20th ICCRTS

"C2, Cyber, and Trust"

Paper 095

Increasing Trust in Network Situation Awareness

Topics

Topic 2: Organizational Concepts and Approaches

Topic 3: Data, Information, and Knowledge

Authors:

Marielle Mokhtari
Étienne Martineau
Régine Lecocq
Valérie Lavigne

Defence R&D Canada – Valcartier
2459 Pie-XI North, Quebec City, QC, G3J 1X5
Canada

Point of Contact:
Régine Lecocq
E-mail: regine.lecocq@drdc-rddc.gc.ca
Phone: 1 (418) 844-4000 ext. 4124
Fax: 1 (418) 844-4538

Increasing Trust in Network Situation Awareness

Marielle Mokhtari, Étienne Martineau, Régine Lecocq, and Valérie Lavigne
Defence R&D Canada – Valcartier

2459 Pie-XI North, Quebec City, QC, G3J 1X5, Canada
Phone: 1 (418) 844 4000 Fax: 1 (418) 844 4538

Abstract

Over the last decade, several defence and law enforcement organizations have considered modern Social Network Analysis (SNA) methods and techniques to assist them in their work. The Intelligence community in support to Command and Control is no exception. Their specific context, which often involves covert networks, makes it particularly difficult to overcome the missing data issue. Furthermore, SNA standard measures hold the premise to be applied on networks when all the links and nodes that compose the network are represented. This premise makes the SNA measures somehow unsuitable to the context discussed here. Indeed, when conducting intelligence activities, new network related data is collected and inserted on a continuous basis creating a certain level of variation in the network measures.

SNA research work conducted at DRDC Valcartier investigates a specific approach called the Spike use case. This approach supports the monitoring of SNA measures variations but more specifically enables the identification of sudden peaks in these variations (spikes). Once identified, these sudden variations are explored to be explained and information about them is provided to the analyst, such as the links and nodes involved in the variation triggering the spike. Subsequent historical analyses may even specify when or where the perceived changes took place in the overall network.

Identifying information about these spikes allows the analysts to refine their network perception by providing an understanding of its stability and the level of trust that can be given to the resulting SNA measures. Indeed, when spikes are detected, the analysts need to be alerted as it means that the shared perception of the network was altered. Such modification can either be caused by a change in the real world network or it may simply be the perception in itself which has evolved due to newly acquired network data. Visualization is a key enabler to comprehending the context when analyzing these spikes. There is a need to provide the analysts with information about measurements that changed, as well as where and why these changes happened. This paper describes how it is envisioned to convey the Spike use case information to the intelligence analysts and augment their trust in the analysis performed.

1 Introduction

For the last decade, the Canadian Armed Forces (CAF) and their allies have been increasingly involved in operations taking place in unfamiliar socio-cultural settings. In addition, their adversaries have also evolved and the changes from conventional to irregular warfare have now set the stage for a combination of both. Undeniably, even conventional operations have components of irregular warfare where groups with mitigated allegiance to blue forces rapidly turn into adversaries. This creates a need for the CAF to be better equipped so they can increase their understanding of such complex environments. Social Network Analysis (SNA) tools and techniques have been more systematically investigated since 9/11, especially within the Intelligence Community (IC). This domain, under research at Defence R&D Canada (DRDC) for the last four years, has progressed significantly (Lavigne, Lecocq and Gouin, 2012; Lecocq, Martineau and Caropresso, 2013; Martineau and Lecocq, 2013; Lavigne, Lecocq, Martineau and Mokhtari, 2014) with expanding momentum following the current events associated to the radicalization phenomenon in western countries. The question is not anymore whether SNA can be of advantage to the IC but rather how it can be best utilized by the intelligence analyst as well as integrated into other intelligence support systems.

2 Problem Definition

CAF objectives, using SNA, cover the full spectrum from white to red awareness, aiming at better grasping the social fabrics of the population in place as well as their connections to either insurgents, the local government, terrorist groups or organized crime cells. The objectives also include identifying some of the relations between these networks of interest and the mechanics supporting their interconnections internally or externally. Lately, the SNA application has also been leveraged as an enabler to the analysis of social media.

Previous work (Lavigne *et al.*, 2014) introduced how DRDC research has envisioned SNA as a full capability rather than as an isolated analysis technique enabled by technology. From this perspective, processes and work performed prior to (network data extraction) and after (sense-making) the analyses are as important as the analyses phase itself. Essentially, extraction of data pertaining to the network under examination is a dynamic process where information is acquired on a regular basis from multiple structured and unstructured data sources. Once the analyses have been executed on the network, the analysts must make sense of the situation and adjust their understanding based on the changes perceived. In an operational context, the decisions and activities grounded in SNA results can have substantial consequences to both our military personnel and the local population; thus impacting the mission success. Some critical challenges are discussed in the following sections along with how the envisioned Spike use case may respond to some of them.

