Total Maritime Domain Awareness

an evolving maritime security solution

A Larus Technologies White Paper
March 2013

Abstract

Maritime domain awareness is the understanding of activities that impact maritime security, safety, economy or environment. It enables quick threat identification, informed decision making, effective action support and knowledge sharing. The core component of maritime domain awareness, persistent surveillance, involves multiple systems corroborating in order to detect, classify, identify, track and assess situations within an area of interest. In this white paper, we concentrate on the challenges presented in applying the concepts of information fusion to the problem of maritime domain awareness, and how we plan to resolve them. We first introduce the conventional techniques and their drawbacks, discuss the contemporary data ecosystem and present a potential solution which learns to closely match the dynamic internal structures present in the data. This solution, developed and patented by Larus Technologies, performs behavior analysis through predictive modeling, is capable of dealing with heterogeneous (i.e. multi-source, multi-sensor) data, is automated yet human-centric and resolves many of the challenges presented by maritime domain awareness.

Maritime Domain Awareness

In a world where more than 40% of the population lives within 100 kilometers of a coast [1] and where traditional and asymmetric threats to physical and cyber infrastructures and borders continue to rise each year, countries are becoming increasingly aware of the gaps that exist in their ability to achieve persistent surveillance and continuous awareness of their maritime domains. Persistent surveillance is an essential component in a global system to ensure *Territorial Security*. The latter being defined as the prevention, detection and response to unauthorized persons and/or goods crossing a physical or virtual perimeter; this problem has recently become a security concern of individual, corporate, and international scope.

In a vast and mostly uninhabited country such as Canada, which borders the Atlantic, Pacific and Arctic oceans, a major component of Territorial Security is Maritime Domain Awareness (MDA), which provides awareness of potential threats from maritime approaches and cueing of military and interagency responders. MDA is defined as the situational understanding of maritime activities that could impact the security, safety, economy or environment [2]. MDA involves people processes and technological tools that together contribute to physical and virtual defences of the country's borders. MDA includes national, provincial and municipal services and organizations that act based on defence strategy and the government policies and procedures. To be effective, MDA also includes systems that ensure domain knowledge is captured and stored for handling future scenarios when they arise.

Typically, multiple loosely connected maritime security systems have been used to patrol and monitor maritime areas of strategic importance, however, a number of challenges exist in this disjointed architecture. First, intra-connected (i.e. linking the sensors that make up a security system) and inter-connected (i.e. linking the security systems themselves) are both inflexible and expensive to setup while not being interoperable from the start (i.e. knowledge sharing between authorized users and systems should be a design objective). Additionally, and more importantly, operators and analysts are being overwhelmed by the tide of incoming data, including sensor outputs, databases, reports and other sources of information. This situation typically leads to operator/analyst fatigue, overload, stress and inattention which, in turn, lead to human errors within the MDA process. State-of-the-art security solutions have been effective in limited scenarios, where the regions of interest were well delineated, the data sources were structured and precise, the events of interest were few and far between and the response was neither time-critical nor calculated. Hence, any proposed solution to these challenges will need to feature constant surveillance of the environment unconstrained by data parameters or geographical boundaries, i.e. persistent surveillance.

Persistent Surveillance

An effective persistent surveillance system of systems (SoS) includes multiple collection, exploitation and dissemination systems that are controlled cooperatively to detect, classify, identify, track, corroborate and assess situations within an area of interest (AOI). A persistent surveillance SoS provides the decision and/or policy maker with a range of information and intelligence products to inform decision making that enables effective mitigation of potential threats and timely response to actual territorial breaches. Persistent surveillance SoS provide three significant benefits:

- **Expertise benefits**: Low-risk points-of-sensing do not require highly trained personnel for monitoring. Technological solutions provide the ability to have an autonomous system cover the low-risk points without the need for highly trained personnel;
- **Remote monitoring benefits**: The use of large numbers of unattended sensors for perimeter security introduces a significant remote monitoring problem. Technological solutions provide the ability to remotely monitor large numbers of geographically dispersed sensors; and
- **Resource benefits**: Critical infrastructure protection in large geographically condensed areas typically suffers from the lack of resources for monitoring and management. Technological solutions provide the ability to secure these physical and/or virtual infrastructures with autonomous monitoring that does not require human management.

