Maximizing Behavior Analysis with Massive Multi-Sensor Networks

by Serge Olszanskyj

The demand for mission-critical, real-time information is exploding in the Intelligence Community. With sensors becoming increasingly sophisticated, inexpensive, and interconnected, the response to this demand is coming in the form of massive multi-sensor networks. While such systems tantalize us with the ever-increasing potential for analytical omniscience, the reality is that analysts and decision-makers are overwhelmed by the large data sets being generated.

The purpose of sensors is gaining insight, not drowning analysts and decision-makers in data.[1] Hiring more analysts is ultimately cost-prohibitive, inefficient, and unsustainable. A precise, automated solution is needed, but current approaches do not adequately scale. The goal of our work at IEM is to develop technology and algorithms that effectively transform the exponential growth in data streams into an asset for analysts, providing automated suggestions of future outcomes and behavioral intent.

The ability to provide analysts with a large number of data streams that include full motion video (FMV) and other data types is closer to the present than the future. The wide-area surveillance sensor capability of a Gorgon Stare consists of a single 80-megapixel image or 10 to 12 separate images and includes an IR channel. There are plans to double this capability.[2,3] Military research labs are testing pilots' abilities to manage up to six drones while monitoring 30 to 40 different information streams, and they anticipate greater loads in the future.[4] Swarms of quadrotors, each with their own sensor suite, are on the horizon.[5] The relatively low cost and small footprints of sensor technologies are driving these advancements. What becomes relatively expensive is the time and effort required to analyze all those parallel data streams efficiently, and transform that data into insight.

The rapid increase in sensor availability and deployment is strongly related to (and partially due to) the recent explosion in inexpensive parallel computing hardware. In the 1980s and 1990s, parallel computing lived in the hallowed echelons of supercomputer enclaves. Now, there are multiple processors (along with a parallel computing-capable graphics processor) in computers, laptops, and handheld devices. Yet, writing software to efficiently harness that computing power is a non-trivial task. This challenge worried the computing industry so much that companies like Intel and Microsoft invested in education and software technology to develop better parallel computing programmers and tools.[6]

Similarly, successfully exploiting the emerging massive sensor capability is a challenging task in which the following questions should be considered:

- Is the Intelligence Community investing sufficiently in sensor exploitation technologies?
- If an analyst never has a chance to use the data a sensor produces, was the sensor worth the expense?

Getting the most out of a network of sensors is similar to exploiting parallel processors in one respect: the goal is to obtain maximum efficiency from the available resources. However, maximum efficiency for the multi-sensor analysis problem has two unique components: *maximum observation* and *maximum intuition.* Maximum observation is making the most of all the data available. Maximum intuition is obtaining the greatest understanding of a situation in the least amount of time. Achieving both in a massive multi-sensor environment requires creative, cutting-edge solutions.

The private sector is developing many different technologies that can be leveraged to build an effective solution set. "Big Data," the concept of using sophisticated analysis of large data sets to drive decisions, is currently big business.[7] Google has implemented self-driving cars, demonstrating that computers can smoothly sense a real-world environment, merge this information with historical data (i.e., street maps), and make automated, real-time decisions in the context of external human actions.[8] Apple's Siri combines natural language processing, artificial intelligence, statistical modeling, and cloud computing to produce a personal assistant with a personality on your mobile phone.[9] All of these advancements could contribute to improving the situational awareness capabilities harnessing massive multi-sensor data sets.

Defining the Mission

Our focus is on maximizing sensor capability to analyze human behavior. From this perspective, a sensor could include detection and tracking functionality, full motion video, and even human intelligence (HUMINT) data gathering elements. We do not focus on detecting and analyzing physical phenomena, such as plume source estimation for a chemical, biological, radiological, or nuclear (CBRN) release, as such analysis does not include the human element.[10]

We have identified two fundamental use cases for human behavior analysis with massive multi-sensor networks, ascertained from discussions with active and retired military analysts. We refer to these use cases as Real-Time Surveillance and Analysis (RTSA) and Post-Surveillance Data Analysis (PSDA). Both use cases can occur in one mission, but most missions fall into one category or the other.*

RTSA represents human-in-the-loop analysis of live multi-sensory data feeds. RTSA missions typically focus on the real-time tracking, monitoring, and intelligence-gathering of persons or events of interest, or provide general "eye in the sky" support to a mission.

PSDA represents the extraction of information from recorded sensor data streams after an event of interest — for example, mining surveillance video after a crime has been committed. This data is typically searched for anomalous behavior leading up to the time of the event. From this data, detailed context of the event is constructed and actors identified.

