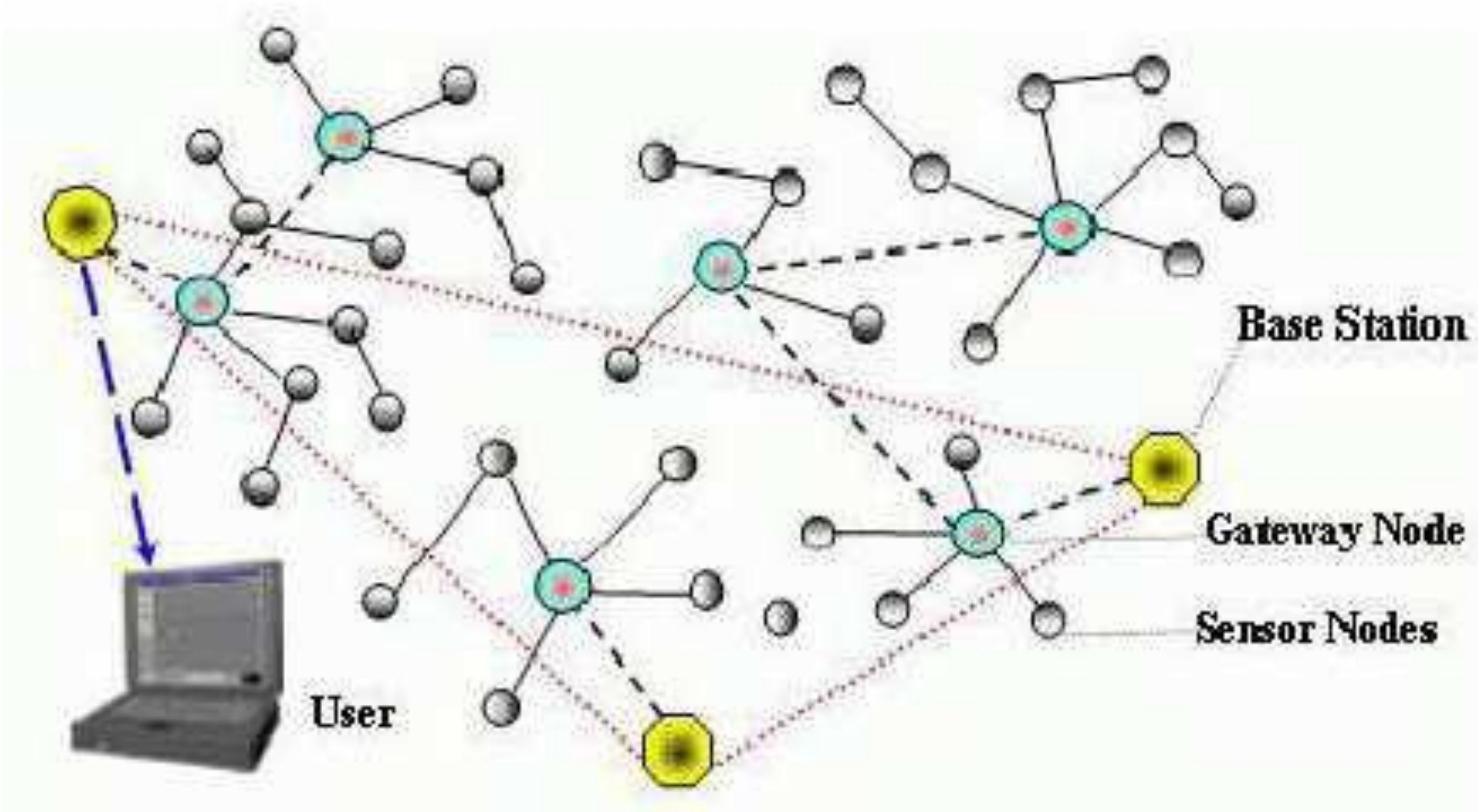
Privacy Homomorphism

Dr. Vivaksha Jariwala

Associate Professor

Wireless Sensor Networks



Wireless Sensor Networks

- A Sensor Network is a network of such sensors that can Sense specified parameter relating to their environment
  - Process them either locally or in a distributed manner Communicate processed information to base station
- WSNs gaining popularity – low cost solution to real world challenges
- Military, environmental monitoring, health monitoring, home appliances, civilian, societal surveillance applications

Challenges in Ensuring Security

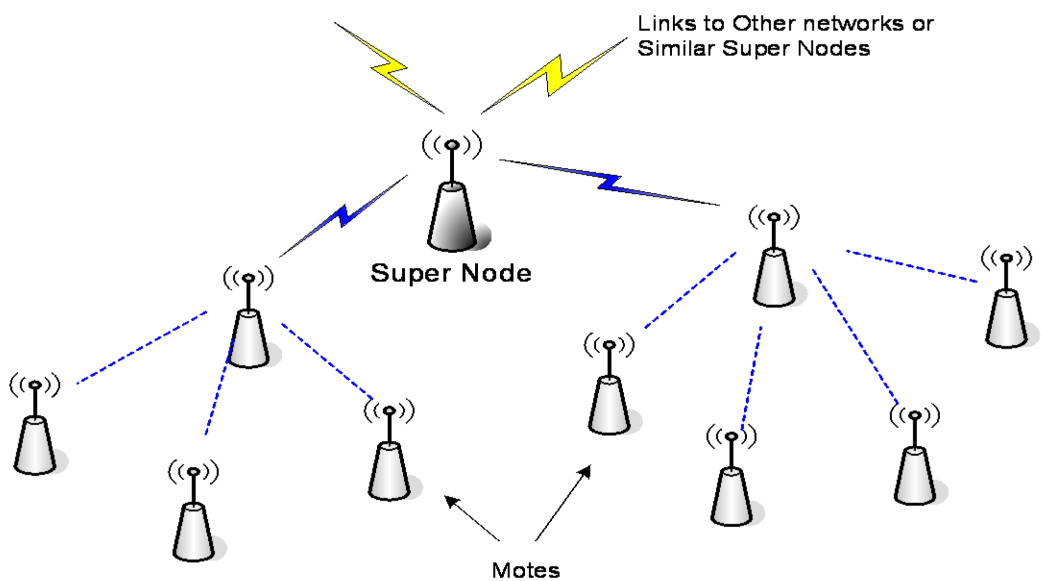
- Ensuring secure communication in Wireless Sensor Networks is a challenge.
- The challenges are due to : ([Akyildiz][Karlof et al])
  - the open-to-all wireless communication, deployment in evasive environments
  - inherently resource intensive security algorithms, inherently resource starved WSN nodes
  - conventional route-centric multihop protocols not directly applicable - data-centric multihop communication
  - In-network processing…..what is it ?????