2.1 Dynamic Network Data

The data collected in order to build a representation of the social networks from a theatre of operation (or at a more strategic level) emerges from numerous different structured and unstructured data sources. Where historically most information was obtained from HUMINT

reports or SIGINT sources; other sources are now increasingly contributing to the overall social networks awareness. Within the currently developed prototype, network components are extracted, organized and stored from continuous data streams in order to build representations that are as close as possible to the real world networks. It is on such representations that network analyses are run in order to refine information about the situation.

While the evolution of these representations needs to be conveyed to the analyst, it would also be inefficient to create an alert for each newly introduced network component. Obviously, prompting the analyst for every new addition would create a cognitive overload. This issue is partially managed in the proposed Spike use case where the analyst is only alerted when major network structures changes are taking place. This approach helps mitigating the information and cognitive overload issue as expressed by (Patterson, Roth and Woods, 2001) and discussed in (Lavigne *et al.*, 2014).

2.2 Missing Data Issue

Several elements have an impact on the IC's ability to collect in a timely manner the data pertaining to a social network under observation.

First, much of the data of interest is changing over time. The most stable social data, as for instance "who is the mother of whom", represents a very little portion of the network. Indeed, military data of interest concerns questions like "who meets with whom" or "who is phoning whom", which is information that can be added almost to the first one. This consideration means that analyzing social networks is always in a definition state.

Second, the social nature of the data makes collecting meaningful information particularly difficult. Indeed, many of the social topics under review are the results of indirect interpretation of indicators alluded to by some data. For instance, if one wants to establish a relation of influence between two individuals in a network, there is no instrument directly measuring the influence of one on another. Therefore, if the analysts want to grasp such social phenomenon, they must first identify the indicators for it, the data supporting the indicators and the sources generating this data. The result is an interpretation that a relation of influence might exist. Consequently, there should be an acknowledgement by the analysts that what is portrayed about the social network is only a representation that is more or less realistic/accurate of the real world network. At this point, the analysts' knowledge and expertise are required to comprehend the networks analyses results; they are the best positioned persons to validate findings emphasized through SNA.

Finally, in our context of operation the social networks under observation are frequently of a covert nature. Indeed, in order to reduce their vulnerability, the adversaries take many precautions to hide or distort information about the network structure or activities. Therefore, the intelligence collection process appears even more complex, heavily relying on counter stratagems in order to acquire the desired data.

2.3 SNA Measure Issue

Based on available theory, analysis of social networks should be performed on networks devoid of missing data (Borgatti, Carley and Krackhardt, 2006). Unfortunately, there are no reliable ways to assess the goodness of fit of the network representation and therefore, the level of missing data will almost always remain/be unknown. Most of the time, analysts may not even be aware of the problem and assume that conclusions drawn from SNA are still valid. The initial trust they place in SNA could lower rapidly if analyses results fail to match real world situations and effects. Analysis performed on real operational data has shown that results provided by most SNAs are not reliable if networks have missing data (Wang, Shi, McFarland and Leskovec, 2012). Unfortunately, as exposed in section 2.2, this is often the case since social networks are hard to observe, even more for covert networks.

The objective of the intelligence process is to provide information about the situation being faced. Since a continuous process can be considered as a sequence of steps repeated over time, the intelligence process provides a continuous flow of new information. However, this process is always in a catch-up mode, and the information flow coming from the different sources has its limitations.

This process first builds an “initial awareness” of the environment. In this step, the learning curve is steep and the information flow is high; the understanding of the social network evolves rapidly and it may appear to grow and morph as missing data is reduced. In this initial phase, it is believed that even robust SNAs can be misleading (Howison, Wiggins and Crowston, 2011). However, it is assumed that when the knowledge base of the network reaches a certain size, its information flow will reduce. At that point, new information should relate more to changes in the real network and most SNAs should provide relatively constant and reliable results. For the analysts to trust SNA, they must either use only analyses robust to missing data or possess a mechanism that allows them to assess if the perceived network is similar to the real network.

2.4 Perception vs. Real World

It has been mentioned that one objective of performing SNA is to increase the intelligence analysts’ situation awareness. Such awareness is passed on to the commander who will take meaningful decisions leading to mission success.

