Energy-Efficient Provenance Transmission in Large-Scale Wireless Sensor Networks

S. M. Iftekharul Alam
Electrical and Computer Engineering, Purdue University
Sonia Fahmy
Computer Science, Purdue University

Emergence of Large Scale Sensor Networks
- Global Sensor Network to fight climate change.
- Sensor based decision support systems to monitor power grid and critical infrastructures:
  - Smart Grid
  - Smart Building
  - Smart Bridge
  - Smart Tunnel

Provenance and Trust Framework
- Trust models assess trustworthiness of data based on provenance similarity and value similarity.
- Provenance of a data item is a tree of nodes that manipulate or forward that item.

Trustworthiness of data affects the quality of decision making.

Probabilistic Provenance Flow (PPF)
Adaptation of probabilistic packet marking (PPM) of IP traceback
Embedding a connected sub-graph of full provenance into a single packet
Two complementary encoding schemes:
- (a) Juxtaposition of ranks and (b) Prime multiplication
Faster decoding and construction of provenance

Provenance Encoding
- \( \text{prime}(n) = \) The greatest prime number less than or equal to \( n \).
- \( \text{offset}(n) = n - \text{prime}(n) \).
- Difference between node ID and \( \text{prime}(ID) \) is less than or equal to 7.
- \( \text{rank}(ID) = \) Position of ID in an increasing sequence of IDs of all member nodes.

Provenance Decoding and Construction
- Rank method: Use the counter to extract partial provenance.
- Use Prime Factorization to extract data from prime multiplication field.
- Use solution to Subset Sum problem to extract data from offset field.
- Exploit ordering information extracted by rank method.

Full provenance: at least one ID for every node is received.

Goals and Challenges
- Probabilistic incorporation of node ID to reduce the expected length of the provenance.
- Number of bits required to represent provenance should be fixed.
- Fast convergence of provenance construction is critical.
- Topological changes should be rapidly reflected in provenance.

TOSSIM Simulation
- PPF requires 33% fewer packets than PPM based approaches of IP traceback.
- PPF consumes 30% less energy than PPM with network coding.
- Trust model integrated with PPF provides high level of accuracy for trust score calculation.