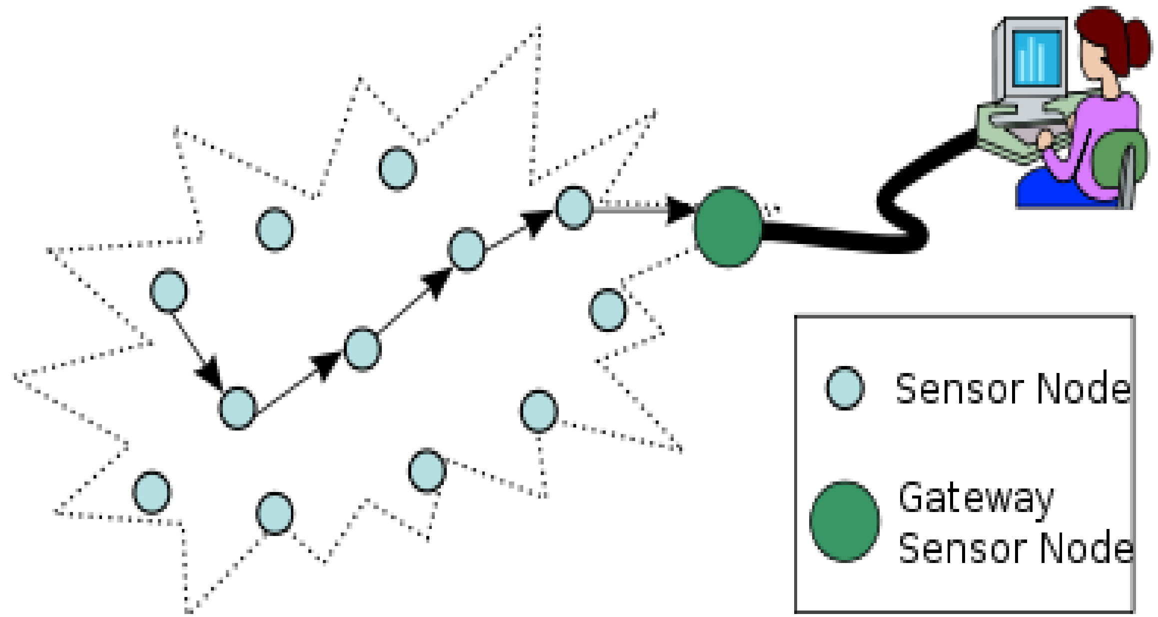
In-network Processing

- In-network processing is…..
  - processing done on-the-fly on a packet in transmission
  - enables reduced packet transmissions to the base station
  - leads to a fundamental distinction between data-centric multihop communication and route-centric multihop communication
- An example to understand better……….

Motivation: In-network Processing

- Major and dominant application scenario for WSNs is
  - environmental monitoring
    - wherein data sensed at different distributed locations is transmitted to a central point viz. base station.