When they have to analyze and understand a situation, the analysts use the tools, technologies and techniques that are at their disposal. Most of these tools are based on data collected either through official intelligence sources or from external and open sources. In all cases they have to rely not directly on the reality but rather on a representation of the real world conveyed through the data and tools exploited. In other words, the analysts have a perception of the real world networks based on the collected data.

From a human factor standpoint, Endsley (1988) describes situation awareness (SA) as “an expert's perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”. The

three SA levels from her standpoint are usually described as “Perception” – “Comprehension” – “Projection” (Figure 1).

Figure 1 – Endsley 3 levels of SA

The term “perception” is defined in the dictionary (Oxford, 2015) as:

1) The state of being or process of becoming aware of something through the senses;
2) A way of regarding, understanding, or interpreting something; a mental impression.

In light of Endsley’s reference framework, the first definition from above would correspond to the first level of SA where the second definition would include some type of interpretation leading to the Comprehension, the second level of SA.

Perception is a process through which human beings recognize some stimuli through their senses but also provides an interpretation for them; it is therefore a mental construct. The analysts create perceptions in order to understand the world around them and its military implications. Figure 2 broadly depicts concepts many times discussed in the military context: the cycle between real world components, their capture and representation, as well as the perception and comprehension of the real world based on such representations. Obviously, it is impossible to capture all the real world components (1) in reports (2) and, as expressed previously, even if such an attempt was made, not all of the information would be available for the military personnel to do so. Moreover, some pieces of information are lost when captured in reports and from those reports only a fraction will be extracted (3) for representation (4) purposes. Finally, the perception and the comprehension (5) that the analysts build are based on partial information creating a gap (6) with the reality. In some cases, the gap concerns information of little importance but in other situations, this lacking information can have major impacts on the mission success.

Figure 2 – Spike use case foundation

3 Proposed Solution

Given the overwhelming quantity of data extracted from sources stream, it is impossible for the analysts to spend all their time acknowledging each new node or relation in the network. The Spike use case involves alerting the analysts only when significant changes are taking place in the network representation. Therefore, the main message conveyed by the alert to the analysts is not that the real world has changed but rather that their "perception" and their "comprehension" of the real world need to be updated. Subsequently, the analysts will have to determine if the changes are taking place either in the real world or only in its representation. For instance, a relation built between two individuals due to a stay in the same training camp might have always existed for some time but never been captured explicitly.

The Spike use case's second foundation is directly related to the SNA measures applied on a specific network in military context. As mentioned previously, the SNA theories and measures presuppose their application on a stable network. Prior to analyzing a dataset obtained after a military operation, the research team formulated the hypothesis that once passed the "initial awareness" phase and despite the network's dynamic aspects, SNA measures taken on the network should become relatively stable over time. The results were quite different; it was rather observed that the measures were repeatedly changing as new information was gathered and inserted in the network. Centrality measures that were taken appeared to change the order of importance of individuals on a daily basis leaving the analysts with very little confidence in the comprehension to be gained about the situation. The Spike use case was investigated to

mitigate this issue. It permits to overlook the changing results of the measures taken to only consider significant changes that should be flagged to the analysts.

3.1 Spike Analysis

3.1.1 Variability of measure in the network

As expressed in the previous section, network changes observed by the analysts are either the result of an evolution of the real network or the convergence of the perception with the said network. In both cases, it is difficult for analysts to assess to what extent these changes affect nodes in the network or the SNAs applied to it. In fact, one may want to answer the following questions regarding changes in a time interval:

1) Which nodes were affected by the changes;
2) To what extent;
3) Is this a real world change or a perception change; and
4) Does it invalidate previously performed analyses?

Unfortunately, there is no way one could answer all these questions correctly by simply looking at the structure of the network. However, the variability of SNAs measures over time can provide an insight into the answers.

Changing some aspects of the network directly alters the answers provided by SNAs. Moreover, the magnitude of the change in the results of an analysis is not directly proportional to the changes in the network. For example, it is possible to add several links and nodes in a network and barely notice the change in an analysis and, on the other hand, add a single relation and invalidate almost all previous analyses. This behavior forbids the use of changes in cardinality of nodes and links as an indicator of change that an analyst could trust. A better method would be to pick a single analysis that provides a quick assessment of the topology and use the variability of this measure to answer the questions mentioned above.

The notion of topology is extremely important for evaluating the changes in the network since global topology affects the importance of individual nodes and links. In fact, the study of the topology is the general idea behind SNA and in particular, centrality evaluation. However, centrality measures can only be applied on a static network and thus cannot be used on their own to study dynamic network. To answer the previous questions, one must find a way to quantify the change based on the topology. Comparing the result of two centrality evaluation calculated on the same network, but at different times, can provide the answer. The key here is that instead of discarding centrality measures results because they vary too much, their variability is leveraged to highlight relevant changes.