Looking more closely at the national maritime surveillance domain, the Arctic region has been one of much discussion in the past few years. As Canada is the chair of the Arctic Council[1] in 2013, it becomes imperative for us to set the stage for advancing our Arctic foreign policy and promoting Canadian Northern interests [3]. Amongst the many priorities that are articulated in Canada's Northern Strategy, MDA supports: (i) securing international recognition for the full extent of our continental shelf and (ii) addressing Arctic governance and related emerging issues, such as public safety. Finally, MDA solutions need to enhance the Canadian Forces (CF) ability to conduct surveillance in the North through the replacement of the Aurora patrol aircraft, the development of unmanned aerial vehicle (UAV)

[1] The Arctic Council is a high-level intergovernmental forum to promote cooperation, coordination and interaction among the Arctic States (http://www.arctic-council.org/index.php/en).

platforms, and the Polar Epsilon capability [4]. The latter is a space-based Canadian radar system that augments surveillance of Canada's Arctic and maritime approaches.

The Polar Epsilon capability has provided Canada with all-weather, day/night persistent surveillance of its Arctic region and ocean approaches. The capability includes ship detection, oil detection, environmental sensing, change detection and oceanic intelligence [5]. Within the Polar Epsilon context, MDA was sequentially defined as Detect → Classify → Identify → Track → Intent. Future efforts in this direction will concentrate on the "Intent" phase, including the development of additional exploitation and assessment capabilities, as well as better utilization of the upcoming RADARSAT Constellation Mission (RCM) which is scheduled for launch in 2018 and will initially include three satellites with capacity to support up to six satellites within the constellation (see Figure 1). The RCM's three main uses will be maritime surveillance, disaster management and ecosystem monitoring. RCM recently received Government approval to proceed to its next and final stage of development and deployment [6].

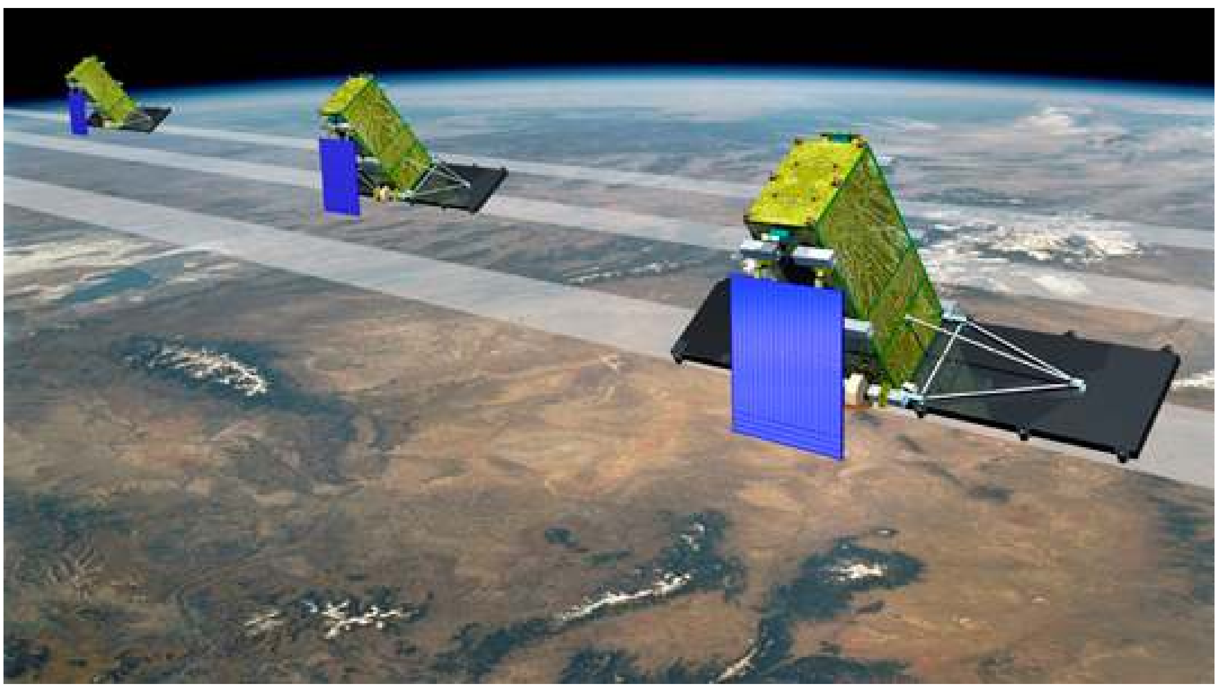

Figure 1. RCM's three satellites (Credit: MacDonald, Dettwiler and Associates Ltd.)