© Maher

Motivation: In-network Processing

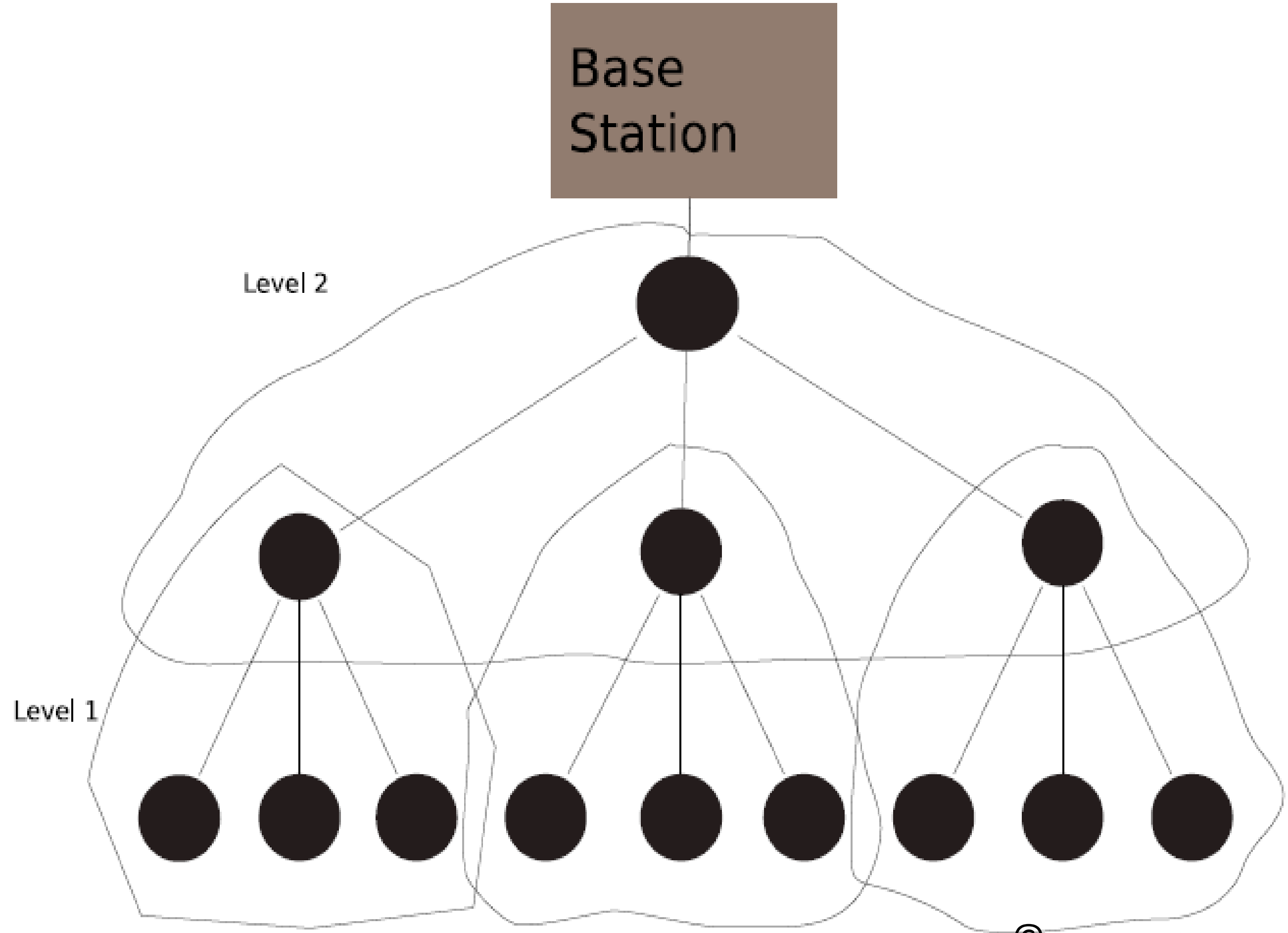
- The data collected is required to
  - be analyzed further, that eventually serves to initiate some action.
  - such analysis is typically based on pre-computation of an optimum e.g.
    - computing the minimum/maximum/sum/average/variance/duplicate elimination….
- Where to do such pre computations ?
  - Two alternatives
    - at the central point i.e. the base-station OR
    - in the network itself
- Which one of the two is a better alternative ?

Motivation: In-network Processing

- Which one of the two viz. computation at the base station or in-network computation is a better alternative ?

- Centralized Pre-computation
  - Leaf nodes….9 messages
  - Level 1……….12 messages
  - Level 2……….13 messages
  - Total messages = 34

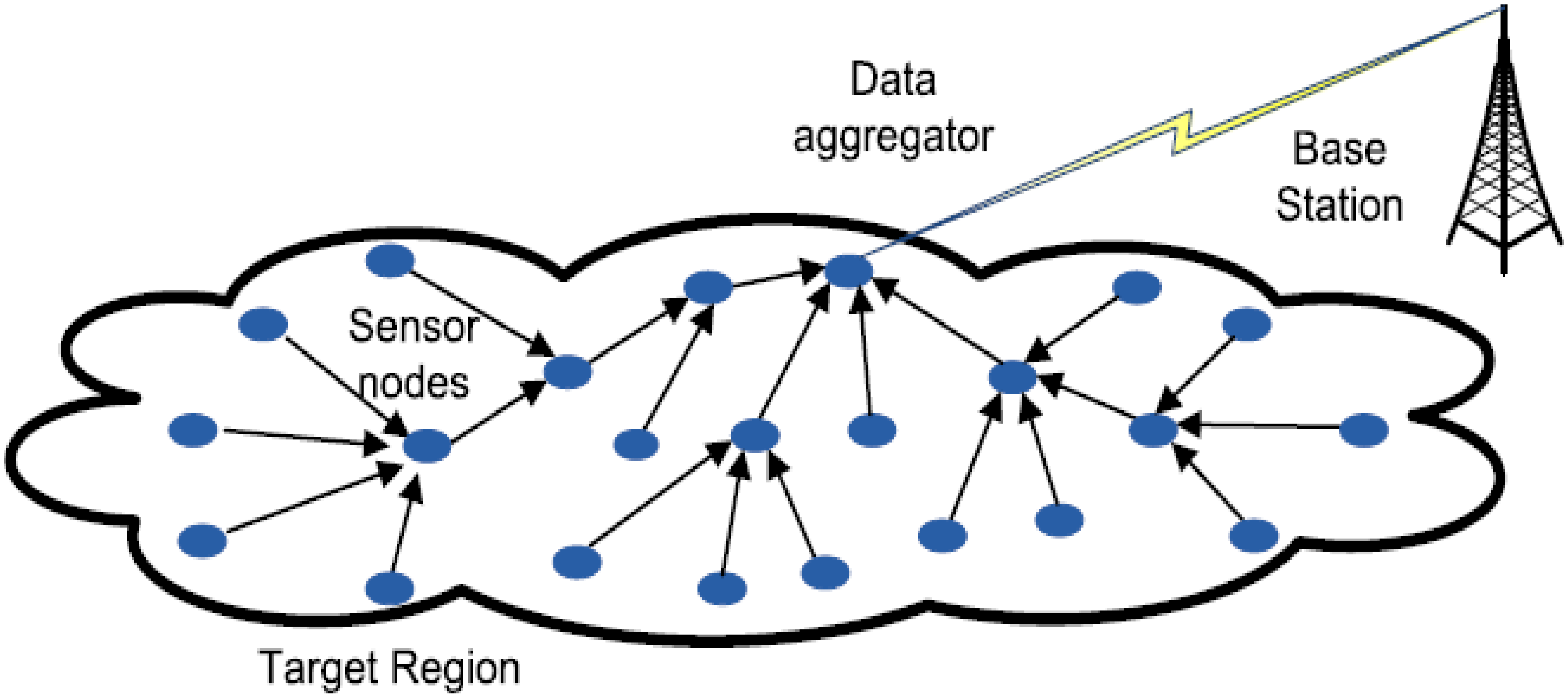
- De-centralized pre-computation
  - Leaf nodes….9 messages
  - Level 1………..3 messages
  - Level 2………..1 message
  - Total messages = 13