3.1.2 Quantifying the change in the network

To detect and highlight topological changes in the network, it is first necessary to define a method for comparing network representations at different points in time. The key factor of this method is that any change detected, i.e. the result of a comparison, must strongly correlate

with the change in the routing of the flow in the network. The quantification of the change in the flow must consider, for each combination of nodes, how possible paths between them have evolved during a time interval. Unfortunately, the mechanism of transfer and propagation in a network is not always available, particularly in conditions where data is missing. It is clear that the optimal solution, in an intelligence context with time and missing data constraints, will be out of reach. However, a good approximation would be adequate to provide the needed information on the network representation stability, and therefore, the level of trust one could place in SNAs.

To achieve this, it was decided to compare ranking of betweenness centrality analysis. Freeman's betweenness measure (Freeman, 1977) sums the proportions of shortest paths from one node to another that pass through a given node and ranks them in order of importance. Thus, a node with high betweenness is responsible for connecting many pairs of nodes via the best path and will be one of the first in the list. This, of course assumes that the flow in the network follows the best path, meaning the shortest paths which is probably not the case. However, as mentioned in (Freeman, Borgatti and White, 1991), a node's betweenness ranking is an indicator of its importance for the efficiency of a network. This observation can be generalized to conclude that a change in the ordering of nodes within that ranking is a good indicator of topological change no matter what the propagation or transfer methods in the network really are. Other centrality methods exist but their focus is less appropriate to quantify topological change. For example, closeness focuses on the position inside a network and degree is only about the number of connections inside an ego network. Also, on the practical side, Freeman's betweenness measure has the following advantages:

1) It has polynomial time execution complexity;
2) It is usable on disconnected network; and
3) It is available in most SNA library (if not all).

To quantify the change, one must select a network and perform Freeman's betweenness at two separate points in time, $(t, t+1)$ to obtain rankings BC_t, BC_{t+1}. The change is given by comparing the two lists using the following formula:

$$\sum_{n \in BC_t, BC_{t+1}} |P_n(t) - P_n(t+1)| = \sum_{n \in BC_t, BC_{t+1}} |\Delta P_n|$$

Where $P_n(t)$ is the position of the node in the ranking BC_t. This formula provides a process to compare two ranking lists. However, depending on the situation, the sum may only be applied to a subset of nodes like:

1) The one at the top of the list;
2) Those related to persons of interest; and
3) Only nodes that have moved more than a certain threshold.

This offers the flexibility to focus the detection on the operational needs in term of situational awareness for the intelligence analysts.

3.1.3 Change in analyst's perception

As mentioned previously, the appearance of a Spike does not directly alert the analysts about a change taking place in the real world, it rather indicates to the analysts that their perception and comprehension of the social network require to be updated. Subsequently, the analysts are the most experienced and best positioned persons to identify if the change is either new data about the real world or else a newly collected one on a previously existing reality. Based on this use case, the analysts can increase their understanding of the situation and entrust the SNA measures and processes.

3.1.4 Rehabilitation of regular SNA measures interpretation

Another significant value of the Spike analysis is to help understand how big the gap is between the real world and the analysts' perception. If the gap is reducing, it can trigger the standard interpretation of the SNA measures. Indeed, the network data extracted on a regular basis should demonstrate a stabilization of the number of nodes discovery while the number of relations could keep on increasing at a slower pace. For instance, when arriving in a new operational environment data flow from the collection phase would be high, unveiling numerous nodes and relations between them. However, once the initial collection effort performed, reality and its representation would converge providing a more stable network to which SNA measures could be applied with higher reliability and be better trusted by the analyst.

3.2 Visualization Positioning

To correctly interpret the spike alerts, their causes and effects must be identified and presented to the analysts. A suitable visual presentation enables analysts to investigate and analyze the why and how in order to enhance and strengthen their network perception. This section proposes a spikes comprehension framework concept that provides consciousness to the analysts of significant changes taking place in the network which could affect their perception and comprehension of the real world.

3.2.1 Visualization concept overview

The proposed framework concept for exploitation of the Spike use case is presented in Figure 3. The space is split into two main areas with different purposes:

- Spikes visualization over time : notifies analysts of significant changes in the network;
- Spikes cause and effect visualization: offers the analysts access to pertinent information in order to understand thoroughly the changes and their impact on the network. This area is in itself divided in three sub-areas:
    - Nodes ranking – ranking change analysis;
    - Network graph – graph change analysis; and
    - Nodes evolution – nodes ranking evolution.

