Innovation Crossover Preliminary Research Report
DoD Technologies – Sensor-Data Fusion

Context/Scope

This paper represents research conducted by OVO Innovation for the NSWC Crane Innovation Crossover event October 12-13, 2016. This research is intended to provide more insight into key challenges that were identified within the four technology clusters (Advanced Manufacturing, Cyber/IT, Life Sciences and DoD Technologies) first documented in the Battelle report. OVO consultants interviewed subject matter experts (SMEs) from the private sector, academia and the government identified by NSWC Crane to gather insights into key challenges in each cluster. This report is meant to inform the participants of the Innovation Crossover event and identify new research and new technologies that might address the key challenges.

This research was collected during August and September, 2016. The reports were submitted by OVO to NSWC Crane in late September 2016.

Introductory Narrative

The Innovation Crossover event, scheduled for 12-13 October 2016 in Bloomington is the culmination of months of planning and hard work. Some of this preparatory work involved the initial Battelle study which identified key technology clusters (Advanced Manufacturing, Life Sciences, Cyber/IT and DoD Technologies) in southern Indiana. From these clusters NSWC Crane and its contractor OVO Innovation conducted further, more detailed research, to examine detailed challenges and opportunities in each technology cluster. The reports attached document the research OVO conducted with subject matter experts identified by NSWC Crane in academia, industry and in the government. The reports are meant to document specific challenges within each technology cluster that could become areas of joint research and cooperation across the three constituents in southern Indiana. The reports are provided to you to help you prepare for your participation in the upcoming Innovation Crossover event and to frame both the challenges and active research underway to address these challenges.

Problem Statement

DoD Technologies: Sensor-Data Fusion

- Data fusion techniques combine data from multiple sensors and related information from associated databases to achieve improved accuracy and more specific inferences than could be achieved by the use of a single sensor alone. Briefly, we can define data fusion as a combination of multiple sources to obtain improved information; in this context, improved information means less expensive, higher quality, or more relevant information. Data fusion techniques have been extensively employed on multi-sensor environments with the aim of fusing and aggregating data from different sensors. *The goal of using data fusion in multi-sensor environments is to obtain a lower detection error probability and a higher reliability by using data from multiple distributed sources.*
- Cognitive, Distributed, Active and Passive Electromagnetic Sensing with Cooperative Electronic Attack. Includes High Fidelity Computational/EM (CEM) Modeling. Specific need: Sensors and track data fusion.

Problem Context

- Problem or Challenge:
  - Millions of sensors deployed worldwide on many platforms
  - Sensors provide a consistent stream of data; many different formats (IR, Acoustic, Radio, Optic, etc)
  - Challenge to aggregate and normalize all of the data
  - Challenge to rationalize the data and convert into information or knowledge
  - Challenge to provide meaningful information to propose an action by a human or autonomous agent
- How do we gather and use all of the sensor output, interpret it and synthesize it effectively so that it creates meaningful knowledge that informs timely actions?