In-network Processing

- In-network processing is…..
  - processing done on-the-fly on a packet in transmission
  - enables reduced packet transmissions to the base station
  - Data-centric multihop communication
    - yielding finer granularity of processing
    - necessary in the resource starved sensor nodes
  - Route-centric multihop communication
    - offers coarse granularity of processing
    - tolerable in the resource rich conventional PCs

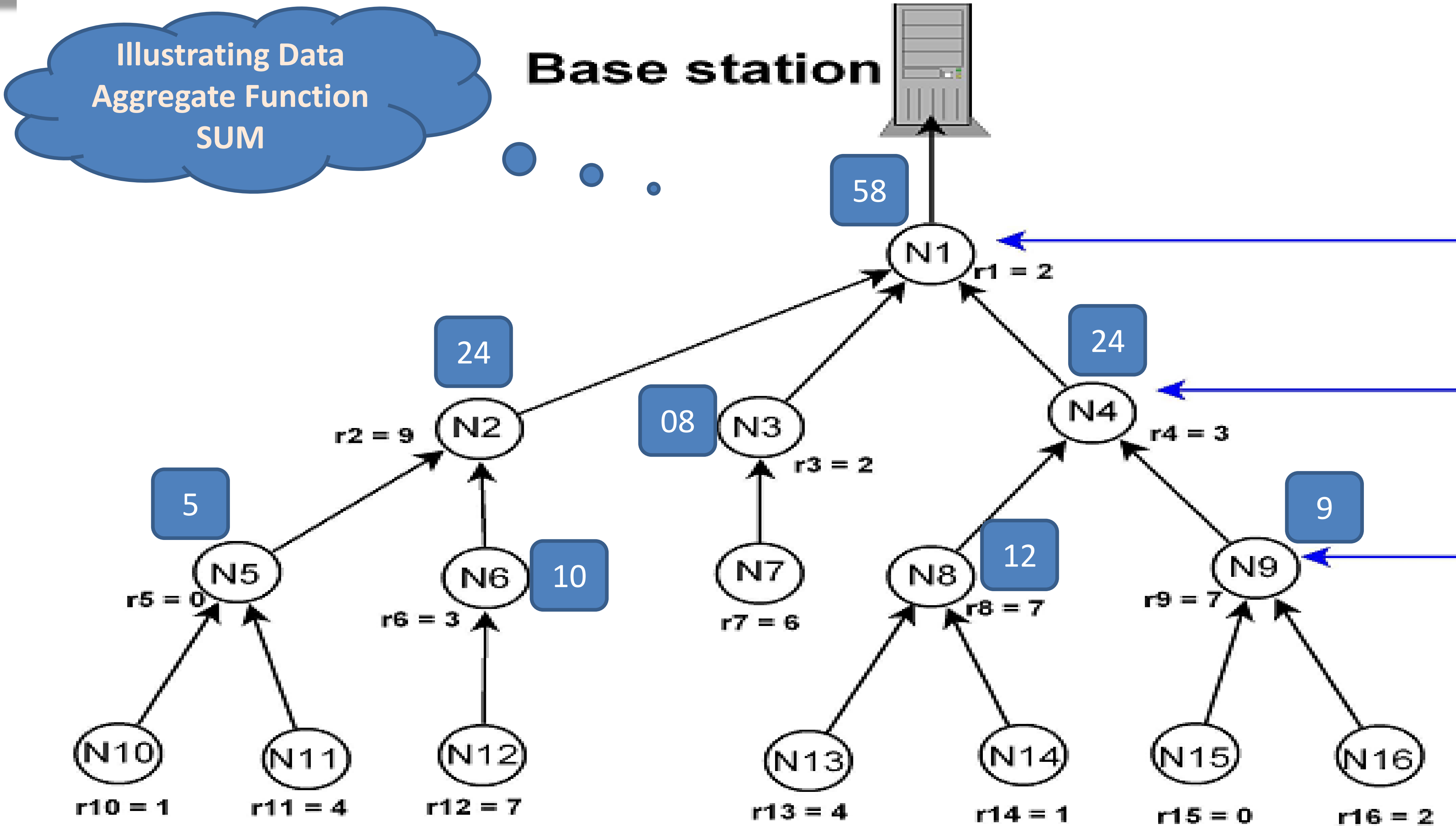
Data Aggregation…



Data Aggregation...

- Aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to the base station
- It usually involved the fusion of data from multiple sensors at intermediate nodes and transmission of aggregated data to base station (sink)
- Desirable Properties
  - Energy efficiency
  - Network Lifetime
  - Data Accuracy
  - Latency

Data Aggregation : An example



Data Aggregation : consequences

- Data aggregation
  - is an efficient way to minimize energy consumption on sensors, but it also creates new security challenges.
  - in a multihop sensor network, a forwarder node
    - by default observes the incoming data, that it has to process
    - can potentially manipulate data coming from its children in the routing tree and affect the aggregation result.
  - this can happen at forwarders as well as the aggregators.
  - identification information of the data is lost once it is aggregated, making the detection of malicious nodes more difficult.