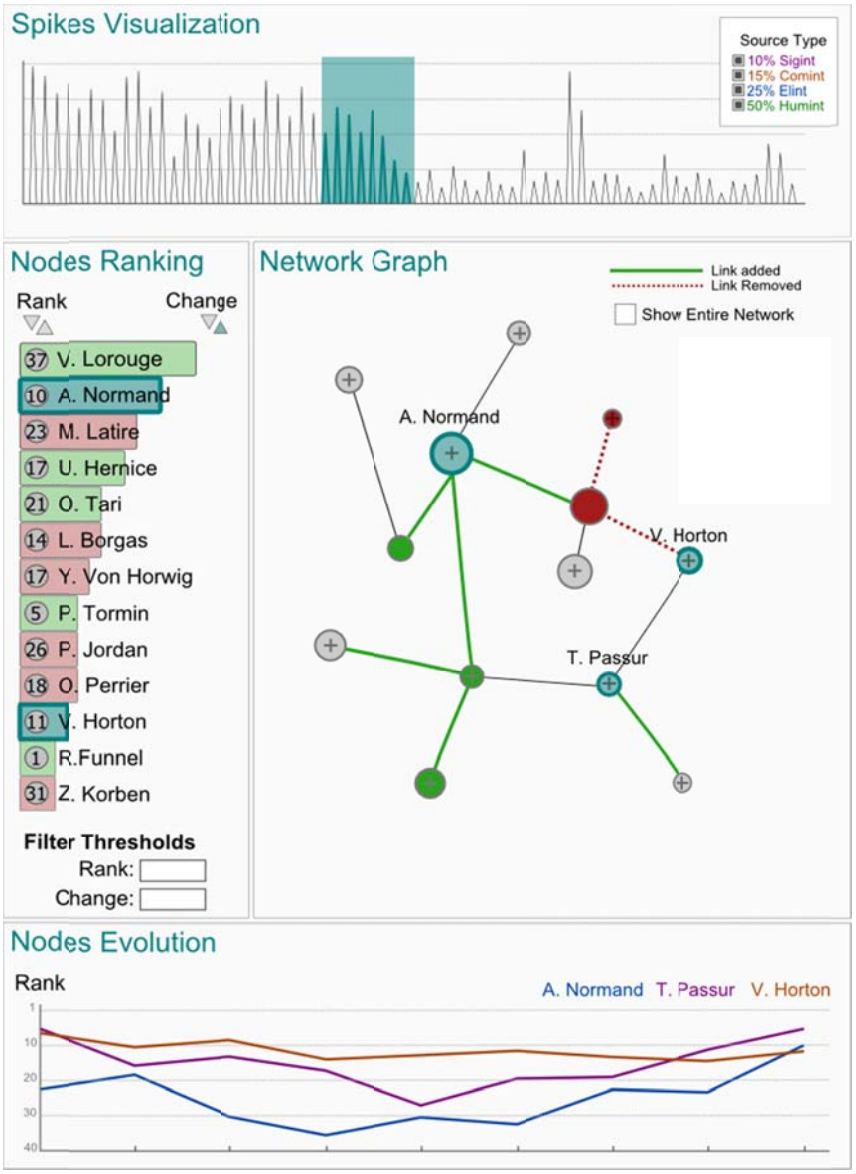

Figure 3 – Concept of the Spikes Comprehension Framework

3.2.2 Spikes visualization chart

The first step in the analysis is to have a look at when spikes occur over time and from what information source they come, using the spikes visualization chart in the topmost part of the display (see Figure 3). Analysts would expect fewer spikes and/or the spikes to become smaller over time as the perception of the network converges towards the real world network topology. While the large spikes occurrences are expected to reduce over time, they remain, at all-time, a notification of significant changes about the perception of the network topology.

Over the chart, measurement of time is built on units, such as years, months, weeks and days. A basic temporal unit – a day – is set by default. Then the spikes computed for each individual modification to the network are displayed for every temporal unit. In the chart, spikes can take

two shapes, a triangle or a line. In the case where a spike is represented by a triangle, it is considered as a *spikes aggregation* representing a set of spikes appearing simultaneously for that temporal unit (e.g. a set of reports/documents analyzed the same day leading to modifications in the graph). By doing a mouse-over a spikes aggregation in the chart, the analyst expands to a more granular level all the *spikes* really computed. At that moment, a new display is provided with a spikes visualization chart focusing on the temporal unit under investigation (see Figure 4). A spike at the lower granularity is represented by a single line. When the spikes visualization chart is zoomed out, the spikes are aggregated for larger time intervals than the basic temporal unit. The value for this spikes aggregation corresponds to the maximum of the spikes composing it, just as it corresponds to the maximum value of the spikes it contains.

