

Motivation

- Emerging large scale dynamic networks (LSDNs), e.g., peer-to-peer (P2P) systems, sensor networks, have a wide range of applications
 - Content distribution
 - storage
 - Information/resource sharing
 - Monitoring (complex systems, habitat, battlefield, and etc.)
- Features
 - Large scale
 - No centralized coordination
 - Dynamic
 - High data volume
- Fundamental challenge
 - Manage the large number of host nodes and the voluminous information to facilitate the deployment of various applications ranging from distributed resource locating to network attack detection.

1

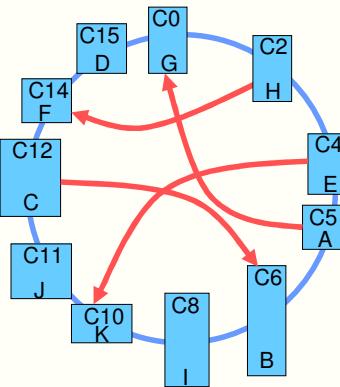
Proposal

- Efficient and robust information management infrastructures
 - Semantic Small World (SSW) – INCP'04
 - Distributed Peer Tree (DPtree) – ICNP'06
 - Multi-level Peer Index (MPI) – MobEA'04
 - DataSocialNet – Patent
- Mechanisms for distributed resource locating
 - Neighborhood signature – IDEAS'03
 - Query processing – in submission
- Mechanisms for distributed knowledge discovery
 - Peer Density-based Clustering (PENS) – InfoScale'06
 - Identify Frequent Items – in submission

2

Semantic Small World (SSW)

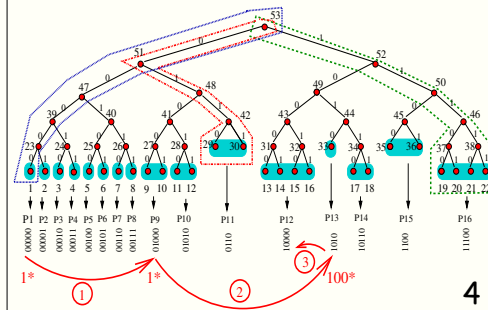
- Motivation
 - Content-based search
 - Resilient to peer join/leave/failure
 - Adaptive to data distribution and users interest
- Design Ideas
 - Small world
 - Semantic clustering
 - Dimension reduction



3

Distributed Peer Tree (DPtree)

- Motivation
 - Support various information management tasks (range query, KNN query, clustering, and etc.)
 - Evolve with dynamic changes on data distribution
 - Balance processing load
- Design ideas
 - Embed a balanced tree in P2P systems
 - Wavelet-assisted distributed mechanism for load balancing



4

Multi-level Peer Index (MPI)

- Motivation
 - Promote P2P information sharing in mobile ad hoc networks
- Design ideas
 - Spatially nested multi-level index
 - Multi-level location services embedded in the index

5

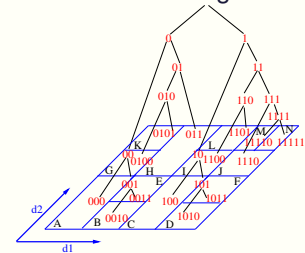
DataSocialNet

- Motivation
 - Autonomy preserving - foundation for trust / incentive / personalization mechanism
 - Search efficient
 - Support various types of data, e.g., numeric, categorical.
- Design ideas
 - Social Identities
 - Social Acquaintanceship

6

Peer Density-based Clustering (PENS)

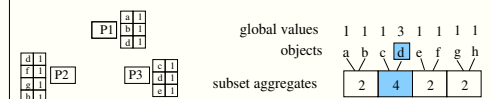
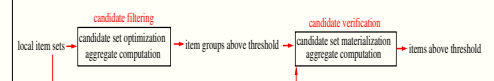
- Motivation
 - clustering in P2P systems: decentralized, incremental
- Design ideas
 - Hierarchy cluster assembly
 - Multi-granularity based cluster membership storage
 - Incremental clustering



7

Identify Frequent Items

- Motivation
 - The problem of identifying frequent items is prevailing:
 - Frequent keywords – cache management
 - Large flow to certain destination – denial of service attack
 - Frequent byte sequences – worm detection
- Design ideas
 - In-network filtering to filter out infrequent items
 - Assign items into groups and obtain the aggregates for item groups
 - Filter out items in light item groups
 - Verify the frequency of candidates



8