An Autonomous Machine Learning Approach for Global Terrorist Recognition

Jerry L. Hill JHILL@AVUM.COM
Avum Inc., Senior Science Advisor
Atascadero, CA 93422, USA

Randall P. Mora RANDALL@AVUM.COM
Avum Inc., President
Malibu, CA 90265, USA

Editor: Rachelle Jensen

ABSTRACT

A major intelligence challenge we face in today's national security environment is the threat of terrorist attack against our national assets, especially our citizens. This paper addresses global reconnaissance which incorporates an autonomous Intelligent Agent/Data Fusion solution for recognizing potential risk of terrorist attack through identifying and reporting imminent persona-oriented terrorist threats based on data reduction/compression of a large volume of low latency data possibly from hundreds, or even thousands of data points.

Keywords: reconnaissance, data fusion, intelligent agent, service-based framework, data compression, derived data, machine learning, sensor

1. Reconnaissance

It was a cool morning in the late summer of 1943, though the temperature that day in North Africa would breech a hundred degrees. The Sun had not yet risen, but objects were clearly visible in the early morning light, as Lieutenant Paul Burkhardt pushed the throttles forward in his P-38 Lightening and began a slow climb into the predawn sky. The Sun had broken the horizon, just as he reached his cruising altitude and settled in for the three-hour flight across the Mediterranean. His mission this day was to fly reconnaissance over the Ploesti oil fields of Bulgaria. As he throttled back to level flight, the familiar rhythmic drone of the engines produced a soothing effect, prompting his thoughts to wander.

A student of History, Paul pondered the historical significance of this unique part of the world and how this cradle of civilization had changed since its awakening. Sentries on hilltops and messengers between encampments had been replaced by high resolution cameras in the belly of a 400 mile per hour twin engine aircraft which would skim the treetops to avoid detection and return to base with the valuable reconnaissance photos to provide the Intelligence needed to destroy one of the Third Reich's primary sources of fuel, needed to run

their war machine. In the months following, this Intelligence would provide the backdrop for the total destruction of these critical refineries by allied bombers in June of 1944.

Lieutenant Burkhardt's aircraft was providing what is now referred to as platform-based Intelligence. How could he have foreseen the reconnaissance platforms of today? Reconnaissance has advanced in step with emerging technology, now with Unmanned Aerial Vehicles (UAV)s controlled from half-way around the world, and Low Earth Orbit Satellites (LEO)s circling the globe at 17,000 miles per hour, sending back encoded digital data streams to be received by the Intelligence Community (IC). Once received and decoded/decrypted, these data must be analyzed for use in making decisions which have a direct bearing on diplomatic relations with other countries, and in providing battlefield commanders the Intelligence they need to effectively support the warfighters. So these are some of the platforms that provide reconnaissance data to those empowered to make life and death decisions; decisions which may change the course of history.

The IC reconnaissance mission has changed over the years and is in a state of flux to accommodate changing Intelligence needs. But this was not always so. The cold war era, which began at the end of World War II lasted until the gradual disintegration of the Soviet Union during the 1980s and its final collapse in 1991, with the resulting 15 separate countries. During that time Intelligence gathering consisted largely of aerial photography, and agents deployed in strategic geographical locations. Reconnaissance by Earth orbiting satellites soon followed, from the mid 1960s on.

In the 1990s the threat was changing, but our Intelligence gathering, which had been focused on the Soviet threat for decades, was not prepared for what was to follow. The terrorist attacks of September 11, 2001 were a wake-up call to the Intelligence Community that we were in a new era, requiring a completely new focus and methods that would employ a rapidly changing technology in reconnaissance and information gathering. It became imperative that we be able to provide situational awareness of this new threat.

2. Quantifying the Problem

The threat, though potentially enabled from sovereign nations, could manifest itself almost anywhere in the industrialized West. The IC, made up of multiple organizations, Including the CIA, NSA, NRO, FBI, Homeland Security, and military organizations all have their Intelligence missions, the primary goals of which are to protect our national assets from attack and assist our warfighters in Afghanistan and other potentially hostile environments. Coordination and sharing of information between and among these organizations is non-trivial.