Problem Context

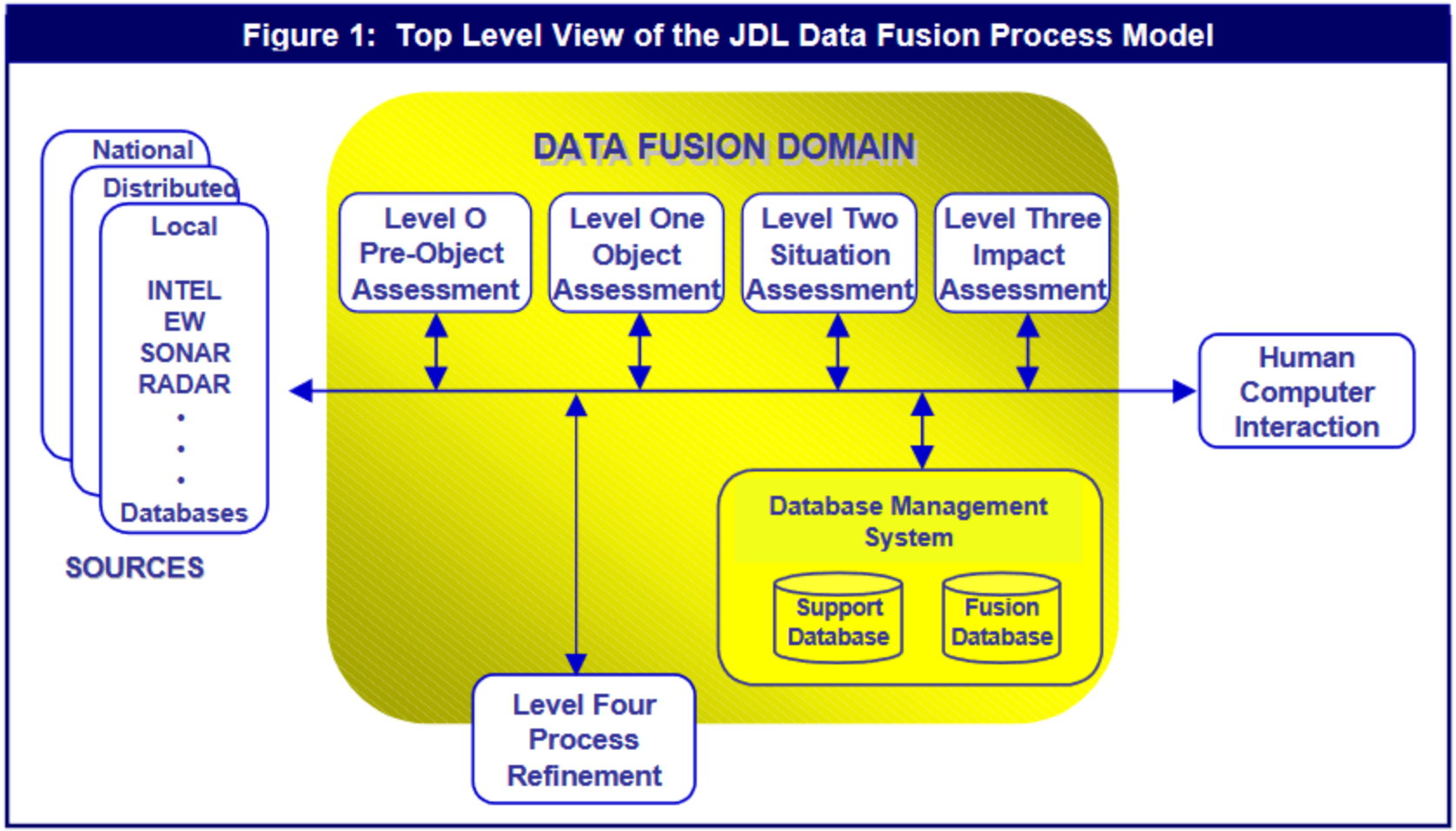
Problems within this challenge are numerous:

- Much of the data generated by sensors is siloed (by type, by owner, by geography)
- Relevance of the source
- Type of data
- Volume of data
- Fuzziness/consistency of data
- Geolocation of data (static/mobile sensors)
- Differing protocols
- Consistency/availability of data
- Normalization
- Teaching the algorithms to interpret the data
- Converting data into actionable information

