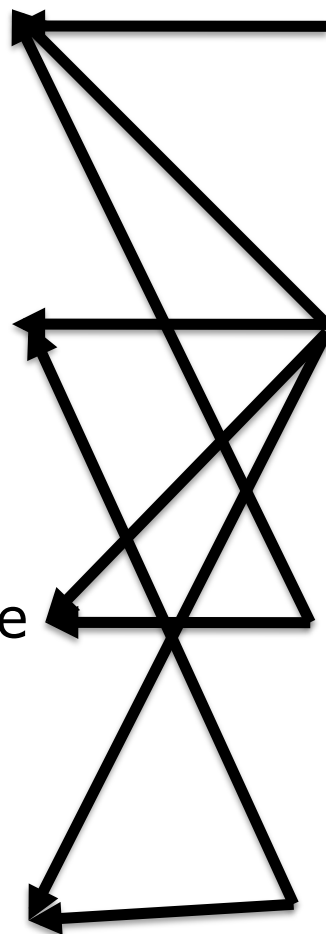
# Performance Metrics

➤ Delay

➤ Bandwidth

➤ Node degree

➤ Uptime



➤ D. England, B. Veeravalli, and J. B. Weissman. A Robust Spanning Tree Topology for Data Collection and Dissemination in Distributed Environments. IEEE Transactions on Parallel and Distributed Systems, 18(5):608–620, 2007.

➤ F. Wang, Y. Xiong, and J. Liu. mTreebone: A hybrid tree/mesh overlay for application-layer live video multicast. In in IEEE ICDCS, page 49, 2007.

➤ S.-W. Tan, G. Waters, and J. Crawford. MeshTree: Reliable Low Delay Degree-bounded Multicast Overlays. Parallel and Distributed Systems, International Conference on, 2:565–569, 2005.

➤ G. Tan, S. A. Jarvis, X. Chen, and D. P. Spooner. Performance Analysis and Improvement of Overlay Construction for Peer-to-Peer Live Streaming. Simulation, 82(2):93–106, 2006.



# Complementary Overlay Support

- Mesh ← ➤ Y. Li and W. T. Ooi. Distributed construction of resource-efficient overlay tree by approximating MST. In ICME, pages 1507–1510, 2004
- Extra links ← ➤ H. V. Jagadish, B. C. Ooi, K.-L. Tan, Q. H. Vu, and R. Zhang. Speeding up search in peer-to-peer networks with a multi-way tree structure. In SIGMOD '06: Proceedings of the 2006 ACM SIGMOD international conference on Management of data, pages 1–12, New York, NY, USA, 2006. ACM
- Gossiping ← ➤ C. Tang and C. Ward. GoCast: Gossip-Enhanced Overlay Multicast for Fast and Dependable Group Communication. In DSN '05: Proceedings of the 2005 International Conference on Dependable Systems and Networks, pages 140–149, Washington, DC, USA, 2005. IEEE Computer Society.
- DHT ← ➤ P. Costa and D. Frey. Publish-Subscribe Tree Maintenance over a DHT. In ICDCSW '05: Proceedings of the Fourth International Workshop on Distributed Event-Based Systems (DEBS) (ICDCSW'05), pages 414–420, Washington, DC, USA, 2005. IEEE Computer Society.
- Central Entities ← ➤ B. Akbari, H. R. Rabiee, and M. Ghanbari. DPOCS: A Dynamic Proxy Architecture for Video Streaming Based on Overlay Networks. In IEEE MICC & ICON '05, 11 2005.
- Skip Graph ← ➤ M. Li, W.-c. Lee, and A. Sivasubramaniam. DPTree: A Balanced Tree Based Indexing Framework for Peer-to-Peer Systems. In ICNP '06: Proceedings of the Proceedings of the 2006 IEEE International Conference on Network Protocols, pages 12–21, Washington, DC, USA, 2006. IEEE Computer Society.



# Build and Maintenance

## Joins, shifts-up and swaps

- G. Tan, S. A. Jarvis, X. Chen, and D. P. Spooner. Performance Analysis and Improvement of Overlay Construction for Peer-to-Peer Live Streaming. *Simulation*, 82(2):93–106, 2006.
- B. Akbari, H. R. Rabiee, and M. Ghanbari. DPOCS: A Dynamic Proxy Architecture for Video Streaming Based on Overlay Networks. In *IEEE MICC & ICON '05*, 11 2005

## Eager and lazy push gossiping strategies

- J. Leitaó, J. Pereira, and L. Rodrigues. Epidemic Broadcast Trees. In *SRDS '07: Proceedings of the 26th IEEE International Symposium on Reliable Distributed Systems*, pages 301–310, Washington, DC, USA, 2007. IEEE Computer Society.

## Bellman Ford, Prim's algorithm

- D. England, B. Veeravalli, and J. B. Weissman. A Robust Spanning Tree Topology for Data Collection and Dissemination in Distributed Environments. *IEEE Transactions on Parallel and Distributed Systems*, 18(5):608–620, 2007.

