Discovery of IoT data streams and IoT services

- Semantic based discovery requires
  - "Semantic description" of objects (IoT data streams and IoT services)
    - Various semantic models have been developed to describe objects in general
      - Single keyword
      - Multiple keywords
        - Bank, Citi, Plano, Texas, ATM-only
      - Attribute-based model
        - Object = Bank, Name = Citi, Location = “Plano, Texas”, …
      - Hierarchical naming (URI, ICN)
  - Semantic models for IoT objects
    - All the general models, SSN, SensorML, Data-Stream-Centric

- Discovery algorithms
  - Centralized: Registry
  - Hierarchical: Bloom filter based
  - Peer to peer
    - SFC+DHT, Query-Covering+DHT
    - CBCB, Cache Summary, Two Level

- A major issue in routing – Routing table size
  - In order to know how to route for a given query, a routing table needs to contain all possible “keys” that may appear in a query
    - For example, in IP routing, the key is the IP address
    - The routing table needs to contain all possible IP addresses that may appear in a routing query, e.g., $2^{32}$ addresses for IPv4 or $2^{128}$ addresses for IPv6
      - Infeasible, needs to reduce the routing table size
  - “Covering” is a very important concept in all routing protocols
    - Example for covering in IP: “128.10.3.*” covers 256 IP addresses
      - Can be used to greatly reduce the routing table size
  - Covering in Query-Covering+DHT, CBCB
    - Attribute based description with covering
      - 1. (object = bank, name = Citi, location = Plano-Texas, property = banking-center)
      - 2. (object = bank, name = Citi, location = Plano-Texas, property = ATM-only)
      - 3. (object = bank, name = Citi, location = Plano-Texas, property = *)
      - 4. (object = bank, name = Citi, location = Texas, property = *)
      - 5. (object = bank, name = C*, location = Texas, property = *)
      - 5 covers 4, 4 covers 3, 3 completely covers 1 and 2
  - Covering in Cache Summary, Two Level
    - Ontology based

- IoT service composition
  - OWL-S specification for IoT services
  - How to compose services based on IOPE

- Potential exam question types
  - Specification models
    - Understand the naming schemes, being able to come up with the naming based on different specification models for some given examples
    - Covering schemes, understand the covering relations
    - Understand ontology coding in ontology based covering
      - How to construct ontology code, how to use ontology code in routing table construction and in query routing
- Discovery algorithms
  - Bloom filter
    - How to use Bloom filter to determine how to route discovery queries
  - SFC+DHT, Query-Covering+DHT
    - How the objects are distributed over the network (where they are, how to determine it)
    - During discovery, how to find the object
    - Given a scenario, being able to follow the algorithm to do object distribution and to do object discovery
    - Analyze the algorithms and understand their pros and cons
  - CBCB, Cache Summary, Two Level
    - Build routing table during advertisement
    - Routing for a discovery query based on the given routing tables
    - Given a scenario, being able to follow the algorithm to build routing table and to route the discovery query
    - Analyze the algorithms and understand their pros and cons
    - Understand the impact of usage, distance, coverage, mobility to routing and routing table
- IoT service composition
  - Able to compose services based on the given IOPE of services

Edge for IoT Computing
- Cloudlet
  - Major concept of Cloudlet
  - Major steps in Cloudlet to minimize the size of the VM-overlay
  - Cloudlet migration issues
- Docker
  - Docker lifecycle
  - Docker image layering based on Docker file
  - Docker file structure
  - Container migration
- Cloud4Home
  - Routing for query involving data and computation
- ECC
  - Specification model
    - Competence, intent, context
  - Matchmaking
- FocusStack
  - GA specification and other specifications
  - FCOP
  - How to follow the protocol to locate the queried objects (IoT devices)
- Geo-distributed data analytics
  - How to place Reduce tasks, how to re-locate input data
- Learning on the edge
  - Pruning, code-book based quantization
- Potential exam question types
  - Cloudlet and Docker
    - Able to compare them and understand their pros and cons
    - Given scenarios, able to follow the Cloudlet steps to compute the VM-overlay for carrying and for migration
    - Understand fully the AUFS structure for Docker
      - Given scenarios, know how Docker AUFS would look like
♦ Understand the Container migration method, being able to follow the migration steps for a given scenario

➢ Cloud4Home, ECC, FocusStack
  ♦ Understand the matchmaking procedures of these algorithms
  ♦ Given a scenario, being able to perform matchmaking following the algorithms or a mixture of these algorithms

➢ Geo-distributed data analytics
  ♦ Given a scenario, being able to place the Reduce tasks by migrating data

➢ Learning on the edge
  ♦ For a given scenario, being able to perform pruning and quantization