Both use cases require maximum observation and maximum intuition. The use cases primarily differ only in how specific technologies are employed.

Maximum Observation

To understand the concept of maximum observation in a massive multi-sensor context, we start with a simple example. Consider a sensor network with three simple sensors. Each sensor offers only binary (yes/no) output. Then, assume an observation from a point in time where two sensors output "yes" and one outputs "no." A simple voting algorithm would determine the conclusion is "yes," but this doesn't give a true picture. Why did the one sensor vote "no"? Is it a different kind of sensor, with different resolution? Is it not geographically co-located to the other two? With three sensors, it is straightforward to report the details of each one to the analyst. But this solution doesn't scale. What is the solution for 100, 1,000, or more sensors, with each producing more complex data than simple binary output?

Consider this problem from the analyst's point of view. What does an analyst really need to see from a sensor network? With a sensor network so large, it will not be feasible for an analyst to examine the data stream from every sensor. There must be a higher-order interface

* When we speak of the role of the analyst we are also including the role of the decision maker. We realize that in many organizations, those roles may be held by separate individuals.

What works today may not work tomorrow, and the software layer has to adapt to help the analyst stay ahead.

present — a layer (e.g., software and algorithms) between the analyst and sensors. What does this software layer need to accomplish for the analyst to achieve maximum observation and maximum intuition?

To achieve maximum observation, the input data stream must be fused into a format that the analyst can easily ingest. There are a number of multivariate data visualization techniques that would work for sensor networks:

- Spatial (e.g., geographical map display)
- Hierarchical (e.g., collapsible trees or graphs)
- Temporal (e.g., timelines)
- Some combination of spatial, hierarchical, and temporal[11]

Organizing input data streams in these ways saves the analyst time by providing a high-level view, but with detailed access to individual pieces when needed. These are good (and necessary) approaches to maximum observation. However, while this solution scales in the presence of many sensor inputs, it will not easily scale in all cases. For example, how does one organize and present 100 FMV data streams?

Maximum Intuition

Naturally, if we move some of the decision-making to the computer, allowing it to decide what information is important and what isn't, and present only the "important" data to the analyst, we can significantly streamline the input to the analyst. This is the realm of classification, queries, dimensionality reduction, and other data mining techniques.[12] Such technologies are important parts of any software layer designed to assist the analyst. The fundamental drawback to these approaches is that they require the analyst to supply the system with a hint (e.g., a keyword or image) to initiate the interpretation of the data.

The problem, of course, is that analysis does not always work that way. *A priori* knowledge of what to look for is often not available. Sometimes a "person of interest" is not interesting until she/he has done something that might be considered unusual. Sometimes it is not what happens that raises concern, but what does not happen. How do you query the concept of "not fitting a pattern I do not know beforehand"? These problems can only really be solved by immersing the analyst *in situ* in the space and time of interest to sense when something is amiss.

This is the crux of achieving maximum intuition — "the power to see the invisible."[13] The ideal for a sensor network is to completely automate the analysis. The software layer would function as the "brain" for the network that would analyze the data autonomously, determine the context and meaning of the situation, and report or even act on the analyses. This is the ideal, but it is not imminently practical. More realistic is to plan for the software layer to assist, not replace, the analyst. For this to work, the analyst would have to give up some tasks to the computer. Of course, there are some tedious tasks that analysts will happily give up if a computer will do it for them. Some tasks that involve a little more thought, intuition, or on-the-job experience may not be so easy to relinquish. The software layer will have to earn the trust of the analyst for that to happen.

There are a number of methods that can give the software layer some intuition skills, such as object identification through training sets[14] and stochastic methods.[15] While these approaches are assets, it is not clear if they can adapt quickly enough to be practical. In an asymmetric, human behavior-based threat space, the antagonists may adjust their methods in response to evolving analytical techniques. Thus, what works today may not work tomorrow, and the software layer has to adapt to help the analyst stay ahead.

Ironically, another important dimension to maximum intuition is the ability to handle "nothing" efficiently. In many missions, "no activity of interest" occurs far more often than critical events. Much of the current research focuses on finding events of interest, but does not examine how to improve the analyst's efficiency when there are no events of interest. Systems that tend to self-monitor, filter, and provide high-level alarms are a typical way to handle this problem, but these approaches leave out important insight.[16] Data streams deemed non-interesting by automated tools may very well provide important context that improves the analyst's overall situational awareness. In these cases, can a system provide maximum intuition and still allow an analyst to turn his or her attention to another task during lulls in the mission?

