

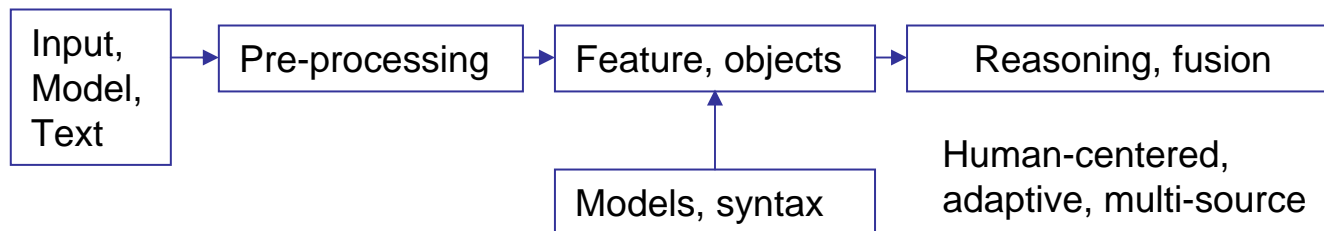
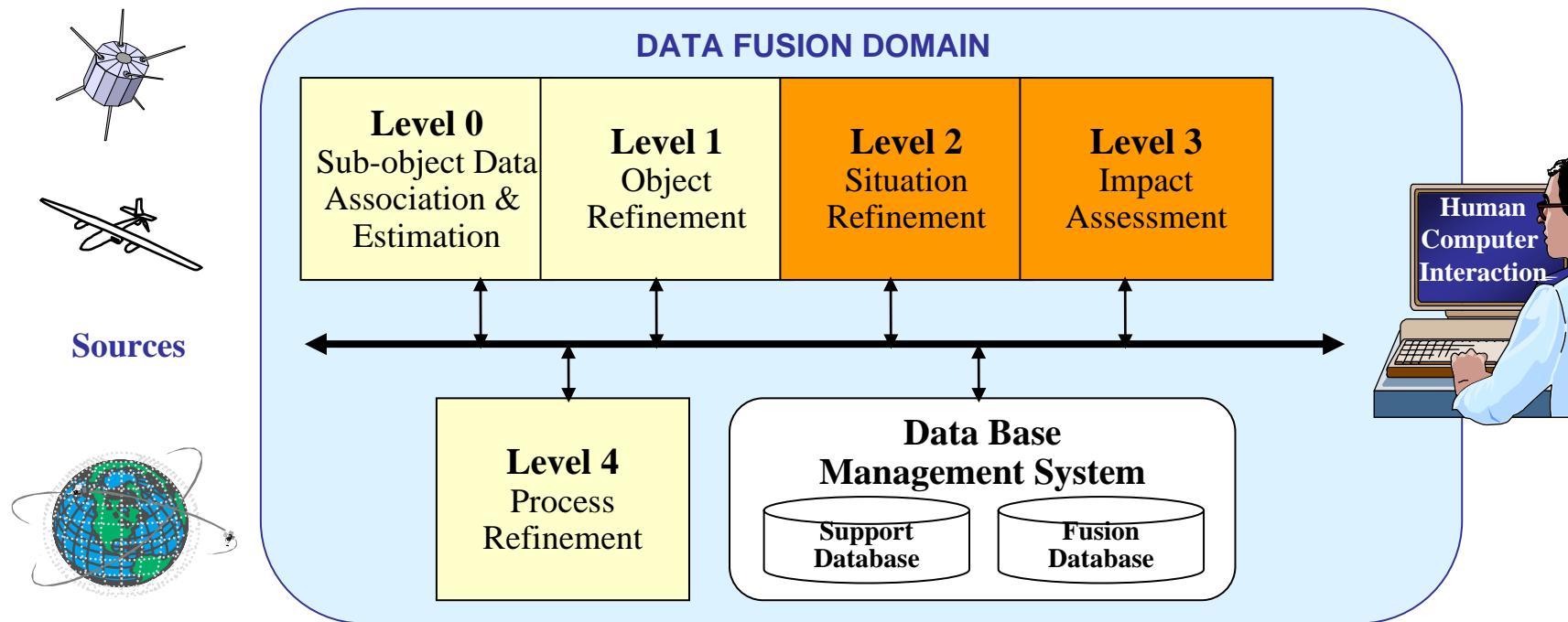
Collaborative Distributed Data Fusion