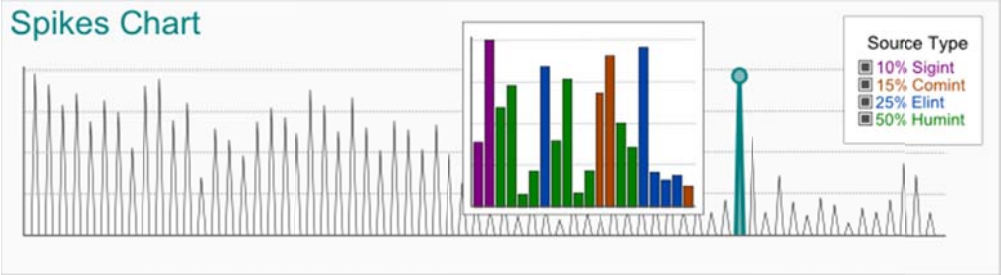

Figure 4 – *Spikes* versus *spikes aggregation*

Spikes can be colored according to the information source, or another relevant property that may help the analysts get a first impression of how trustworthy this information is. The different color categories can be individually turned on or off. Furthermore, the exploitation percentage per information source for the whole period of time or for a time interval is given.

The spikes chart allows the analysts to make two different types of selection that impact the nodes ranking area and the graph analysis area: selecting a single spike, or selecting either a spikes aggregation or multiple spikes aggregations and/or even spikes (selected under a rectangle).

3.2.3 Ranking change analysis

The nodes ranking is explored, for a selected time interval, in the left part of the central area of the spikes comprehension framework (see Figure 3). The nodes ranking list can be sorted based on one of two following criteria:

- The ranking of the node (higher ranked nodes appear at the top – see Figure 5) : this view lists the nodes, their rank and an icon indicating whether the ranking increased (a green up arrow associated to the number of ranks gained), decreased (a red down arrow associated to the number of ranks lost), moved but came back as before or did not change at all; and
- The absolute value of rank change (nodes for which the change in ranking is large appear at the top – see Figure 3) : this view lists the nodes, their rank and a colored

ribbon of variable length under node and rank indicating whether the rank change increased (green) or decreased (red).

Additionally, this list can be filtered to reduce the number of nodes displayed. For example, nodes that have a ranking or an absolute rank change lower than a given threshold can be filtered out. In this latter case, nodes keep their rank when displayed. The time interval is selected by means of the spikes visualization chart, if there is no selected time interval, the ranking is given either for the last selected time interval or for the whole analysis period. At any time, the analyst can hide the list. The effect is that the space dedicated to the graph is enlarged.

As shown in Figure 3 in the network graph area, the selection of a node in the list highlights this node in the current displayed graph if the node exists. If the displayed graph is big in terms of nodes and links, gaze will be focused on the right node. If the node is not included in the current displayed graph, then the graph is replaced by the whole graph and the selected node is highlighted.

Figure 5 – Nodes listed by rank

3.2.4 Graph change analysis

The largest part of the display presents the network graph. The analyst is able to alternate between a global view of the graph or a subset containing only the nodes that changed – for

the specific spike under investigation – and their immediate surrounding network (up to a level selected by the analyst, by default this level is 1).

As mentioned previously, the spikes chart allows the analysts to make two different types of selection:

- Selecting a spike (directly in the spikes chart or from a spikes aggregation), by simply clicking on it, to explore its causes and effects:
    - The nodes' ranking list appears as it exists at the time of the spike;
    - A (sub-) graph reflecting the spike's meaning appears in the graph change analysis display. Possibility is given to the analysts to explore the graph by mouse hovering the node or the link to see relevant information (description box appears) and to explore a larger graph if necessary either by adding level(s) or by expanding node(s);
- Selecting a spikes aggregation – thus a basic time interval – by clicking on it or a specific time interval by selecting multiple spikes aggregations and/or single spikes to analyze the overall change:
    - The nodes ranking list appears as it exists at the time of the first spike (first time of the time interval);
    - A (sub-) graph reflecting the first spike's meaning appears in the graph change analysis display, and graph exploration possibility is given to the analysts as explained beforehand.

Nodes for which not all links are showed display a plus sign to indicate that they can be expanded. After expansion, nodes and links appear and a subset of these *new* nodes might display a plus sign. A node or group of nodes can be also selected to be hidden and collapsed. No matter which view is used, the analysts are able to expand, or collapse nodes, in order to browse the graph according to their interests. Of course, the analysts have the possibility to zoom in and out (from a specific node or for the graph displayed) and pan.