Technologies

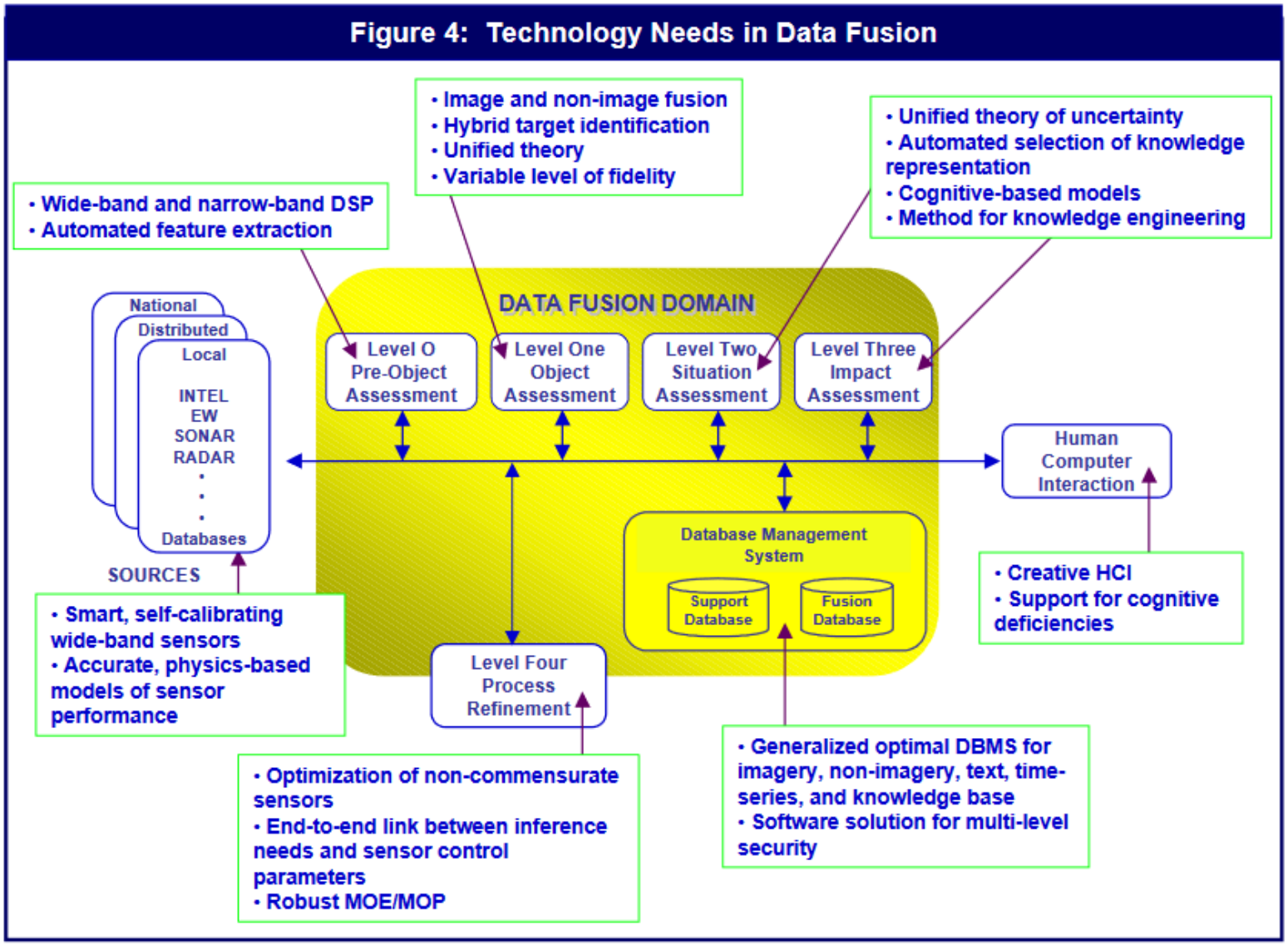
- Several technology categories contribute to this opportunity or challenge:
  - Sensors: fixed and mobile, deployed in every setting
  - Data streams: different data formats, different periodicity, different relevance. Also a cybersecurity problem.
  - Normalization/Synthesis: Big data problem
  - Learning: a machine learning/algorithm problem
  - Analysis leading to knowledge: finding information in the data, and using the knowledge to gain an advantage

Process Model



From: Dirty secrets in multisensory data fusion, http://www.dtic.mil/dtic/tr/fulltext/u2/a392879.pdf

Technology Needs



Problem Relevance

- The sensor-data fusion problem has relevance for both military and civilian entities.
- In the civilian space, the challenge occurs in situations like autonomous vehicles, which must gather data from a range of sensors and make thousands of decisions in near real time
- In the military sector, the challenge is to obtain information from widely dispersed sensors to gain an advantage in intelligence or combat and to create the best possible outcomes based on the information we can extract and process