Prof. Wu Jian Kang
Graduate School, Chinese Academy of Sciences
jkwu@gucas.ac.cn

Definition

- Data Fusion, Sensor Fusion, Information Fusion.
- “Information fusion is an Information Process dealing with the:
 - association, correlation, and combination of data and information, from
 - single and multiple sensors or sources, to achieve
 - refined estimates of parameters, characteristics, events, and behaviors, for observed entities in an observed field of view
- Models: The most referenced are probably variations of the OODA-loop (Observe, orient, decide, acts) and the JDL-model.

Data Fusion Functional Model (JDL)



Energy → Signals → Data → State vectors → Labels → Knowledge

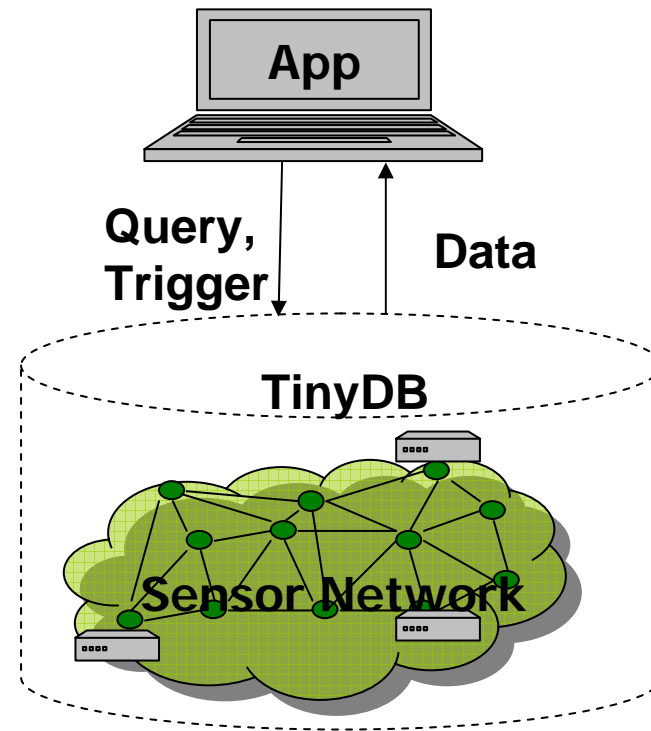
Challenges of Data Fusion in Sensor Networks

- **Uncertainty of sensors:**
 - no perfect sensors available
 - difficult to predict sensor performance
 - to effectively use 10^N sensors
 - Heterogeneous sensors
 - Power constraints
 - Package loss rate is high with wireless communication
- **Dynamics of environments**
 - The sensor network is embedded into the real world. Its structure and data fusion strategy and algorithm must adapt to the environment changes.
 - to effectively task geographically distributed non-commensurate sensors
- **Dynamics of targets**
 - Targets can appear any time, any where, with any speed, under any weather conditions.
 - There is insufficient training data
- **Human computer interface (HCI):**
 - to know how to link decision needs to sensor management
 - Incorporating human knowledge into the decision process
- **Conclusion: Challenges lie on level 0 and level 1 fusion**

TinyDB: fusion by aggregation

- UCB and Cornell
 - Sensor network is a database
 - Descriptive language
 - Subscription / query
- Issues to be addressed
 - Data model is very simple, managing simple data types such as temperature, noise, etc.
 - Does not address data stream, which is the nature of sensory data.

```
SELECT MAX(mag)
FROM sensors
WHERE mag > thresh
SAMPLE PERIOD 64ms
```