Gathering of data from hundreds, if not thousands of geographical locations to provide situational awareness on a global scale, and to detect and act upon emerging and imminent threats, is now being integrated into our Intelligence process. Sensor data reduction and analysis may now be performed on those data to monitor and track terrorist movement in relation to our national assets; assets such as our national power grid, water reservoirs, population centers, seaports, airports and any assets that, if attacked, could cause a significant

number of casualties and/or disrupt one or more aspects of our national infrastructure, with potentially catastrophic results. Intelligence must now expand to include Cyber attacks on assets like the national power grid.

Data from all sources, platform-based, and information based - must be acquired from field agents and, strategically located sensors, into an Intelligence-driven engine that autonomously determines relationships between data sources, integrating the data and blending them to draw a logical conclusion that analysts can provide to government agencies in making tactical and strategic decisions. This backdrop motivated development of autonomous systems to assist in coordinating the several organizations making up the IC in a homogeneous Intelligence operation.

Combining the existing service-based frameworks with innovative concepts in integrated software applications, autonomous software agents can now be considered to help analysts reach conclusions in time to intercept a terrorist threat before it materializes into a devastating reality.

3. Devising an Innovative Solution

An approach has been devised to provide situational awareness through fusion of data from multiple sensors, possibly with wide geographical separation, loosely coupled to the concept of telemetry acquisition, reduction and presentation for analysis from Low Earth Orbit (LEO) satellites. We are now adapting new technology and processing concepts to the terrorism threat that must be addressed by the IC, in addition to traditional strategic and tactical applications familiar to battlefield commanders.

An Intelligent Agent (IA) Framework is now being modeled to expedite an autonomous data fusion process, to make validated results available in real-time, providing inference to analysts in support of strategic and tactical decisions to counteract global terrorist threats.

The IA Framework will process input from multiple sensors to provide situational awareness for the IC and battlefield commanders, in identifying and reacting to terrorist threats, in both domestic and warfighter environments. WMD, including nuclear "dirty bombs" and suicide bombers can be detected, given multiple reference points. Suspicious activities can be correlated such that otherwise unremarkable events and movements, when combined through data fusion, uncover an imminent threat.

4. Solution Description

Design and implementation is incremental, with the ultimate objective being a designed, documented IA Framework to acquire and process thousands of heterogeneous sensor inputs, quickly aggregate, integrate, intelligently fuse, and disseminate relevant Intelligence data in time for the IC to thwart an imminent terrorist attack. Refer to Figure 1 - Intelligent Agent Data Flow, below, for the following functional description of a preliminary model.

The Terrorist Threat Recognition model demonstrates multiple IA types functioning independently or as a team at logical nodes, allowing distributed technology to use an extensible model for probabilistic reasoning and machine learning. Each node can accommodate volumes of disparate data, fusing relevant sensor data to create composite Intelligence forms. Nodes might represent reconnaissance and surveillance sensors in aircraft, at ground locations, synthetic aperture radar (SAR), satellites (including interferometrics), and manually furnished Intelligence data. The service-based framework will allow an analyst to configure a selectable number of sensor proxies and IAs in providing a model that can be expanded upon when actual sensors are available.

Figure 1 represents the main components of a potential IA architecture. The core Framework and base classes contain the functionality for service-based communications, task management, and a set of extensible classes for workflow, model management, data mining, and probabilistic algorithm parsing.

