Redefining Military Intelligence Using Big Data Analytics

HARIDAS M

National security and defence related data being generated from multiple sources will have to be fully analysed for national decision-making, military operations and better situational awareness in the future, especially at the national and joint Services levels. Today, machine data is generated by the movement of ships, aircraft and vehicles, satellites in space, drones, Unmanned Aerial Vehicles (UAVs), reconnaissance aircraft, sensors and Battle Field Surveillance Radars (BFSR). Human generated data include data from social media sites like YouTube, Facebook, Tweeter, etc. Business data is generated from all e-commerce transactions. All these data have intelligence which we cannot afford to miss in the future. Big data analytics will be utilised for intelligence gathering in the near future since inputs for national and military intelligence are obtained continuously during peace and war, the quantum increasing exponentially during crises and war. Human analysis of this information and intelligence data is well beyond physical capability; therefore, big data analytics-based intelligence will provide the requisite output for decision-making and conduct of operations. Big data analytics can give a big boost to intelligence gathering when there is an information deluge because of the following aspects associated with future data:

- **Volume:** Data will grow at scale ranging from terabytes to petabytes.
- **Variety:** Data will be in structured, semi-structured, unstructured formats and from different sources.

- **Veracity:** Managing the reliability and predictability of inherently imprecise data types will be challenging.
- **Complexity:** Relationships, hierarchies, and multiple data linkages will have to be connected and correlated

Discovery of useful, possibly unexpected patterns in data, non-trivial extraction of implicit, previously unknown and potentially useful information from data by automatic or semi-automatic means of analysis to discover meaningful patterns leading to data simulation and modelling will be the order of the day as part of the Intelligence, Surveillance and Reconnaissance (ISR) philosophy.

Case Studies

The US **Argus** ground surveillance system collects more that 40GB of information per second. It is an autonomous Real Time Ground Ubiquitous Surveillance Imaging System and has a 1.8 gig pixel video camera with 12 frames per second (fps) and 368 sensors. It collects 6,000 terabytes of data/imagery per day and feeds it to Home Land Security. Post 9/11, this has served as a big deterrent.

Memex developed by the Defence Advanced Research Projects Agency (DARPA), which stands for a combination of "memory" and "index" is a high-tech internet tool for internet crawling, internet searching, data aggregation, data analysis, data visualisation, data extraction and image analysis. Data flowing in the networks of the USA is analysed on a real time basis for proactive actions.

The **Green Line Vessel Selection System** (VSS) helps analysts and decision-makers identify targets of interest and make interdiction decisions. Complementing existing Command and Control Systems (C2) with operational analysis, VSS is a comprehensive decision-support platform that uses Green Line's proprietary risk methodology to provide a clearer understanding of the maritime domain's actors, assets and actions. Robust information sharing capabilities allow partner agencies and allies to share information selectively and securely.VSS develops a holistic maritime picture by integrating data from multiple sources, such as terrestrial and satellite-based ship position/AIS data, vessel and company characteristics data; entity or vessel watch lists; and agency databases containing details on vessels' cargo and crew. This data is analysed and presented in a user-friendly interface which supports a collaborative work environment.

How can Big Data Analytics Enhance Military Intelligence Productivity?

In the time to come, we will be swimming in sensors and drowning in data. Big data collected as part of the ISR grid can be analysed automatically. Engaging with social media can give us answers to questions like why, when, what, where, who, how? to many a security incident which otherwise may go unnoticed. Big data analytics has a potentially significant role in helping to manage the data deluge and assisting analysts in focussing their efforts. Of course, there will be complexity and challenges in preparing infrastructure, changing mindsets, introducing skills and understanding, and in ensuring that use of big data is handled responsibly and sensitively with regards to individual privacy and protection.

The Defence Intelligence Agency should build on the huge investments being made in this area by the commercial sector and in doing so, ensure it is well-positioned to track and exploit further commercial technological developments as and when they occur. In modern combat scenarios, a data scientist helping interpret and analyse data could save many more lives than a hundred troops on the ground. Big data, with computational analysis, can deliver insights enabling commanders to proactively identify hot spots for operational planning. Sophisticated analysis of the massive datasets will be possible in the long run with analytical tools developed specifically for this purpose. Dedicated resource monitoring systems will provide better asset visibility, a must for network-centric warfare. Big data analytics capability using machine learning as a tool will ensure no hidden information escapes the eyes of the commander howsoever huge the data set is.