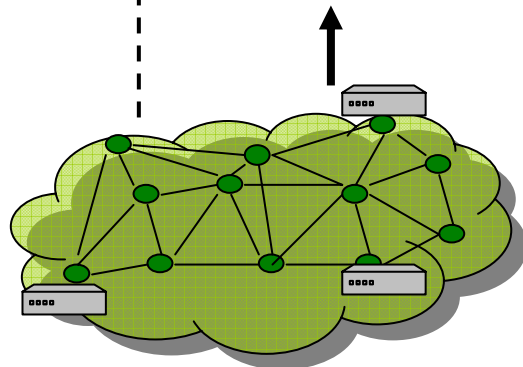
Our Solution – Spatiotemporal Stream Segmentation

Web / Database & Applications
Well-formed data: Formatted with semantics



There is no link now

Stream Data Segmentation



Embedded and Networked Sensors
Multimodal, space-time stream with uncertainty and embedded world models

Applications

Converting continuous accelerometer data stream into activity information l^i

$$\{WearerID, (a_{1,t}^{waist}, a_{2,t}^{waist}), (a_{1,t}^{left}, a_{2,t}^{left}), (a_{1,t}^{right}, a_{2,t}^{right})\}$$

$$\Rightarrow \{wearerID, (l^1, t_{start}^1, duration^1, a_{sample}^1), \dots, \lambda_{activity}\}$$

Segment and label time sequences (Bayesian Networks)

$$\text{MAP: } p(Y_k^s | Z_{1:k}^s)$$

Z^s is composite measurement of a neighborhood at s

Spatial data modeling: Random Field for real world

1. Spatial location v and relation e ;
2. Random measurements z_v^j at v for j^{th} sensor modality
3. A neighborhood of v for a task

Random Field – a world data model

- Random Field and Graph Model
 - Sensor network can be best represented by a graph $G=(V, E)$, with V denoting vertex set and E the edge set. Where v denotes the spatial location of a sensor node, with its coordinates in the geographic space.
 - Let $Z_v=\{z_v^1, \dots, z_v^M\}$ be random value measurements indexed by vertex v and modality m . z_v^m represents the measurement for the m^{th} sensor mounted on the sensor node at location v .
 - The edge between two vertices represents the relationships and dependences between the two data vectors

Neighborhood system for collaborative processing

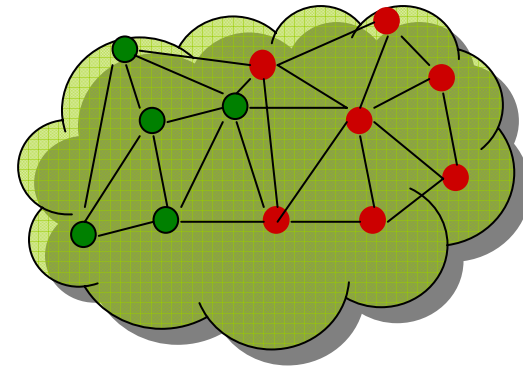
- Vertices are related to one another via a neighborhood system. A neighborhood consists of one or a group of sensor nodes, performing a common task or forming a homogeneous region.

- A neighborhood system is defined as:

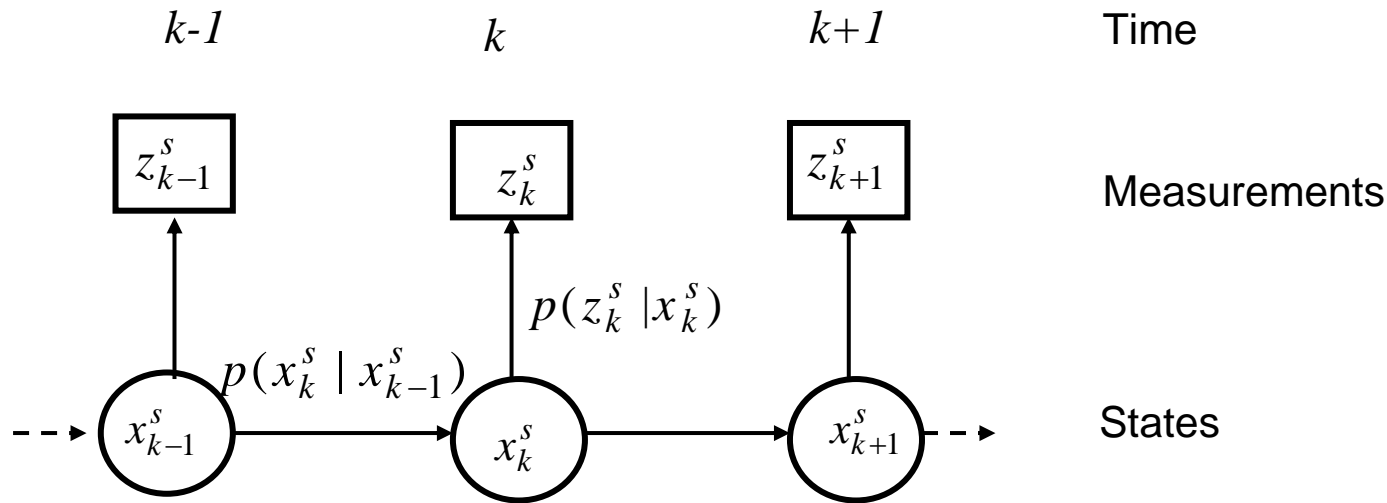
$$S = \{S_i \mid \forall i \in V\}$$

- Where S_i is the set of vertices neighboring vertex i , and vertex i is referred to as site of a neighborhood S_i .
- A vertex is not neighboring to itself, and Neighboring relationship is mutual

$$i \notin S_i \quad \text{and} \quad i \in S_{i'} \Leftrightarrow i' \in S_i$$



Dynamic System Model based on Bayesian Network



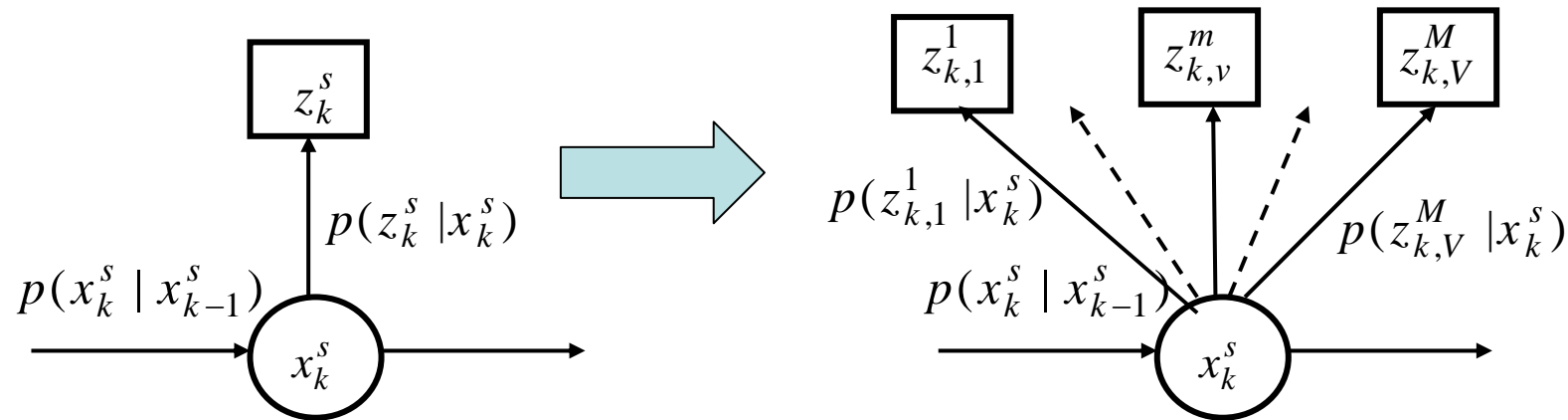
Prediction
$$p(x_k^s | z_{1:k-1}^s) = \int p(x_k^s | x_{k-1}^s) p(x_{k-1}^s | z_{1:k-1}^s) dx_{k-1}^s$$

Update:
$$p(x_k^s | z_{1:k}^s) = \frac{p(z_k^s | x_k^s) p(x_k^s | z_{1:k-1}^s)}{p(z_k^s | z_{1:k-1}^s)}$$

The recursion requires the specification of a dynamic model describing the state evolution, $p(x_k | x_{k-1})$, and a model that gives the likelihood of any state in the light of the current observation, $p(z_k | x_k)$.