The graph highlights changes causing the spike – as well appearance and disappearance of node(s) and link(s) as changes in the node and link status – by playing both with the size and color of nodes, as well as the line type and the color of links. Mouse hovering a node in the graph causes the appearance of a summary card displaying all the relevant characteristics of a node visually using images and icons to represent the information (Lavigne *et al.*, 2014), as shown in Figure 6.

Select a spike leads to explain/demystify every aspect explaining its causes and effects: a node removed impacts node(s) and link(s) and possibly the nodes ranking; a link added (meaning node and link added) or a link removed impact node(s) that has(have) its(their) rank modified. Select a spikes aggregation (the basic time interval) or a time interval (1 to *m* spikes aggregations or *n* single spikes) can be considered as selecting a set of 1 to *p* spikes. Consequently, causes and effects of every spike are represented in the graph, as well appearance and disappearance of node(s) and link(s) as node rank changes. The analysts can go through the *p* spikes either one by one by controlling the animation from the spikes chart or

launch the animation to show changes propagation. Animation is reflected in the ranking display as well.

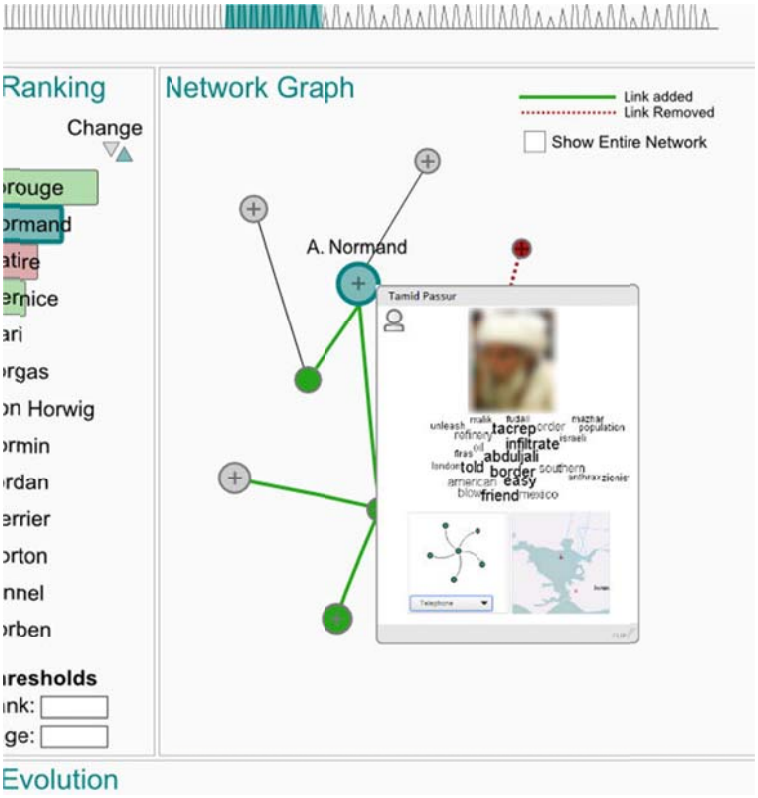

Figure 6 – Information about a node displayed as a summary card appearing when mouse hovering on the node

3.2.5 Nodes ranking evolution over time

This chart located on the bottommost part of the display (Figure 3 and Figure 7) is activated on demand (show/hide at any time). This chart displays the evolution of the ranking values for the selected node(s) either over the full time interval (starting/ending at the first/last time of the analysis or at the time a node appears/disappears in the network graph) or during the selected time interval from the spikes visualization chart. If a spike is selected or if no time interval is selected – no spikes aggregation and no set of spikes selected – then the full time interval is considered. When new node(s) is(are) selected and/or new time interval is defined, the chart is redrawn.

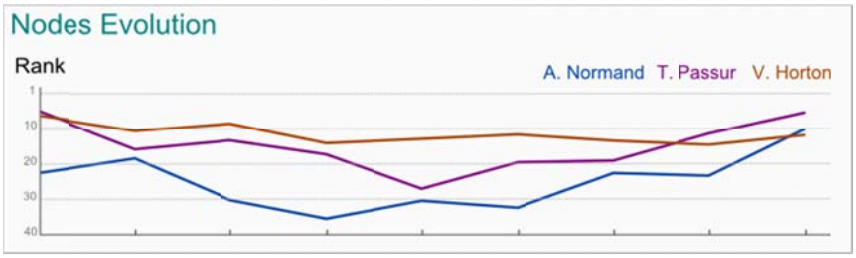

Figure 7 – Nodes evolution display

Individual node(s) can be selected in either the graph view or the ranking view and the selection will be reflected in the adjacent view. As soon as a node is selected and the evolution/comparison action is chosen, the nodes ranking evolution chart appears.