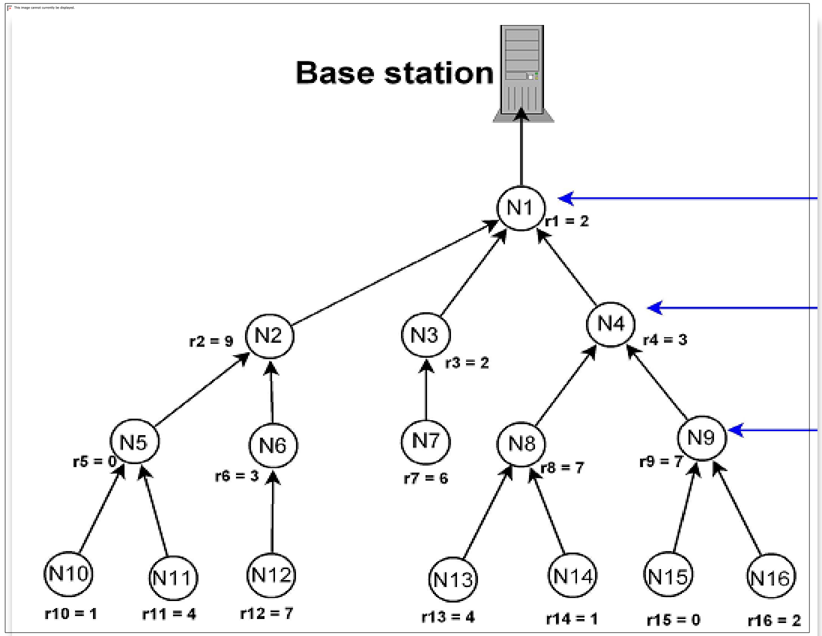
- WSNs are deployed in hostile environments, making the sensors susceptible to attack by an adversary.

What could be the solution strategy ?

- Use Secure Data Aggregation
- What could be Secure Data Aggregation ?
- What could it be based on ?

Secure Data Aggregation

- def:
  - Secure Data Aggregation is the efficient delivery of the single processed/summarized result reported to an off-site user (or a base station), obtained from a number of raw sensor readings,
    - maintaining their privacy in such a way that ensures these reported raw readings have not been altered in the process. (adapted from [Przydatek et al]).



- Primary Objectives
  - Confidentiality
  - Robustness of data
    - Message/entity authentication
  - Privacy of the data sensed

Why Secure Data Aggregation ?

- There is a strong conflict between data security and data aggregation protocols.
  - security protocols at the application layer are end-to-end
    - sensor nodes prior to its transmission encrypt/authenticate sensed data
    - to be decrypted only at the base station [Alzaid et al][Lingxuan et al].
  - On the other hand, data aggregation protocols natively use plain data to implement data aggregation at every intermediate node
    - end-to-end integrity check not viable
      - data aggregation results in alterations in sensor data
    - necessary to provide source and data authentication along with data aggregation.

Secure Data Aggregation : Paradigms

- Broadly two paradigms in the literature to ensure Secure Data Aggregation.
  - Using Hop-by-Hop encryption i.e.
    - Secure Data Aggregation using a Link Layer Security Architecture (LLSA)
      - e.g. TinySec, SenSec, MiniSec, FlexiSec, OR IEEE 802.15.4….
    - Using specific adhoc approaches [Sang et. al.]
  - Using End-to-End encryption [Sang et. al.] i.e.
    - Secure Data Aggregation using Homomorphic Encryption
    - focus on imposing security operations on the processed data
    - also known as Concealed Data Aggregation (CDA)

[Ozdemir] [Castellucia] [Piotorwski].

Hop-by-hop Secure Data Aggregation

Hop-By-Hop Secure Data Aggregation

- Secure Data Aggregation using a Link Layer Security Architecture (LLSA)
  - requires multiple encryption-decryption i.e. security operations at each link
    - e.g. TinySec [Karlog et al], MiniSec[Luk et al], SenSec, FlexiSec[Jinwala et al], IEEE 802.15.4 based radio chips…
  - increases overall resource overhead……why ?
  - increases vulnerability to attacks
    - repeated encryption/ decryption at each hop in the network
  - offers only security (and thereby robustness) and not privacy

Questions

- Privacy???
- Difference between confidentiality and privacy??

Privacy

- **Privacy** is the control over the extent, timing, and circumstances of sharing oneself (physically, behaviourally, or intellectually) with others.
- Examples of activities considered private might include
  - a medical examination;
  - activities within your home;
  - using a restaurant bathroom;
  - entering the office of a reproductive health provider;
  - generally any action for which you have the reasonable expectation of privacy.
- Most things done in public places would not be considered private.

