# A Letter Soup for the Quality of Information in Sensor Networks

<table>
<thead>
<tr>
<th>Chatschik Bisdikian</th>
<th>Kin K. Leung</th>
<th>Robert I. Young</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joel Branch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM Corporation</td>
<td>Electrical and Electronic</td>
<td>Defence Science and Technology Lab</td>
</tr>
<tr>
<td>T. J. Watson Research Center</td>
<td>Engineering</td>
<td>Sensor Systems Group</td>
</tr>
<tr>
<td></td>
<td>Imperial College, London</td>
<td></td>
</tr>
</tbody>
</table>
Overview

- Need for (sensory) information
- The Quality of Information (QoI) domain
  - World, states, QoI, VoI
- Sensor-enabled applications
  - Dynamic binding
  - Common Context of Operation (CCO)
- The 5WH principle
- Quick example
  - When: Spatial relevancy
  - Where: Temporal relevancy
- Summary
Why do we collect information?

- Information is gathered to build knowledge, and hence understanding of parts of the world that are of interest
  - …so appropriate actions can be decided upon

- The degree of understanding of the world depends on the degree of “pertinence” of the information gathered
  - …depends on the quality of the gathered information (QoI)

- Information is altered (for better or worse) along the path from generation to consumption
  - Information gathering, processing (fusion, storage, curation, retrieval), and dissemination
The domain of Quality of Information (QoI)

- The world comprises objects and linkages between these objects
- The “world” state vector (wsv) describes the world as a function of time
- Elements of the wsv are inferred using data sources
- Information (in our case) is the estimated value wsv elements

\[ x_1 = \{ \text{location, velocity, type, allegiance} \} \]
\[ x_2 = \{ \text{location, velocity, type, allegiance} \} \]
\[ x_3 = \{ \text{location, direction, strength, allegiance} \} \]

Information: estimate of object state \( x \)
QoI: collection of information attributes: accuracy, latency, provenance

Information gathering
Information processing
Information distribution
Information delivery

\( x_1 \) \( x_2 \) \( x_3 \) QoI?

sensor nodes
The domain of Quality of Information (QoI) (cont.)

- **QoI**: Measure of inherent characteristics of information
  - **Accuracy**: the spread and biases of the inferred elements of the wsv
  - **Latency**: the time difference between the arrival of the inference and time associated to the inferred wsv
  - **Provenance**: the history of measurements used for the inference

- **VoI**: Value of Information
  - The ability to take action on the information in light of its QoI
    - (sufficient) accuracy, timeliness, relevancy, completeness, usability, trustworthiness…

- Knowledge of the above allows us to
  - Pass judgment on the information regarding its fit-for-use
  - Make educated decisions and take actions based on this judgment
Sensor-enabled applications

- In “monolithic” deployments, sensor-generated information and sensor-enabled applications are statically bound and share
  - Common Context of Operation (CCO)
  - Applications have prior knowledge as to what their sensors are trained to observe, when, and where

- But…
  - Economies of scale improve as responsibilities split
  - Applications and sensor networks are deployed at their own pace
  - Dynamic binding

- How do we enable CCO with dynamic binding?
To establish CCO
- Information needs
- Temporal context
- Spatial context
  - Dictated by mission objectives

Establish CCO with who?
- A sensor information provider

Dynamic binding
- Bind to the semantically rich information flows from information providers
  - Not necessarily sensors
- Interpreted relatively to information needs
  - The information needs of a fusion element could be the raw sensor measurements

The 5WH principle

why

when

where

what

who

how

application needs

sensor information provider(s)

operational context

placeholder for future proofing!
The 5WH principle: A simple example

Why
Secure territory → requires e.g., monitor for suspicious activities
generate sensing tasks

Resolve to collection of atomic sensing operations, e.g., monitor for trucks of particular behavior, smoke, gunfire, etc.

Task/task needs

who  what  where  when

Agent that produce the “what” using the “how” resources

how
Sensing resources

For each task…

• Task 1: detection of intruders (NW)
  – When: 24 hours-a-day, within 30-seconds of an incident, until further notice
  – Where: NW area of camp
  – What: (detection features/information needs): intruder location, track, strength, type of vehicles, …
  – Who: A service (fusion resources) that can extract and deliver the aforementioned “what” features
  – How: Sensors (to be) deployed in support of the “who” service
The 5WH principle (cont.)

- Simple application agnostic framework for describing
  - Information consumption needs
  - Information production capabilities

- Basis for searching and exchanging information

- Enables dynamically binding applications to sensor information providers
Where: Spatial relevancy

- Information needs of an information consumer (application) pertain region \( R_c \)
- An information provider provides (related) information in region \( R_p \)
- The overlap of the regions can be used in a measure of spatial relevancy
  \[
  r_{\text{spat}} = \frac{f(R_p \cap R_c)}{f(R_c)}
  \]
  - \( f(.) \) some appropriate non-decreasing, non-negative function, e.g., the area of the region
- The expression can be generalized for “nice” vs. “good” to have requirements
  - Suppose that the information consumer would like information about region \( R_c^g \) (good to have) but it can still use information from the larger area \( R_c \) (nice to have)
    \[
    r_{\text{spat}} = \frac{\alpha f(R_p \cap R_c^g) + (1 - \alpha) f(R_p \cap (R_c - R_c^g))}{\alpha f(R_c^g) + (1 - \alpha) f(R_c - R_c^g)}
    \]
    - The expression can be generalized even further up to a continuum of good to have regions
Similarly, we can define temporal relevancy

- A mission needs information about an incident within time $T_1$ of present time while a (sensor information) service provider provides the information within time $T_2$.
- Alternatively, a mission may need information about what happened during a time period $(t_1, t_2)$, while the provider has information only for $(t_3, t_4)$.

Temporal relevancy can be measured on its own right:

- Or, in relation to the aging of the information $X$ desired:

$$r_{\text{temp}} = \frac{g((t_3, t_4) \cap (t_1, t_2))}{g((t_1, t_2))}, \quad g(\cdot) : \text{non-decreasing, non-negative}$$

$$r_{\text{temp}} = h\left( \hat{g}_X \{(t_1, t_2) | (t_3, t_4)\} \right), \quad h(\cdot) : \text{monotonic, decreasing: } 1 \rightarrow 0 \text{ (e.g., } e^{-x})$$

where: $\hat{g}_X \{(t_1, t_2) | (t_3, t_4)\}$ is the error in estimating $X$ over the interval $(t_1, t_2)$ given knowledge of it in $(t_3, t_4)$. 
Summary

- Discussed the reason for information gathering
  - Gain knowledge / situation awareness
- Introduced QoI and VoI
  - Enables the process of judging information
    - Inherent properties of information
  - Taking actions based on the information
    - Role of a piece of information within its context of use
- Referred to the challenge of dynamically deployed applications
  - The need for establishing CCO
- Proposed the 5WH principle for information summarization
- Discussed Where and When and spatiotemporal relevancy
THANK YOU!