This evolution visualization will enable a trend analysis over specific nodes and comparisons between the selected nodes. A benefit of the evolution chart is that it could make the slow rise or decline of a node more salient. This kind of slow change is unlikely to cause a large spike.

4 Conclusion and Future Research

During recent decades, social components of military operations have been increasingly investigated for mission success purposes and specifically to respond to intelligence analysis requirements. While SNA has been gradually more valued, there still remains much to do, predominantly research on its adaptation to the military context. As previously expressed, there is a waterfall of several challenges faced when applying network analysis in military settings. Amongst others, some of these challenges are: the dynamic aspect of the collected data, their social nature, and the covert aspect of many networks of interest. These challenges result in the missing data issue and variability in network measures, which in turn generate a lack of trust being conferred to SNA measures.

The situation described above is of great concerns in an operational environment where on one hand SNA has an undeniable value for intelligence analysis, but on the other hand its related technics and methodologies still remain to be mastered in order to reach their full potential. The orientation taken within the Spike use case is to value measurements' versatility for the benefits of the analysts. Three critical attainments are made through the Spike analysis: first, it reinstates the "perception" stance that should be the one of the intelligence analyst; second, it behaves as an indicator of the possibility to apply the standard interpretation of SNA measures; and finally, it brings back the proper level of trust to be ascribed to the measures.

The research team has planned on conducting a number of activities in a near future. An initial activity consists in testing the observed results but based on a different data set in order to evaluate their homogeneousness. It will also be required to determine if Freeman's betweenness centrality measure could be combined to other analyses and network measures, and also if it is the most appropriate one to use in this setting. Subsequently, this analysis tool will be instantiated in the current sense making prototype along with the corresponding visual analytic components investigated in this research paper.

5 References

[1] Borgatti, S.P., Carley, K.M., and Krackhardt, D., (2006), *On the robustness of centrality measures under conditions of imperfect data*, Social Networks, Vol. 28, # 2, pp. 124-136.

[2] Endsley, M., (1988), *Design and evaluation for situation awareness enhancement*, In: Proceedings of Human Factors Society 32nd Annual Meeting, vol. 1.

[3] Freeman, L.C., Borgatti, S.P., and White, D.R., (1991), *Centrality in valued graphs: A measure of betweenness based on network flow*, Social Networks, vol. 13, pp. 141-154.

[4] Freeman, L.C., (1977), *A Set of Measures of Centrality Based on Betweenness*, Sociometry, vol. 40, #1, pp. 35-41.

[5] Howison, J., Wiggins, A., and Crowston, K., (2011), *Validity Issues in the Use of Social Network Analysis for the Study of Online Communities*, Journal of the Association for Information Systems (JAIS), vol. 12, #12, pp. 767-797.

[6] Lavigne, V., Lecocq, R., Martineau, E., and Mokhtari, M., (2014), *Graph Analyzer Widget Closer to Agility through Sense-Making*, 19th International Command and Control Research and Technology Symposium (ICCRTS), June 16th, Alexandria, USA.

[7] Lavigne, V., Lecocq, R., and Gouin, D., (2012), *Visual Apps for Counter-Insurgency Social Network Analysis*, Joint Symposium on Persistence Surveillance: Networks, Sensors, Architecture, NATO SET183-IST112, April 30th, Quebec City, Canada.

[8] Lecocq R., Martineau, E., and Caropresso F., (2013), *An Ontology-based Social Network Analysis Prototype*, Proceedings of the IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA 2013), February 26th, San Diego, USA.

[9] Martineau, E., Lecocq, R., (2013), *Automated extraction and characterisation of social network data from unstructured sources – An ontology-based approach*, Proceedings of the 18th International Command and Control Research and Technology Symposium (ICCRTS), June 19th, Alexandria, USA.

[10] Oxford Dictionary on line, (2015), DOI: http://www.oxforddictionaries.com, visited on February 2015.

[11] Patterson E.S., Roth E.M. and Woods D.D., (2001), *Predicting vulnerabilities in computer-supported inferential analysis under data overload*, Cognition, Technology & Work, vol. 3, pp. 224-237.

[12] Wang, D.J., Shi, X., McFarland, D.A., and Leskovec, J., (2012), *Measurement error in network data: A re-classification*, Social Networks, vol. 34, #4, pp. 396-409.