Problem: only a concept solution; integrals are not tractable in most cases.

Composite Measurements



The composite measurement of the neighborhood S at time k can be modeled graphically as shown on the right: M – number of modalities, V – number of sensor sites

If measurements are conditionally independent each other

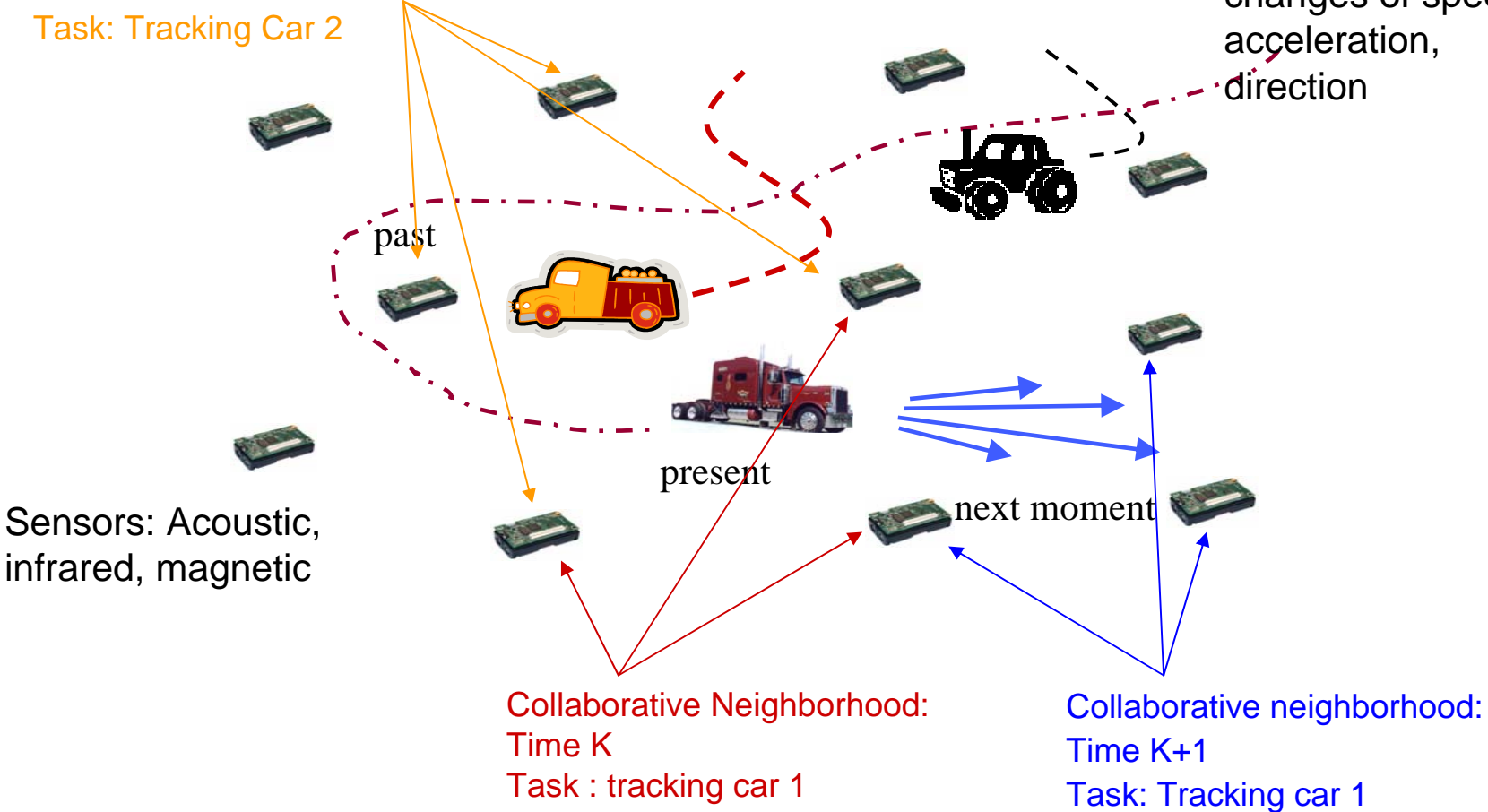
(given state). We have
$$p(z_k^s | x_k^s) = \prod_{v,m} p(z_{k,v}^m | x_k^s)$$