With the RADARSAT Constellation Mission (RCM) satellites now a high priority for the Canadian Space Agency's (CSA), we can assume that maritime persistent surveillance sensing capabilities will be enhanced in the coming years. In Canada, this means surveying 10 million km^2 across the Pacific, Atlantic and Arctic oceans, over 200 thousand km of coastline and 5 million km^2 of Arctic landmass (refer to Figure 2) and the inherent challenge of monitoring and controlling the vast amount of data and information that will be generated.

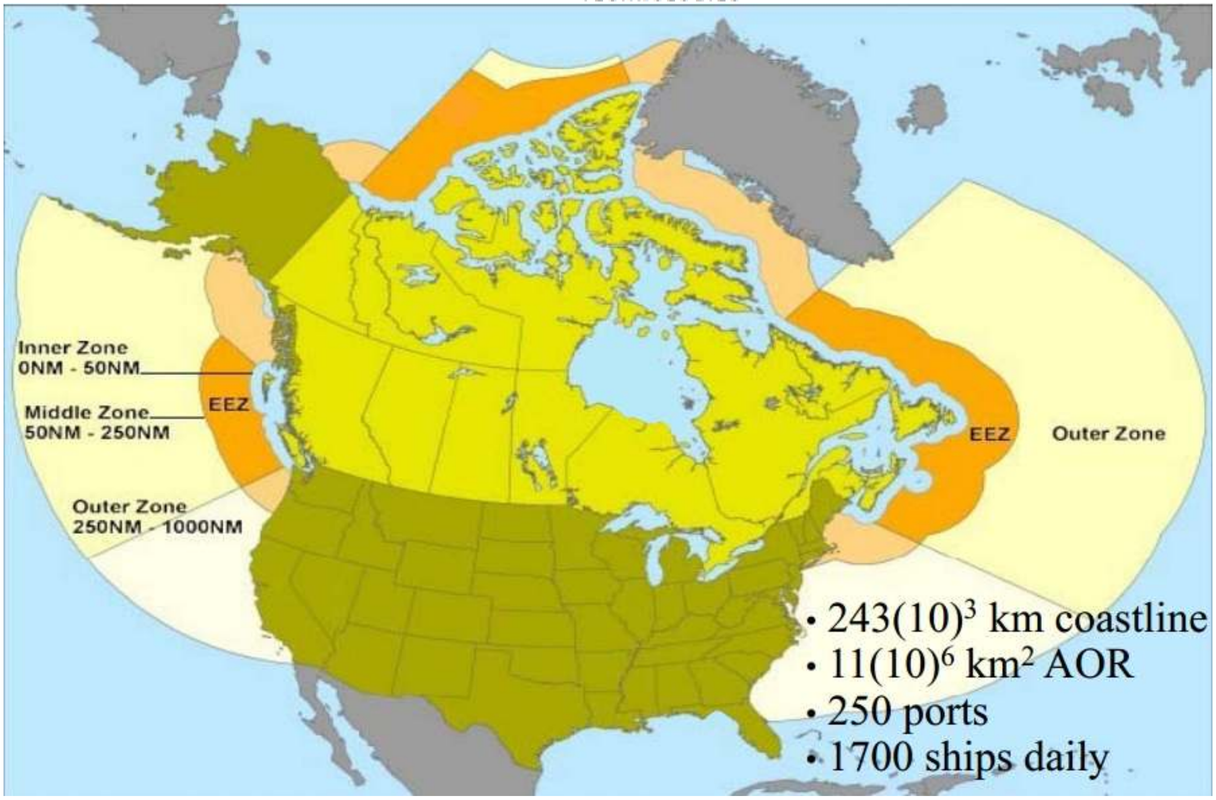

Figure 2. Canada's areas of responsibility and surveillance zones (extracted from [7])

In Canada, MDA is within the jurisdiction of the Marine Security Operations Centres (MSOCs) and three CF Regional Joint Operations Centres (RJOCs), i.e. RJOC Atlantic, RJOC Pacific and RJOC Northern, which are responsible to detect and assess Canadian marine security threats and provide support to responders. Threats include individuals, vessels, cargo and infrastructure performing any activity that could pose an injury to the safety, security, environment or economy of Canada.