## Proactive Maintenance, repair strategies

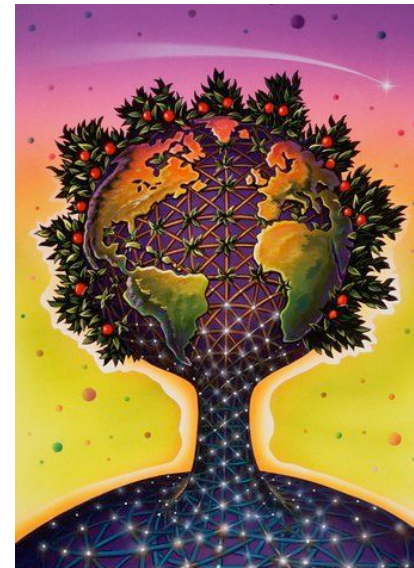
- Z. Fei and M. Yang. A proactive tree recovery mechanism for resilient overlay multicast. *IEEE/ACM Trans. Netw.*, 15(1):173–186, 2007.
- D. Frey and A. L. Murphy. Failure-Tolerant Overlay Trees for Large-Scale Dynamic Networks. In *P2P '08: Proceedings of the 2008 Eighth International Conference on Peer-to-Peer Computing*, pages 351–361, Washington, DC, USA, 2008. IEEE Computer Society.



# Decentralization Level

Most illustrated approaches are distributed  
Self-organizing tree overlays with autonomous nodes

- Hybrid systems
  - Overlay Control Server (OCS)
  - Multicast Server Node (MSN)
  - Content Server
  - Backbone system
  - Super-nodes



- B. Akbari, H. R. Rabiee, and M. Ghanbari. DPOCS: A Dynamic Proxy Architecture for Video Streaming Based on Overlay Networks. In IEEE MICC & ICON '05, 11 2005.
- S. Banerjee, C. Kommareddy, K. Kar, S. Bhattacharjee, and S. Khuller. Construction of an Efficient Overlay Multicast Infrastructure for Real-time Applications. In INFOCOM, 2003.
- J. Liu and M. Zhou. Tree-assisted gossiping for overlay video distribution. Multimedia Tools Appl., 29(3):211–232, 2006.
- mTreebone: A hybrid tree/mesh overlay for application-layer live video multicast. In IEEE ICDCS, page 49, 2007.
- G. An, D. Gui-guang, D. Qiong-hai, and L. Chuang. Bulk- Tree: An overlay network architecture for live media stream- ing. Journal of Zhejiang University, 7(1):125–130, 2006.

# Proactiveness vs. Reactiveness

## Proactive level 1: Support from an underlying overlay

- Y. Li and W. T. Ooi. Distributed construction of resource- efficient overlay tree by approximating MST. In ICME, pages 1507–1510, 2004.

## Proactive level 2: Sorting the nodes according to performance metrics

- G. Tan, S. A. Jarvis, X. Chen, and D. P. Spooner. Performance Analysis and Improvement of Overlay Construction for Peer-to-Peer Live Streaming. Simulation, 82(2):93–106, 2006.

## Proactive 3: Know beforehand the neighbor to connect to in case of failure

- Z. Fei and M. Yang. A proactive tree recovery mechanism for resilient overlay multicast. IEEE/ACM Trans. Netw., 15(1): 173–186, 2007.



## Reactive level 1: detect failures and reconnect (heartbeats)

- J. Liu and M. Zhou. Tree-assisted gossiping for overlay video distribution. Multimedia Tools Appl., 29(3):211–232, 2006.

## Reactive level 2: swaps and shifts-up operations during build

- B. Akbari, H. R. Rabiee, and M. Ghanbari. DPOCS: A Dynamic Proxy Architecture for Video Streaming Based on Overlay Networks. In IEEE MICC & ICON '05, 11 2005

## Reactive level 3: Reactive level 2 in dynamic environments

- A. Walters, K. Bauer, and C. Nita-Rotaru. Towards Robust Overlay Networks: Enhancing Adaptivity with Byzantine-Resilience. Technical Report CSD TR 05–026.

# Discussion and Open Issues

- Peer-to-peer tree overlays can enable the effective utilization of a wide range of application
- Robustness is related to the environment and the application type
- Generic tree overlay (service) for different application types?
- Underlying overlays should enhance the effectiveness of trees and not move their vulnerabilities
- Bridging and gap and unifying proactiveness and reactiveness



# Why do I investigate tree overlays?



# Why do I investigate tree overlays? (cont.)

Energy Management

Stabilization of energy consumption

How?

## **Aggregation and decision-making over a tree overlay**

E. Pournaras, M. Warnier, and F. Brazier. A Distributed Agent-based Approach to Stabilization of Global Resource Utilization. In Proceedings of International Conference of Complex Intelligent and Software Intensive Systems (CISIS'09), March 2009

# Questions?