Problem Relevance

- The challenge only grows more acute in both the civilian sector and the military sector as the Internet of Things (IoT) increases exponentially. Thousands of sensors and data streams exist, generated from millions or even billions of sensors and other technology.
- In some cases we have people or technology go into harm's way to install or deploy a sensor – the data is not always "free"
- Yet as the range and amount of data generated increases, we don't have good mechanisms to gather the data, normalize it, assess it and convert it to information and knowledge
- The data is also subject to cyberattack. If enemies or opponents hack the sensor or change the data stream, we could be left analyzing data that corrupts the information and causes us or autonomous devices to make bad decisions

Scope

- For the purposes of this research, we've defined the scope of the problem as:
  - Sensors
  - The data streams they generate
  - Aggregating and normalizing the data
  - Managing and processing the data, converting it into information and knowledge
  - Ensuring the data streams are reliable and accurate

Implementation Challenges

- There's no substitute for a good sensor
- Downstream processing can't make up for errors in upstream processing
- Sensor fusion can result in poor performance if incorrect information about sensor performance is used
- There is no such thing as a magic or golden data fusion algorithm
- There will never be enough training data
- It is difficult to quantify the value of a data fusion system
- Fusion is not a static process

From: Dirty secrets in multisensory data fusion, http://www.dtic.mil/dtic/tr/fulltext/u2/a392879.pdf

New Research

- There are a number of groups doing research into sensor-data fusion
- These include academic researchers, government agencies and industry

Research

- Government / Military Research
  - NSWC Crane conducting research into soft computing led by Dr. Bob Cruise. James Stuart leading a team as well.
    - NSWC Crane hoping to construct a data fusion / sensor fusion test bed
  - AFRL doing research – classified, so we weren't able to learn much

Research

University Research

- Indiana University, decision model research led by Jerome Busemeyer
- Purdue has active research led by Dr. Ed Delp, Dr. Carolyn Frueh, Dr. DeLaurentis and Dr. David Ebert
- Virginia Tech is conducting research, as is Rice University and Carnegie-Mellon
- IUPUI conducting research, led by Dr. Paul Salama

Research

- Industry research potential
  - Raytheon and Harris are building algorithms to address data fusion

Summary

- Sensor-Data Fusion is an important challenge and its complexity will only increase as the number of sensors increases and the Internet of Things (IoT) becomes a reality
- We cannot adequately manage and interpret the amount of information we currently receive, much less scale to the amount of information that will be available
- While there is a lot of information, we must also understand its value. We may not have sensors in areas where we need data, and have too many sensors or information sources that are duplicative

Summary

- Beyond the wealth of data generated by sensors, we have a significant normalization challenge, because these sensors track different kinds of information, at different time intervals and in some cases are periodic (satellites, sensors on drones)
- Differing data formats, different time intervals and incomplete coverage make normalizing the data streams difficult
- Even after the normalization of the data, we have to be able to analyze and interpret the data appropriately, which becomes a “big data” problem, to find information or knowledge in the data

Summary

- The analysis and insight impacts both the civilian world, as the Internet of Things and concepts like autonomous vehicles become a reality, as well as the military and governmental sectors, as we seek to keep our vigilance and our ability to respond at a high level
- Further, Sensor Data Fusion issues also overlap with Cybersecurity and Machine learning topics, because we must be able to trust the data sent by a sensor and the data we analyze. Competitors could provide false streams of data to confuse our analysis.

Sources

Subject Matter Experts consulted / interviewed

- Greg Reece, NSWC Crane
- Kyle Werner, NSWC Crane
- Rob Walker, NSWC Crane
- Dr. Dan DeLaurentis, Purdue
- Dr. Razi Nalim, IUPUI
- Bill Harrison, AFRL

Sources

- Introduction to multisensor data fusion techniques
- Dirty secrets in multisensory data fusion
- Soft Computing Approaches for Information Fusion and Control Diffusion