Data Sources

There are many data sources within MDA. These sources are divided into two major categories with varying nomenclatures: structured vs. unstructured, hard vs. soft, sensed vs. unsensed, etc. The first set of names indicate calibrated, precise and structured ("hard") data such as imagery and radar sensor data, while the second set of names indicate uncalibrated, imprecise and unstructured ("soft") data such as operator reports and open source intelligence available from internet web pages.

Hard data typically has a high observational sampling rate, is easily repeatable and provides attributes for single sub-objects or objects. The structure that resides within hard data allows system integrators to easily interface to such data sources and extract the known features to perform further data processing. Examples of hard MDA data sources include, but are not limited to:

- Radar-based (e.g. Synthetic Aperture Radar – SAR, Automatic Radar Plotting Aid – ARPA);
- Tracking-based (e.g. Ground Moving Target Indicator – GMTI, LINK 11/16/22, Over The Horizon – OTH-Gold);
- Contact-based (e.g. Automatic Identification System – AIS, Global Positioning System – GPS, National Marine Electronics Association 0183 – NMEA 0183);
- Electro-Optical-based (e.g. day/night cameras, thermal sensors, infrared cameras);
- Environmental-based (e.g. temperature, humidity, pressure, precipitation, dew, smoke);
- Ranging-based (e.g. sonar, Light Detection and Ranging (LIDAR), laser);
- Orientation-based (e.g. magnetic compass, gyroscope); and
- Ontology-based (e.g. Wikipedia, Linking Open Data Project [8]).

Soft data typically has a low observational sampling rate, is not easily repeatable and hence is less precise and provides relations between discovered entities. The lack of structure forces system integrators to develop techniques for feature extraction and data source ingestion. Examples of soft MDA data sources include, but are not limited to:

- Weather-based (e.g. weather reports, weather patterns);
- Human observation-based (e.g. field reports, interviews, intelligence reports, logs);
- Map-based (e.g. navigational charts, climate maps);
- Web-based (e.g. web sites/pages, forums, RSS feeds); and
- Social-based (e.g. Facebook pages, Twitter feeds, personal blogs).

Other issues with soft data sources, including source and report credibility, handling of uncertainty, natural language processing, need to be better defined. Additionally, issues that need to be resolved are; (i) fusion point delineation (i.e. where is the best layer for hard and soft data sources to be integrated?), (ii) how to modify models as new strategies are required for the orientation, observation and decision phases of the decision-support system, and (iii) how to best perform contextual information extraction and integration as soft data sources typically contain limited inferential knowledge. For these reasons and more, hard-soft fusion has become a hot topic of research. Solutions to hard-soft fusion have included the introduction of soft data exploitation within existing hard data fusion systems and novel paradigms that attempt to start out with separate hard and soft streams at the ingestion point that move towards a harmonized situational understanding as the model gains situational experience through its real-world embodiment.

Information Fusion

In order to accurately and effectively monitor an AOI, the tide of incoming data must be interpreted and properly managed. Often referred to as the "Big Data Problem", this situation is best handled through the creation and maintenance of a real-time representative model of the world. Contemporary solutions have attempted to resolve this challenge through complex mathematical formulations or brute force number crunching, however, these solutions are not adequate because the 4-dimensional vector (variety, volume, velocity and veracity) representing this type of data changes too quickly for such approaches to remain relevant.

Over the years, researchers and organizations have attempted to tackle the Big Data problem by trying to keep up with the growing and changing data. This used to be a manageable solution approach as the data itself was limited in volume, involved few types (low variety), did not frequently change in mission-critical applications (low speed) and was somewhat trustworthy (high veracity). Hence, solutions that involved low level Information Fusion (IF) modules were quite capable in adapting to the changing structures within the incoming data streams. Today's data, however, can be expressed in terabytes when it comes to its size, in millions per second when it comes to speed, in tens, if not hundreds of types when it comes to diversity and in jams and interferences per second when it comes to trustworthiness. This renders low level IF-based solutions no longer capable of coping with dynamic behavior (whether accidental or intentional) of today's datasets meaning that a new computational paradigm is required.

To address these aforementioned challenges, *High-Level Information Fusion* (i.e. Level 2 and above in Figure 3), or HLIF as it is better known, has become the focus of contemporary research and development efforts. HLIF uses a mixture of numeric and symbolic reasoning techniques running in a distributed fashion, over a secure underlying backbone while presenting its internals through an efficient user interface. HLIF allows the system to learn from experience, capture human expertise and guidance, analyze contextually and semantically, lower computational complexity, automatically adapt to changing threats and situations, and display graphically inferential chains and fusion processes.