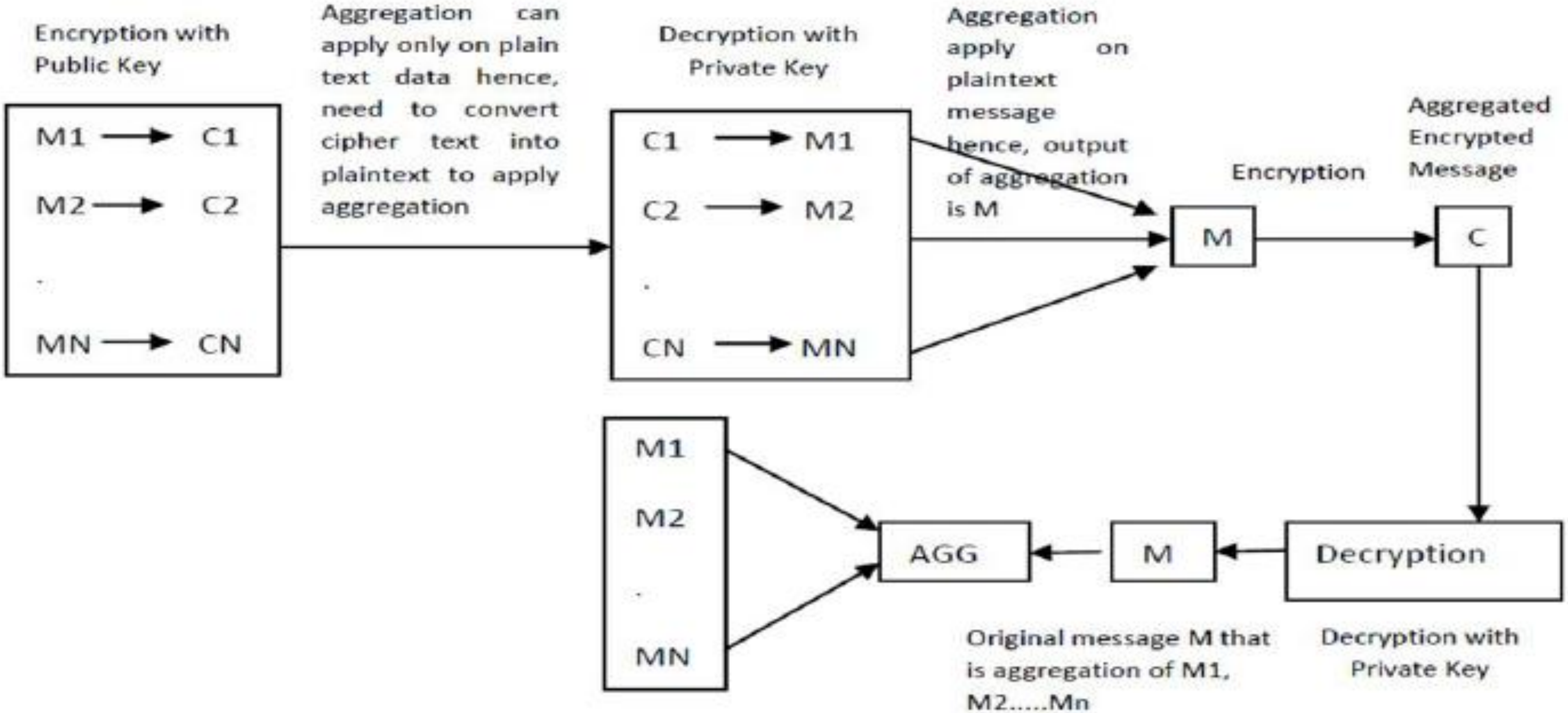
Privacy…

- **Huge databases** exist in various applications
  - Medical data
  - Consumer purchase data
  - Census data
  - Communication and media-related data
  - Data gathered by government agencies
- **Can these data be utilized?**
  - For medical research
  - For improving customer service
  - For homeland security

Methods for Privacy

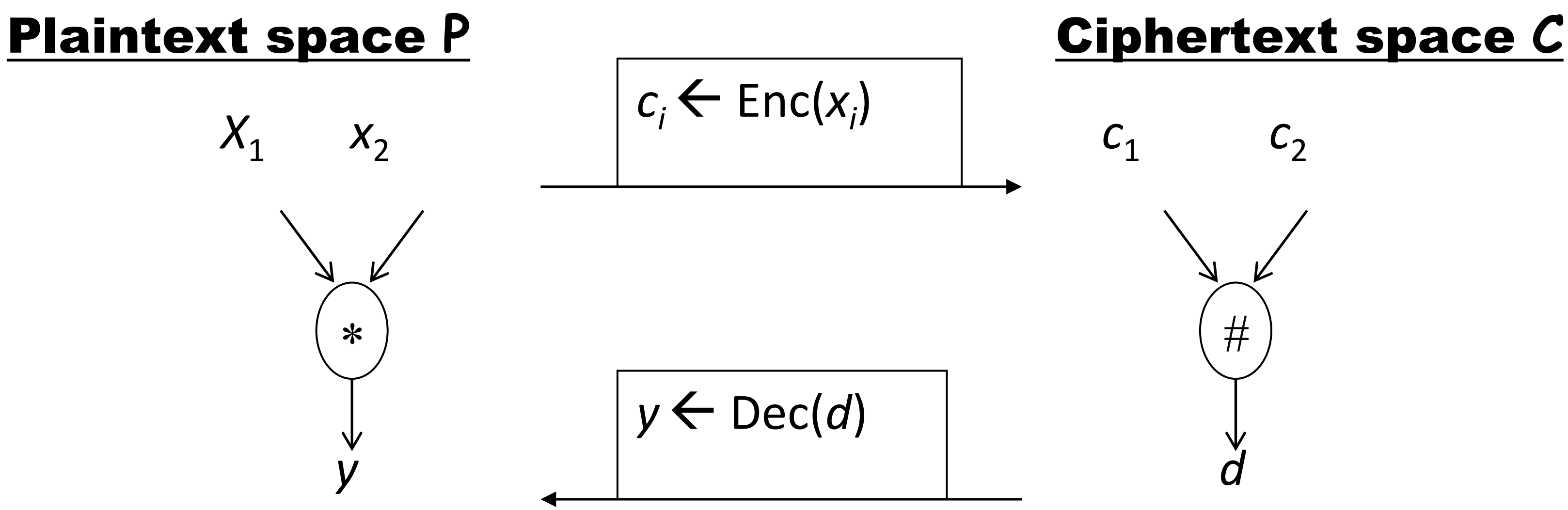
- Multiparty Computation
  - Secret Sharing
- Privacy Homomorphism