The software layer will not only have to learn, it will have to learn on-the-fly and adapt. Maximum observation then becomes maximum situational awareness for both the analyst and the software layer. Maximum intuition occurs with both analyst and computer developing and exchanging insight in real-time, adaptively.

Conclusion

Given the evolving technological advancements in sensors, taking full advantage of their capabilities without overwhelming analysts is a critical problem. In particular, we consider the problem of human behavior analysis with massive multi-sensor networks. Traditional approaches of data visualization, data mining, and probabilistic models are key components in the analyst's tool box, but they may not scale gracefully in the presence of massive, multimedia data sets. Our key focus areas are adaptive intuition in the presence of evolving situational awareness without prior insight and the appropriate handling of large blocks of data that are sparse in critical events. The goal is to elevate the automation assistance for the analyst from a data stream level to an intuition level. To that end, IEM is developing a robust software layer that provides maximum observation and maximum intuition with the ability to suggest intent with precision, in a form that is efficient, adaptive, and trustworthy. Q

Dr. Serge Olszanskyj *is a Computational Scientist at IEM with a Ph.D. from Cornell University and a B.A. from Dartmouth College, both in the fields of Electrical and Computer Engineering. He has worked in the industry on high-performance parallel computing, atmospheric dispersion models, active noise control, flight simulators, virtual reality military trainers, and quantitative risk models. He is the co-author of a patent in synthetic environment generation.*

REFERENCES

1 Adapted from "The purpose of computing is insight, not numbers.," R. W. Hamming, *Numerical Methods for Scientists and Engineers,* 2nd ed. (Dover Publications, 1987).

2 "Gorgon Stare Broadens UAV Surveillance | AVIATION WEEK", November 3, 2010, http://www.aviationweek.com/aw/generic/story_generic.jsp?channel=dti&id=news/dti/2010/11/01/DT_11_01_2010_p30-261179.xml&headline=null&next=10.

3 "New Reaper Sensors Offer a Bigger Picture," *Air Force Times,* n.d., http://www.airforcetimes.com/news/2009/02/airforce_WAAS_021609/.

4 "Military Faces Info Overload from Robot Swarms," *Msnbc.com,* September 7, 2011, sec. Innovation, http://www.msnbc.msn.com/id/44430826/ns/technology_and_science-innovation/t/military-faces-info-overload-robot-swarms/.

5 "Swarming Quadrotors Get Nano-ized," *IEEE Spectrum Automation Blog,* February 1, 2012, http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/swarming-quadrotors-get-nanoized.

6 "Intel and Microsoft Invest in Parallel Programming Research - SD Times: Software Development News," n.d., http://sdtimes.com/SearchResult/31854.

7 Steve Lohr, "Big Data's Impact in the World," *The New York Times,* February 11, 2012, sec. Sunday Review, http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html.

8 John Markoff, "Google Cars Drive Themselves, in Traffic," *The New York Times,* October 9, 2010, sec. Science, http://www.nytimes.com/2010/10/10/science/10google.html.

9 "What Makes Siri Special?," *PCWorld,* October 24, 2011, http://www.pcworld.com/article/242479/what_makes_siri_special.html.

10 Chunfeng Huang et al., "Bayesian Source Detection and Parameter Estimation of a Plume Model Based on Sensor Network Measurements," *Applied Stochastic Models in Business and Industry* 26, no. 4 (July 2010): 331-348.

11 Alexandru C. Telea, *Data Visualization: Principles and Practice* (Wellesley, Mass.: A K Peters, 2008).

12 Chandrika Kamath, *Scientific Data Mining: a Practical Perspective* (SIAM, 2009).

13 Gary Klein, *Sources of Power: How People Make Decisions* (The MIT Press, 1999).

14 Deepak Khosla and David J. Huber, "Method of Recognition and Pose Estimation of Multiple Occurrences of Multiple Objects in Visual Images" (presented at the Automatic Target Recognition XXI, Orlando, Florida, USA, 2011), 804903-804903-11, http://link.aip.org/link/PSISDG/v8049/i1/p804903/s1&Agg=doi.

15 Abhijit Mahalanobis, Robert Stanfill, and Kenny Chen, "A Bayesian Approach to Activity Detection in Video Using Multi-frame Correlation Filters," in *Proceedings of SPIE,* vol. 8049, 8049P, Automatic Target Recognition XXI (SPIE, 2011), 80490P-1 to 80490P-12, http://link.aip.org/link/PSISDG/v8049/i1/p80490P/s1&Agg=doi.

16 N. Finne et al., "Experiences from Two Sensor Network Deployments—Self-Monitoring and Self-Configuration Keys to Success," *Wired/Wireless Internet Communications* (2008): 189–200.