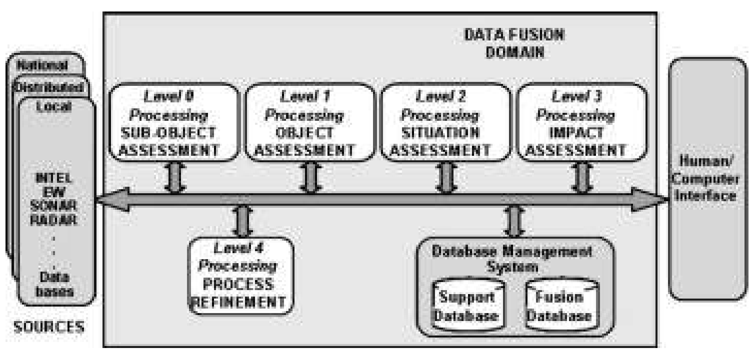

Figure 3. Information Fusion process (extracted from [9])

Instead of attempting to keep up with the ever increasing complexity of the 4-dimensional data streams, HLIF allows one to better understand (i.e. model) the source of those streams, and therefore better adapt to the dynamic structures that exist within the data. Let us now take a look at some algorithms, based on Computational Intelligence (CI), that greatly improve the operation of HLIF systems.

Computational Intelligence

Computational Intelligence involves the design of computational architectures, methodologies and processes to address complex real-world problems using nature-inspired approaches. There are three main divisions within the CI domain, namely Neural Networks (NNs), Evolutionary Computation (EC) and Fuzzy Systems (FS), with a few more emerging trends. Each is described in more detail in the following sections.

Neural Networks

The first theory on the fundamentals of neural computing was published by W. McCulloch & W. Pitts [10] in 1943 which described an all-or-none threshold device that made up the basic processing unit called a *neuron*. When a collection of neurons is connected via weighted links, the result is a Neural Network (NN), where the activity of one neuron can be amplified or reduced and summed with the activity of other neurons to affect the behavior of yet another. NNs replaced the centrally executed, symbolic logical system of artificial intelligence (AI) and offered distributed processing based on sub-symbolic continuous activation levels. See Figure 4 for a depiction of one typical neuron as well as the generic architecture of a NN consisting of three layers: one input, (at least one) hidden and one output layer.

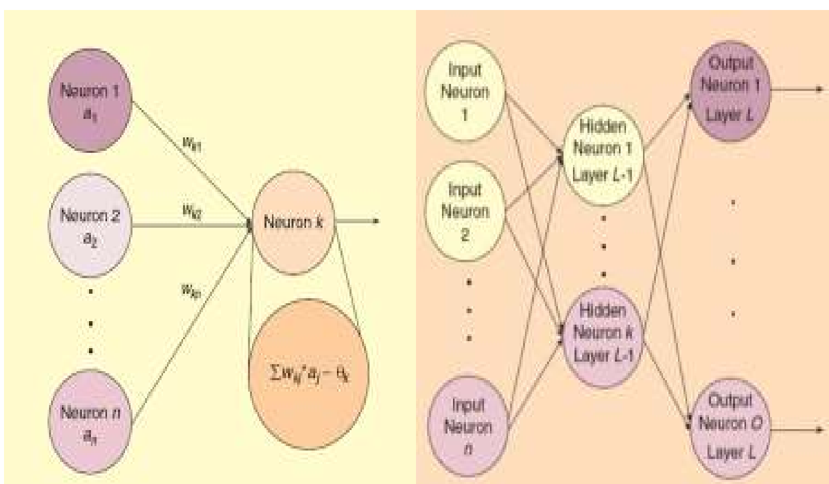

Figure 4. Neural network generic architecture

There are many types of NNs that have been devised over the years; some of the most popular and useful ones include feed-forward networks, such as the Multilayer Perception (MLP) networks, where information flow is strictly unidirectional and recurrent networks, such as the Hopfield and NARX networks, where information is allowed to feed back to nodes in earlier layers of processing.

Evolutionary Computation

NNs were found to perform successful distributed processing; however, information flow between the subcomponents was completely fixed and predetermined by the network topology. Along came evolutionary algorithms, loosely based on the interpreted operation of natural evolution, which essentially represented a distributed system of simple agents with no a priori designed communication flow pattern. Evolutionary Computation (EC) became the field of investigation in the 1960s into all evolutionary algorithms (EAs), including Evolution Strategies (ES), Evolutionary Programming (EP), Genetic Algorithms (GA) and Genetic Programming (GP).