Conventional Encryption

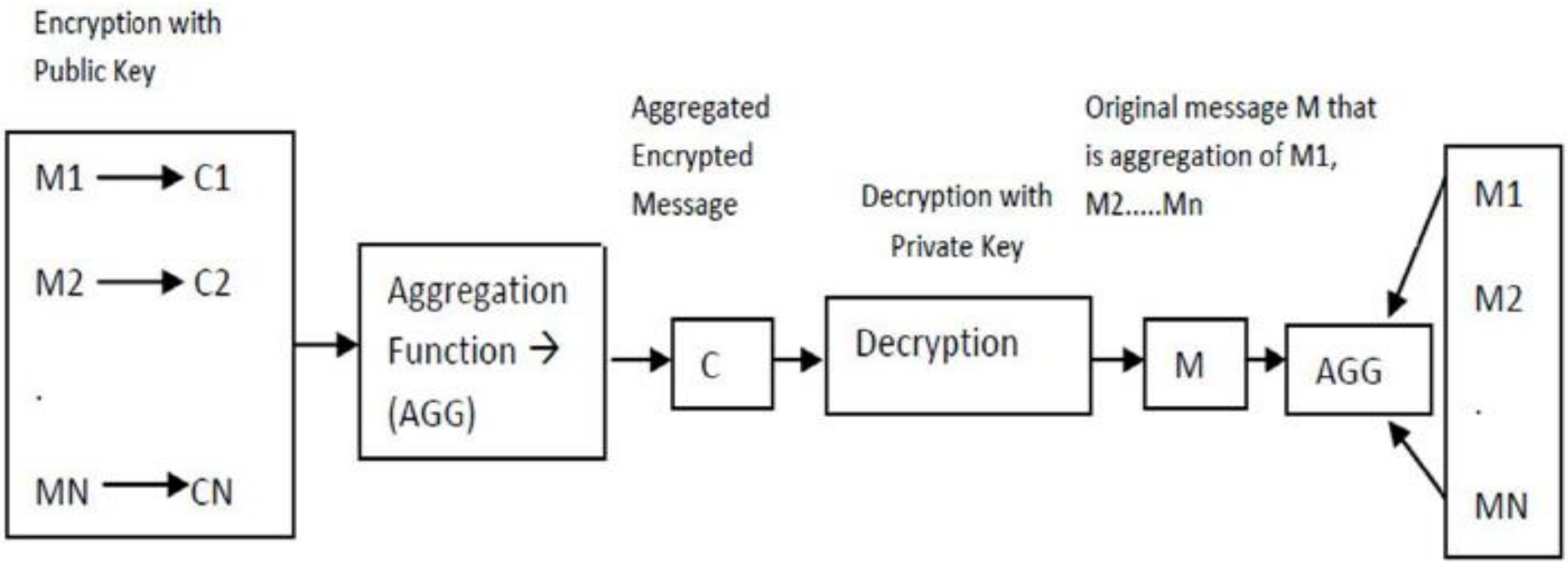


Privacy Homomorphism

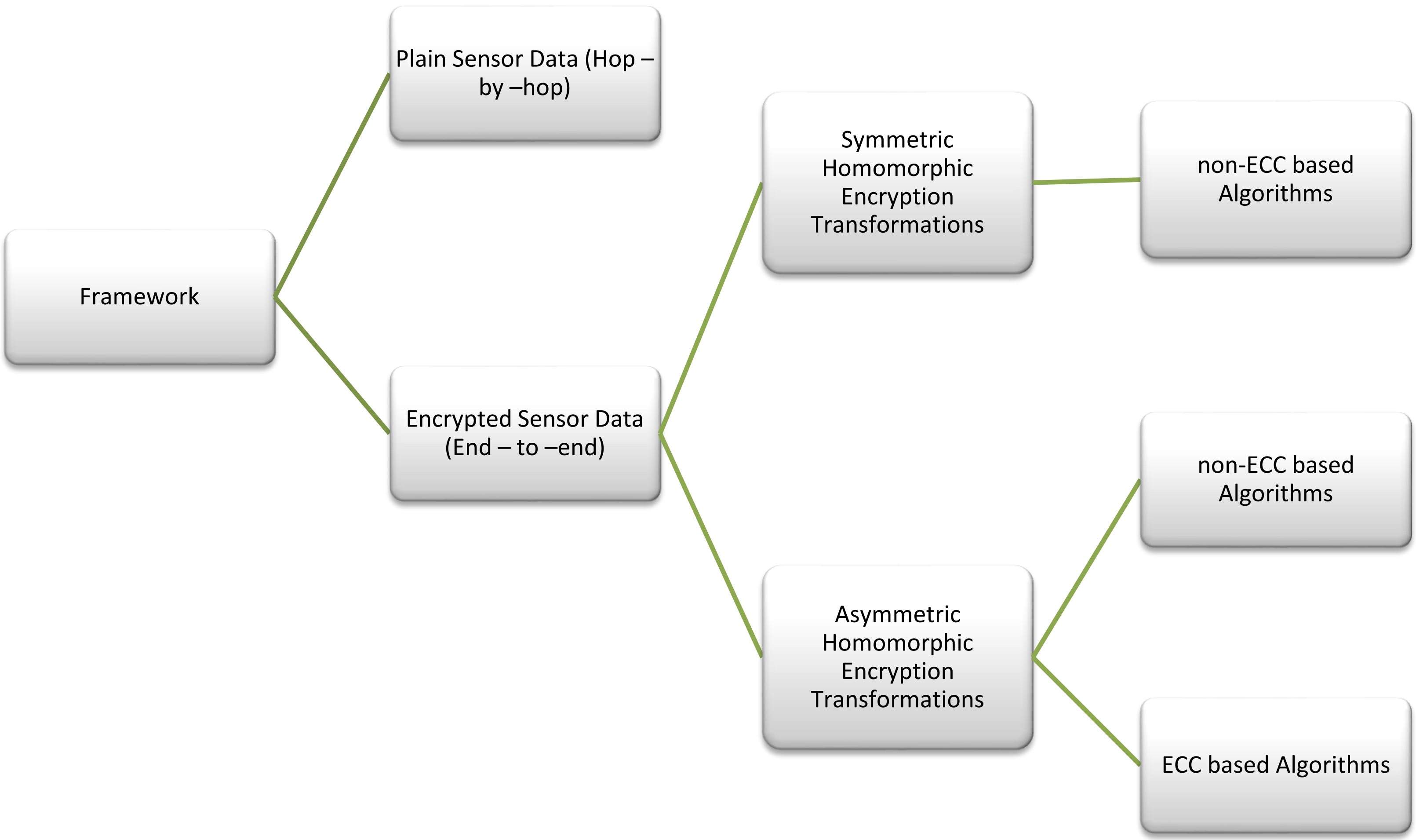
- Privacy Homomorphism is encryption transformation that allows direct computation on encrypted data.
- An encryption algorithm E() is homomorphic, if for given E(x) and E(y) one can obtain E(x * y) without decrypting x,y for some operation *.
- Ek(a + b) or Ek(a x b) from ciphertexts Ek(a) and Ek(b) without the knowledge of the decryption key