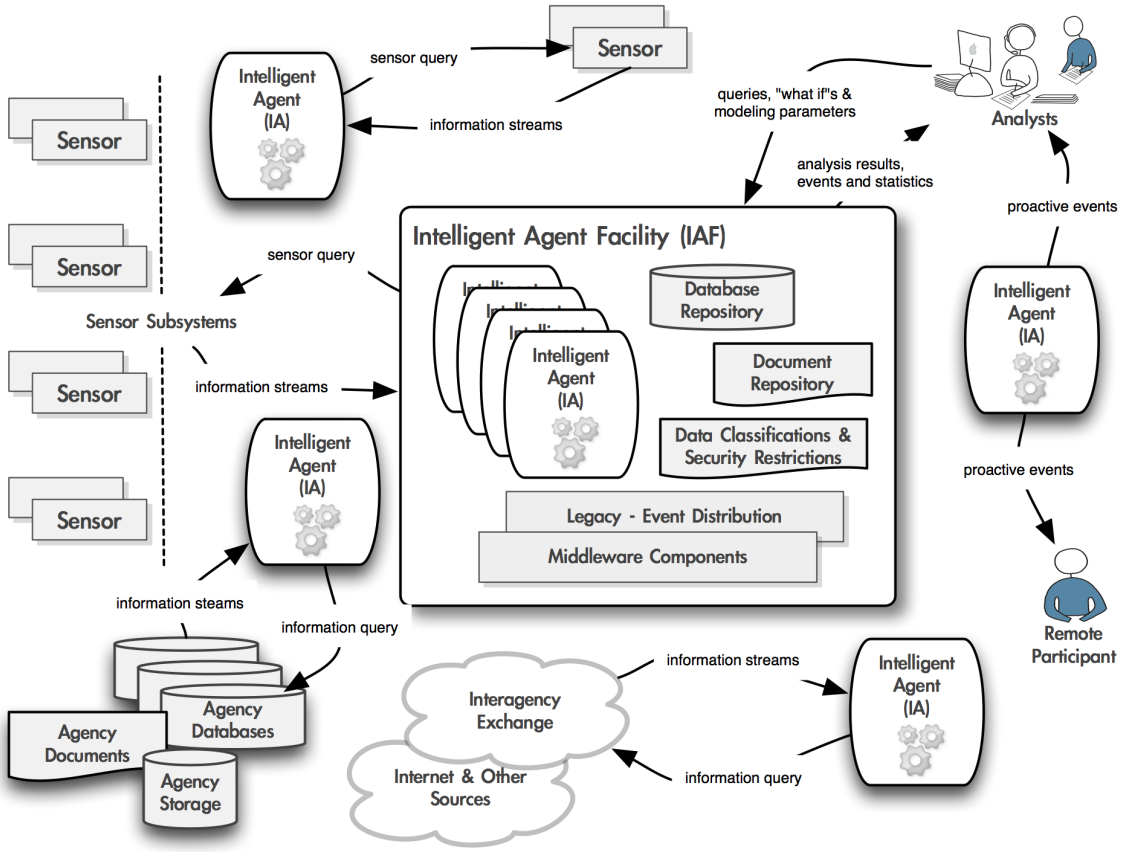

Figure 1: Intelligent Agent Data Flow

This IA Framework provides a prototype for multiple agent types that function independently or as teams or teamlettes at logical nodes, each node processing volumes of disparate data possibly representing thousands of sensors, fusing relevant information to create new forms of easily discernible intelligence. These are the nodes that might represent reconnais-

sance and surveillance sensors and manually input data from locations on land, sea, and air.

These logical agents will interact, applying data compression to eliminate redundant data and concentrate on the relevant data points for data fusion. Data fusion will consolidate multiple sensor inputs to provide validated inferences for action by field commanders and the intelligence community.

The agents will "learn" from experience and from human/machine interaction, which data are relevant and which to compress out of the processing loop to expedite wanted results. Those results would provide visual reference for maximum situational awareness. This team learning process is implemented through the update of Boolean strings, which further enable the agents to modify their rules for acquiring, matching, and fusing data based on success/failure history.

This prototype Framework would demonstrate the agents functioning as a team allowing the distributed technology to use an extensible model for probabilistic reasoning and machine learning. The end result will be the demonstration of "... a more semi-automated and integrated solution for enhanced decision-support capabilities."

This model provides a state-of-the-art analysis backbone for dynamic data streams originating from multiple sensors and observables. Sensors might be transducers on Intelligence platforms such as high altitude piloted aircraft, Unmanned Aerial Vehicles (UAV)s, missiles, satellites, land and sea surface vehicles, radar, radio telescopes or a variety of strategically placed scientific instruments. Derived measurements and manually entered data can be blended with sensor input to provide comprehensive and reliable inference to data analysts.

Dynamically changing data are input as constant data streams into the IA Framework. Since in the target system there can be thousands of sensors, and corresponding databases are very large, IAs must perform data compression on the input data to reduce the amount of data to inspect and process. Multiple algorithms at the sensor entry point eliminate redundant or irrelevant data as they enter the system. Initial parameters are provided for this processing, but the IAs soon update them based on rules-based learning.