Contingent on agents constructing new hypotheses about a solution to the problem, EC uses a random variation and recombination of the information about the old/previous hypothesis and performance-related evolutionary pressure which is biased towards retaining better hypotheses in the next cycle/generation of operation. EC is typically applied to problems where heuristic solutions are not available or generally lead to unsatisfactory results, where, through iterations of random variation and selection, the population can be made to converge asymptotically to optimal solutions (derived from schemata theory).

Genetic Algorithms (GAs) are search techniques modeled after natural selection, including the associated genetic operators and were developed by John Holland at the University of Michigan in the early 1970s [11]. GAs are stochastic algorithms with very simple operators that involve random number generation, and copying and exchanging string structures. The three major operators are: selection, mutation and crossover, with fitness evaluation acting as a control factor in the feedback path [12]. GAs fare well in large search space problems because better solutions tend to "grow old with time". See Figure 5 for a depiction of a GA process as well as a pictorial of the *genetic pipeline* present at the heart of every GA.

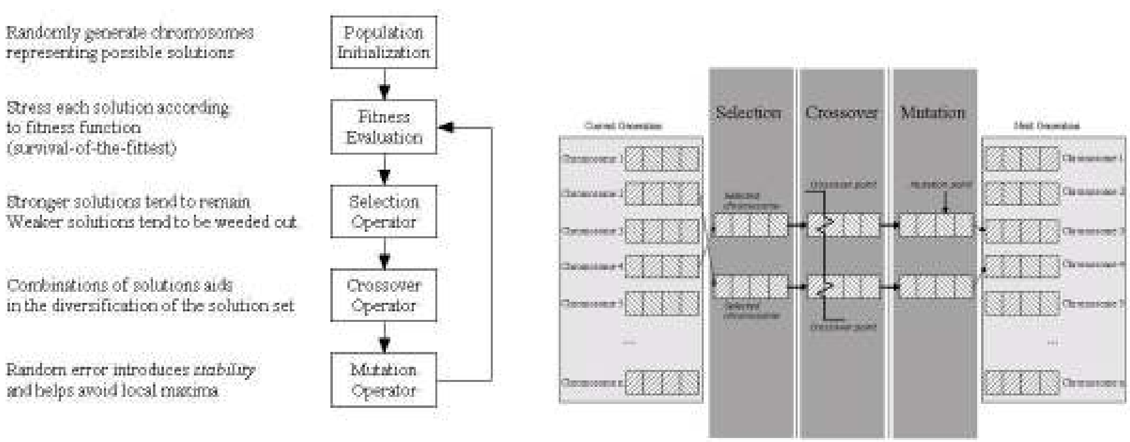

Figure 5. Genetic Algorithm process flow and the genetic pipeline

Fuzzy Systems

The mathematic notion of fuzzy sets was introduced by Lotfi Zadeh [13] in 1965 based on the concept of imprecision. Instead of presenting precise rules or instructions, the system is guided by *fuzzy rules* that describe tasks more easily, such as, "when you are close to the door, open it". This became the foundation of fuzzy computation which stipulated that the interaction between computers and humans can be greatly facilitated by the use of words. Fuzzy Systems (FS) or Fuzzy Inference Systems (FIS) became the physical manifestations of fuzzy computation.

By crafting rules or describing data in terms that are easily understood, a system designer can simplify the design of a very complex system Measurements need only be described using fuzzy terms such as “very often” or “quite high” while membership functions can be intricately designed for fuzzification of crisp inputs. The defuzzification of fuzzy output variables into crisp values uses methods such as the center of gravity or mean of maxima methods. See Figure 6 for the typical process flow of a FS as well as a sample membership function which represents the degree of truth of an element to a particular fuzzy set. For example, a value of 0 indicates that the element does not belong to the fuzzy set, a value of 1 indicates that the element fully belongs to the fuzzy set, and a value in between indicates that the element partially belongs to the fuzzy set. This powerful concept aids in the processing of imprecise data in order to arrive at adaptive, yet rigorous, systems that yield human-assisted and interpretable solutions.