Privacy Homomorphism



Secure Data Aggregation : Types



Classical Algorithms

- Domingo Ferrer
- Castellucia
- Combined Approach

Symmetric Key

- Okamoto Uchiyama
- Goldwasser Micali
- Benaloh
- Elgamal
- RSA
- Paillier

Asymmetric Key

Castellucia

Algorithm Casstelluccia ()

Parameters: Select large integer M

Encryption: Message $m \in [0, M-1]$,

Randomly generated key stream $k \in [0, M-1]$

$c = (m + k) \bmod M$

Decryption: $m = (c\text{-}k) \bmod M$

Aggregation: $c_{12} = (c_1 + c_2) modM$

Castellucia...

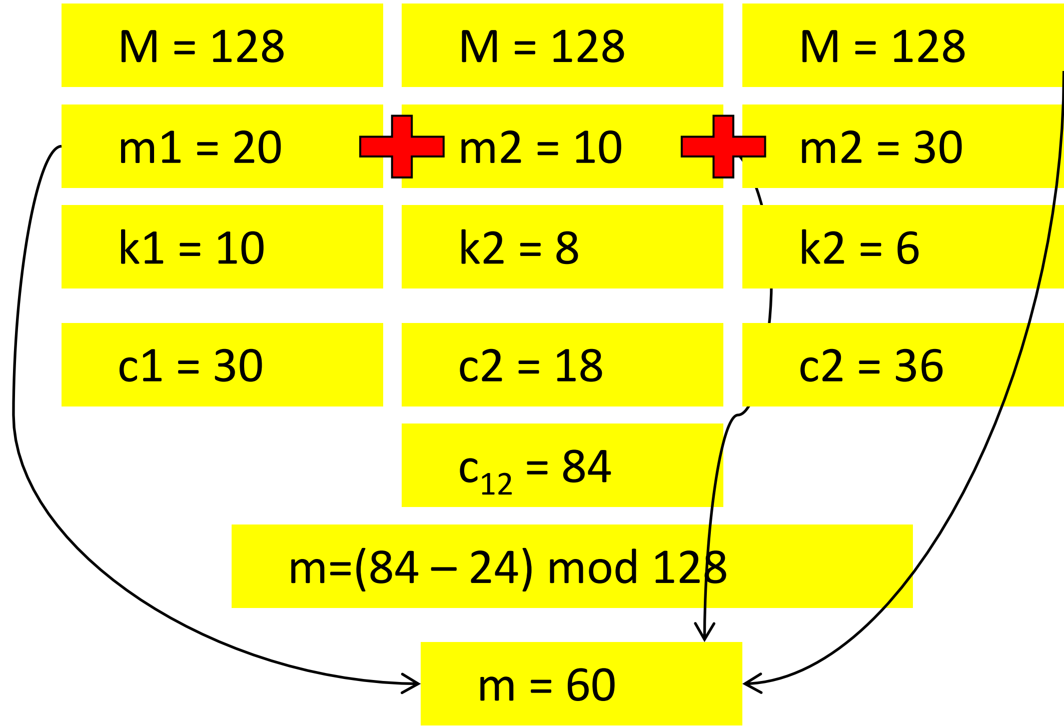
Parameter: select large integer M

Encryption: Message m ϵ [0,M – 1],
randomly generated key stream k ϵ [0,M – 1]

c = (m + k) mod M

Aggregation: $c_{12} = (c_1 + c_2) \bmod M$

Decryption: m = (c – k) mod M



ECC Based Algorithms

- Elliptic Curve Okamoto Uchiyama (EC-OU)
- Elliptic Curve Paillier (EC-P)
- Elliptic Curve Naccache-Stern (EC-NS)
- Elliptic Curve ElGamal (EC-EG)

Thank You !!!!