Terrorism/ proxy war/ left wing extremism/ conduct of training by opponents/ deployment of forces/ sponsoring of non-state actors are a few examples of the inputs likely to be received and analysed. Conversion of satellite data and technical encrypted intercepts require special tools which will be provided by big data analytics. Algorithms can be developed to analyse hundreds of thousands of open-source documents generated each hour and compare them with the human intelligence gathered and billions of historical events, and then have predictive capability to anticipate specific incidents and suggest measures proactively. The information about a suspicious person can be queried with big data tools on social networks, shopping sites, and entertainment sites, and from the web logs of the search carried out in the net, and actions can be taken. Big data analytics can provide the following platforms as value additions:

Threat Alert System: Algorithms can be designed which can alert the commanders on the mention of concepts such as terrorism, bombs, riots, etc. across various information sources. Trending offensive videos/content specific to a person, organisations, geography, etc. can be undertaken. For example, trending video related to the Muzzaffarnagar riots. Any unusual hike in social activities in a specific source of information can be studied. For example, a surge in discussions on Twitter related to a specific topic. Websites promoting unlawful activities can be identified in time to take action.

Big data analytics has a significant role in predictive capability to anticipate specific incidents

Social Media Monitoring: Prime topics/concepts being discussed in the social media can be monitored and studied specific to geography, person, and organisation, etc. Analytics of information sources, for example, affinity of information sources to a specific user group, geography, etc will have high intelligence value. The sentiments of people regarding a policy or concepts can be known and proactive actions taken, as required.

Information Mining: Information in the news/documents related to a specific person, concepts, etc. can be searched to find out the concepts related to a given concept. For example, “riots” might be related to “Gujarat” during 2002. Related documents providing different representations of the same information can be found, for example, the different ways explosives are mentioned in articles. New information regarding a specific topic, for example, new articles, publications, white papers, patents related to missile defence comprise inputs that will be of great help in planning the intelligence strategy at the higher level.

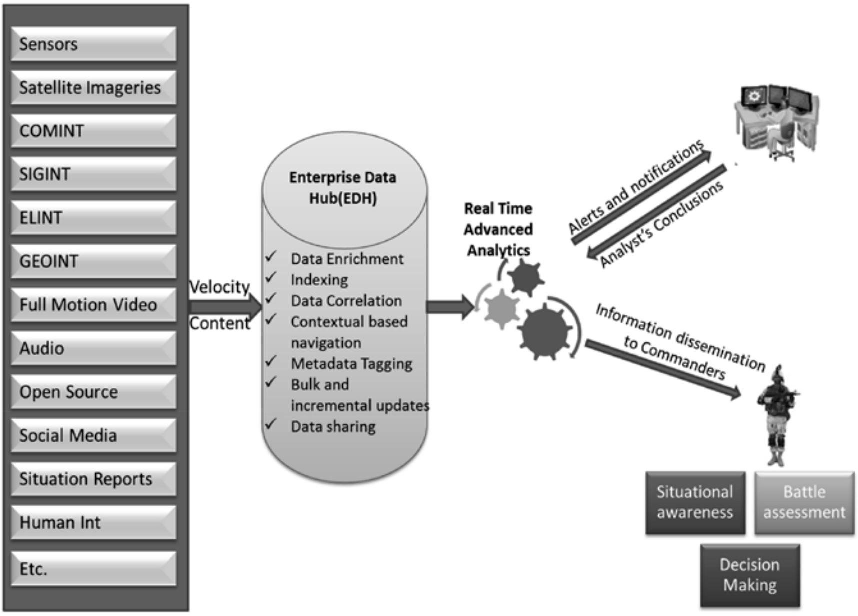
Social Network Monitoring: A study of related people based on social networks can be undertaken. Topics discussed by related people which represent information about the behaviour of a person will have very high intelligence value. Different social profiles of a person in Twitter, Facebook, and LinkedIn, etc. can be studied and websites related to a specific person based on his social profile, created content, friends circle, etc. can be analysed. Macro analysis of social media graphs to identify users’ groups active on a website can be analysed. For example, the age group division of users active on Twitter and Facebook from Kashmir.

Document Analytics: Concepts, topics discussed in a document collection can be studied. For example, the topics that the Foreign Ministry of Pakistan is focussed on, based on analysing the articles/documents on its website. The documents can be grouped in separate segments: documents discussing politics, sports, games, foreign affairs, etc. Finding trends related to specific topics might be an important way to discover the hidden insights in a document.