To cite an IA an example of data reduction (data compression), an IA might eliminate a measurement even though it changes constantly, if changes are too small to affect the resultant processing in any meaningful way. In a battlefield environment, change could be attributable to movement of enemy vehicles. Such movement would be ignored, if contained within an area of no tactical significance. By setting limits on such data, a derived measurement would only be processed when that limit is exceeded, at which point the limits would be moved to accommodate the new position or other measurable object attribute. This data compression becomes significant if there are thousands of sensors.

In general, our model provides logical agents, which interact to remove redundant and irrelevant sensor data and process relevant data for data fusion. Data fusion blends multiple sensor inputs to provide validated inferences for action by field commanders and IC analysts. IAs "learn" from experience with human/machine interactions, which data are relevant, and which are not. Those results provide visual reference for maximum situational awareness.

IA team learning is implemented through a service-based extensible framework, enabling agents to plug in fully vetted and potential order-of-magnitude innovative machine learning methods for determining relevance in acquiring, matching, and fusing data. IA teams can be built/deployed to support real time missions, as well as used to fully vet new potential machine learning techniques that bring higher success rates to a given problem domain.

Data fusion is performed on reduced sensor data, blending related data to provide inferences to person-in-the-loop analysts and to other IAs. Legacy databases and computational frameworks are maintained and queried or executed, allowing IAs to 'learn' from both success and failures. Learning repositories are updated to provide a learned result and modify subsequent processing according to the learned behavior.

As IA interaction provides relevant data to be used to update situational awareness for battlefield commanders or the intelligence community, results are used to update a data repository. This repository is kept current by replacing old data with new. Old data are retired to a legacy database for future reference when needed by IAs or human analysts.

5. Conclusion

Our solution represents the convergence of research and emerging technologies to reach more accurate conclusions based on correlation of seemingly unrelated data from multiple sensors. IA Frameworks, through autonomous analysis, can identify threats that might otherwise be overlooked.

So, in this paper, we have demonstrated a real-life application for a service-based extensible framework that can provide a vehicle to fulfill our Intelligence Community's need for low-latency Intelligence gathering and rapid analysis. Intelligent Agents and Data Fusion, blend input from multiple sensors to foster conclusions that enable tactical and strategic decisions that must be made quickly enough to identify and counter terrorist threats before they materialize. The foundation for the service-based framework must be extensible to allow scientists, engineers, and university researchers to insert emerging technology and applications. The example given in this paper was the Intelligence Community's ability to identify and counteract multiple dynamically changing terrorist threats.

6. Acknowledgments

The authors are grateful to the late Paul Burkhardt and all the unsung heroes who set the stage for our current and future reconnaissance and intelligence gathering capabilities, and who have inspired this paper. To them we say thank you.

7. Biography

Mr. Hill has over 45 years experience in computer systems architecture, design, process-driven systems development, and project management. After serving 4 years in the Air Force where he taught computer systems and programming to Air Force officers "blue suiting" NORAD in the late 1960s, he moved to Sunnyvale, where he provided analysis and programming support for the Air Force Satellite Control Facility (AFSCF). Among his accomplishments are: leadership of projects to replace ground telemetry systems at Vandenberg, and Edwards Air Force bases, and Pax River NAS, in addition to managing winning proposals for those and Pt. Mugu NAS ground telemetry systems. He led a team to replace the weapons procurement system for the Army at Ft. Hunter Liggett in between projects at NASA/JPL, where he managed software development for uplink commanding and telemetry acquisition for Topex Poseidon, which was followed by Jason, both providing oceanographic data for predicting global weather patterns. In the 10 years from 1999 to 2009, Mr. Hill provided consulting services for the Illinois Institute of Technology Research Institute, ITT, Raytheon, Boeing, Lockheed-Martin, and Quintron Systems. In 2009 he joined Avum Inc. where he is currently managing two small Navy contracts; NASIS, and SAPR CMS. Now 72, Mr. Hill looks forward to another 20 years at Avum in this exciting and rewarding career field.