Quality of the measurements (noise, occlusion, etc) should be considered in the fusion process

Application 2: Tracking Multiple Targets

Collaborative Neighborhood:
Time K
Task: Tracking Car 2

Dynamic movement:
changes of speed,
acceleration,
direction



Collaborative Tracking in Sensor Networks

Cross-Sensor Fusion

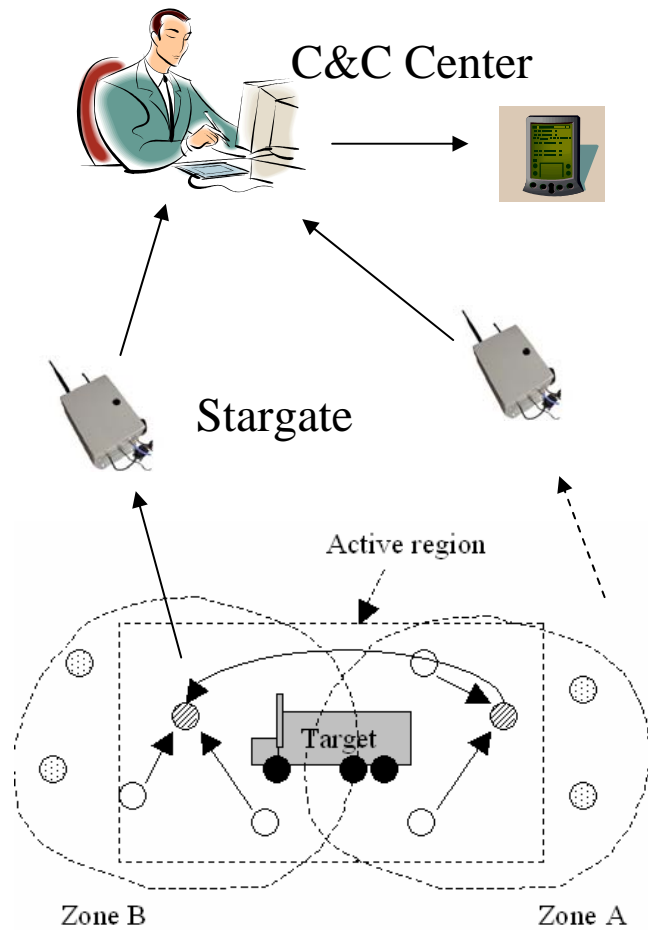
$$p(z^{aco} | x^i) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(GDS - z^{aco})^T (GDS - z^{aco})}{2\sigma^2}\right)$$

Cross-Modality Fusion

$$\begin{aligned} w_k^i &= p(z^{aco} | x_k^i) p(z^{sei} | x_k^i) \\ &= \exp\left(-\frac{c^{aco} e^{aco} + c^{sei} e^{sei}}{2\sigma^2}\right) \end{aligned}$$

- Monte-Carlo multi-modal data fusion.
- Agent-based Hierarchical tracking of multiple targets

Detection, Classification and Tracking



- Dynamic active sensor groups for collaborative tracking, and producing trajectory for each target
- Aggregation of tracking results and other related information to form a summary of behavior
- Distributed storage, management and utilization of data tracking results
- Alerting for suspicious behavior, sensitive regions
- Subscription for specific event (a particular car arrive at particular location)
- Query for summary of a time period, or a region

Thanks!