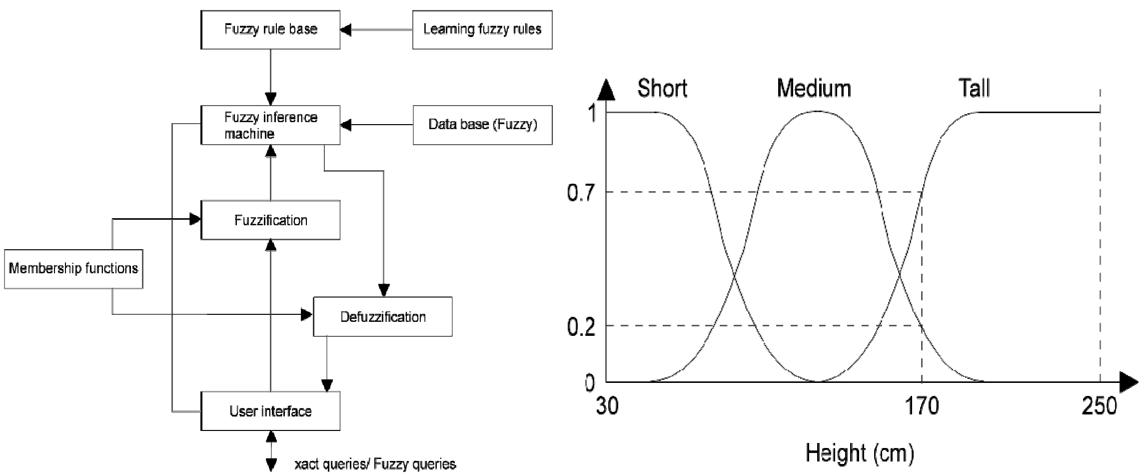

Figure 6. Fuzzy System process flow and sample membership function

Distributed Algorithms

Recently, Distributed Artificial Intelligence (DAI) has become viable for certain applications due to numerous reasons, including:

- It is costly to spend all your efforts on one entity;
- Problems are physically distributed;
- Problems are complex and require local points of view; and
- Systems must be able to adapt to environmental changes.

An example of a distributed algorithm is swarm intelligence [14], which includes ant algorithms, Particle Swarm Optimization (PSO) and diffusion search. The method is based on the operation of a population of simple agents, each of which explores the space possible solutions, until an overall solution emerges from the interactions between the agents.

Behavior-based networks respond to stimuli through a successive activation of a number of nodes in the network, mapping the activation set to a behavior through learned association. They can represent states such as an association in memory between two stimuli or a reaction to external stimuli.

Multi-agent systems (MAS) [15] consist of individual agents coordinate their activities and cooperate with each other to avoid duplication of effort as well as exploit other agents' capabilities. They are typically applied to many areas including spacecraft control, social simulations, ecommerce and industrial systems management. Distributed sensing and sensor networks are a major application area of multi-agent systems [16].

Finally, hierarchical networks consist of probabilistic learning networks that are used to deal with problems of uncertainty and complexity. These complex systems are typically built by combining simpler parts. Examples include Bayesian networks and Hidden Markov Models.

Finally, it is important to mention that there are two ways to extract regularities from presented patterns, namely (i) *supervised learning*, where networks are provided with quantitative information on their performance, the latter being used to adjust the weights to achieve better performance and (ii) *unsupervised learning*, where no provision of feedback is provided to the network and the process is mostly based on an appropriately defined cost function which uses local interactions between the processing elements to arrive at a desired solution.

HLIF Capabilities

As previously mentioned, HLIF deals with Level 2 and up (refer to Figure 3) and has become the focus of contemporary research and development efforts in order reduce the stress on operators/analysts and the burden being placed on computational systems dealing with Big Data streams.

HLIF capabilities are continuing to evolve to alleviate the challenges presented by today's data ecosystem. These include *(i) anomaly detection*, a process by which patterns are detected in a given dataset that do not conform to a pre-defined typical behavior (e.g. outliers), *(ii) trajectory prediction*, a process by which future positions (i.e. states) and motions (i.e. trajectories) of an object are estimated, (iii) *intent assessment*, a process by which object behaviors are characterized based on their purpose of action, and *(iv) threat assessment*, a process by which object behaviors are characterized based on the object's capability, opportunity and intent.

Additionally, real-time adaptive learning becomes an imperative feature of any MDA solution deployed in the field. *Situational learning* (shaping future responses to already seen situations based on human feedback) and *procedural learning* (minimizing the error between predicted and actual events) are two methods that enable a system to better understand real-world dynamics.