Cyber Security: Network managers will be dealing with millions of attacks every day, and big data analytics can be applied to spot advanced persistent threats such as socially engineered attacks designed to steal government information as has happened in the past (the Chinese attack on our websites). Most hackers have a *modus operandi*, which once identified can be used to predict the form of future attacks and put appropriate defensive measures in place. The application can also be used for offensive aspects of cyber security.

In Counter-Intelligence/Counter-Terrorism (CI/CT) operations, big data collected by drones, satellites, UAVs/ technical intercepts, etc. can be automatically analysed based on the big picture provided by ISR data which can be created as part of the big data initiative. This will allow the CI/CT operations to be carried out in real time. Data intensive Ground Information System (GIS)-based applications can be used with big data platforms in the hinterland to assist the forces deployed to fight left wing extremism.

Fig 1: Conceptual Layout of Big Data Application Based Intelligence Gathering System



Implementation Model

In future combat situations, a data scientist could save more lives than troops on ground

The threat landscape is getting more serious day by day with cyber threats, Wiki leaks, whistleblowers' acts all leading to an ambiguous ambience admidst the uncertain policies and laws that are open to interpretation. The military leaders of the future cannot survive if they cannot handle ambiguity. Assignments and responsibilities coming the way of military leaders today, were not even dreamt of a decade ago. The social media is an enabler now and no longer a search engine. It is a medium of communication that allows a multi-dimensional view against the earlier uni-dimensional view

Implementation should follow an evolutionary path within the wider information management/information exploitation strategy and be coherent with the defence Information and Communication Technology (ICT) strategy. It should not rely solely on big data capabilities for the defence forces. Rather, the defence forces should build on the huge investment in this area being made by the commercial sector and in doing so, ensure that they are well positioned to track and exploit further commercial technological developments, as and when they occur. It can be started a small step using existing data with which the defence forces are comfortable. Things should be kept simple initially and then move to more complex uses subsequently. The initial focus should be on day-to-day functional issues since this will help build up a case for the usefulness of big data. Some guidelines are as follows:

- Prepare the defence intelligence agencies culturally to transform operations based on big data analytics through change management applied top down.
- Incorporate data analytics/data mining as a module in all the ongoing projects where better delivery can be achieved.
- Initially use Commercial off-the-Shelf (COTS) software on an experimental basis and then opt for own Research and Development (R&D) for product development after gaining sufficient experience to translate the requirement technically.
- Technology and non-disclosure agreements on data sharing, to develop algorithms developed by industry, R&D institutes and academia on live data which will be more realistic than development on synthetic data.
- In-house skill set development, from the data analyst to the data scientist.
- Pilot projects to validate concepts.

Conclusion

Big data analytics-based intelligence gathering has a potentially significant role in helping to manage the data deluge and assisting analysts in focussing their efforts. Evidence-based information superiority and fully analysed intelligence outputs will give increased situational awareness which will be the primary requirement of commanders at all levels to fight a war of any type in the future. The intelligence systems should be able to collect, collate, filter and process all types of input, from structured to unstructured, including live feed, and display it to commanders at the hierarchical level, through Tac C3I systems to the field commanders and through Command, Control, Communication, Computers, Intelligence, Information (C4I2) systems to the leaders at the strategic level. Intelligence of such class, based on big data analytics, will enable commanders to assess the battlefield situation in real time and in a better way to take appropriate and timely decisions. Such a system can integrate the efforts of all intelligence gathering functionaries to optimise the resources and results. Intelligence report generation and distribution to local, national and international partners can become quick and easy. Such systems can be continuously refined in response to the changing behaviour of the target and stakeholders.

Col **Haridas M** is a Senior Fellow at CLAWS. The views expressed are personal.

References

- @war The Rise of Cyber Warfare by Shane Harris
- http://www.forbes.com/sites/thomasbrewster/2015/04/17/darpa-nasa-and-partners-show-off-memex/
- military.com (for Memex)
- http://www.gizmag.com/argus-is-darpa-gigapixel-camers
- http://www.greenlinesystems.com/
- http://www.sigmoidanalytics.com/casestudies/
- www.ilearning.oracle.com
- www.ibm.com
- www.sap.com
- www.sas.com
- www.wikepedia.com
- Research papers at www.ijarcsse.com
- www.rusi.org