Larus Technologies has developed a patent-pending HLIF architecture that performs behavior analysis through predictive modeling, is capable of dealing with heterogeneous (i.e. multi-source, multi-sensor) data, is mostly automated yet human-centric and is mainly targeted as a MDA solution. It combines CI algorithms situated within a persistent surveillance environment in order to classify, identify, track, assess and support decision makers. This risk-aware decision-support system (DSS) provides the aforementioned HLIF capabilities while alleviating the strain on the operators and analysts by reducing the influx of information to manageable levels. For more information on the DSS, the reader is pointed to the online article entitled "Why High-Level Information Fusion?" also authored by Larus Technologies (http://www.larus.com/knowledge/why-high-level-information-fusion).

Conclusions

Maritime Domain Awareness is a complex and involved process meant to provide true and timely information concerning maritime activity to decision makers. Persistent surveillance involves multiple systems collaborating in order to detect, classify, identify, track and assess situations within an area of interest. Larus' patented HLIF architecture combines Computational Intelligence algorithms and behavior analysis, incorporating a multitude of hard and soft data sources, to aid in persistent surveillance for the purposes of MDA. The solution eases operator overload, provides a more accurate and reliable world model, offers interoperability from the start as well as delivers the requisite HLIF capabilities.

References

[1] NASA Socioeconomic Data and Applications Center (SEDAC). "Percentage of Total Population Living in Coastal Areas." http://sedac.ciesin.columbia.edu/es/papers/Coastal_Zone_Pop_Method.pdf, New York, NY, USA. 25 July 2007.

[2] International Maritime Organization, Amendments to the International Aeronautical and Maritime Search and Rescue (IAMSAR) Manual, pp. 1, last accessed on February 24, 2013.

[3] Government of Canada. "Canada's Northern Strategy: Our North, Our Heritage, Our Future", 2009.

[4] A., Clive, "Vision for the Arctic," FrontLine Security, 6(1), June 2011.

[5] R.J. Quinn, "Project Polar Epsilon: Joint space-based wide area surveillance and support capability", Proc. 8th International Conference on Remote Sensing for Marine and Coastal Environments, 17-19 May, 2005, Halifax, Nova Scotia, Canada, 2005.

[6] Canadian Space Agency, "The Government of Canada Launches Final Stage of RADARSAT Constellation Project", last accessed on February 23, 2013.

[7] P.W. Vachon, "New RADARSAT Capabilities Improve Maritime Surveillance", October 2010.

[8] Linking Open Data Project, http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData, last accessed on February 23, 2013

[9] D. A. Lambert, "A blueprint for higher-level fusion systems," Information Fusion, vol. 10, pp. 6-24, January 2009.

[10] W. McCulloch, and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophysics*, vol. 5, pp. 115–137, 1943.

[11] J. Holland, Adaptation in Natural and Artificial Systems, The University of Michigan Press, Ann Arbor, Michigan, 1975.

[12] D.E. Goldberg, *Genetic Algorithms in Search, Optimisation and Machine Learning.* Reading, MA: Addison-Wesley, 1989.

[13] L.A. Zadeh, "Fuzzy sets," *Inform. Control*, vol. 8, no. 3, pp. 338–352, 1965.

[14] J. Kennedy and R.C. Eberhart, *Swarm Intelligence*. San Mateo, CA: Morgan Kauffmann, 2001.

[15] Gerhard Weiss (editor), "Multiagent Systems," The MIT Press, 2001.

[16] R. Abielmona, E.M. Petriu, and T. Whalen, "Multi-Agent System Environment Mapping by Entropy Reduction" Proc. ROSE'2005, IEEE Intl. Workshop on Robotic Sensing, pp. 30-35, Ottawa, ON, Canada, Oct. 2005.

About Larus Technologies

Through our culture of innovation and research, Larus Technologies has developed the next generation of embedded technology for developers of mission-critical C4ISR Systems and Security Systems.

With a solid foundation pioneering high level information fusion (HLIF) for the ever-changing defense and security industries, Larus is perfectly positioned to help Original Equipment Manufacturers (OEMs) make a world of difference. Working from the higher levels of the US Department of Defense's Joint Director of Laboratories (JDL) information fusion model, our technology not only delivers more knowledge, its adaptive learning algorithms deliver more accurate and more predictive information — faster.

58 Antares Drive, Unit 1, Ottawa, ON K2E 7W6, Canada
Phone: +1.613.244.8916
Fax: +1.613.244.9941
http://www.larus.com