SMART DIAGNOSTIC TOOL FOR COMPLEX SYSTEM-OF-SYSTEMS

LEE Shie Yen Elaina, ONG Xiu Hui

ABSTRACT

Managing the Island Air Defence (IAD) System-of-Systems (SoS) is a challenge, given the magnitude of engineering effort required to integrate constituent systems from different technical domains (e.g. sensor, network, command and control, and weapon) into the SoS. To address this, a DSTA team developed a SoS Diagnostic Tool with centralised monitoring, root cause analysis, impact analysis and self-learning fault prediction to maximise IAD mission availability. It employs a fusion of Bayesian network and machine learning to equip operators with a tool which uses both engineering knowledge and lessons learnt to identify root causes. This article describes the background for the development of the tool, the methodologies and its potential adaptability for future SoS.

Keywords: system-of-systems, fault diagnosis, root cause analysis, Bayesian network, machine learning

INTRODUCTION

The Republic of Singapore Air Force (RSAF) Island Air Defence (IAD) System-of-Systems (SoS) comprises sensors, weapons and command and control (C2) systems networked together. While each system had its own monitoring tool, there was no centralised tool to provide a global view of the IAD SoS health status, and guide fault diagnosis in an expeditious way in the event of a complex failure[1].

The IAD SoS Diagnostic Tool developed by DSTA aims to ease troubleshooting and accelerate system recovery, thereby ensuring high SoS availability. The Root Cause Analysis (RCA) engine within the IAD SoS Diagnostic Tool aims to deliver a seamless decision-support solution for fault analysis under the context of incomplete information, allowing the tool to infer the most probable root cause from past failures.

COMPLEXITIES OF A SoS

The RSAF IAD SoS overcomes the limitations of disparate sensors and Ground-Based Air Defence systems by networking the C2, sensors, and weapon systems together (see Figure 1). It is designed and implemented to ensure a high level of robustness, resiliency and supportability to the IAD SoS.

The IAD SoS provides an integrated real-time air picture to support a multi-layered air defence by allowing operationally independent systems to work together and engage targets efficiently and effectively. This cannot be achieved by any individual constituent system.

However, the integrated nature of the IAD SoS makes it highly complex, especially when coupled with the constantly evolving new threat environment and the addition of new technologies and capabilities. System faults can potentially affect other downstream systems or result in unexpected behaviours. As such, speedy and accurate diagnosis, as well as the rectification of failures become even more critical to ensure mission success.

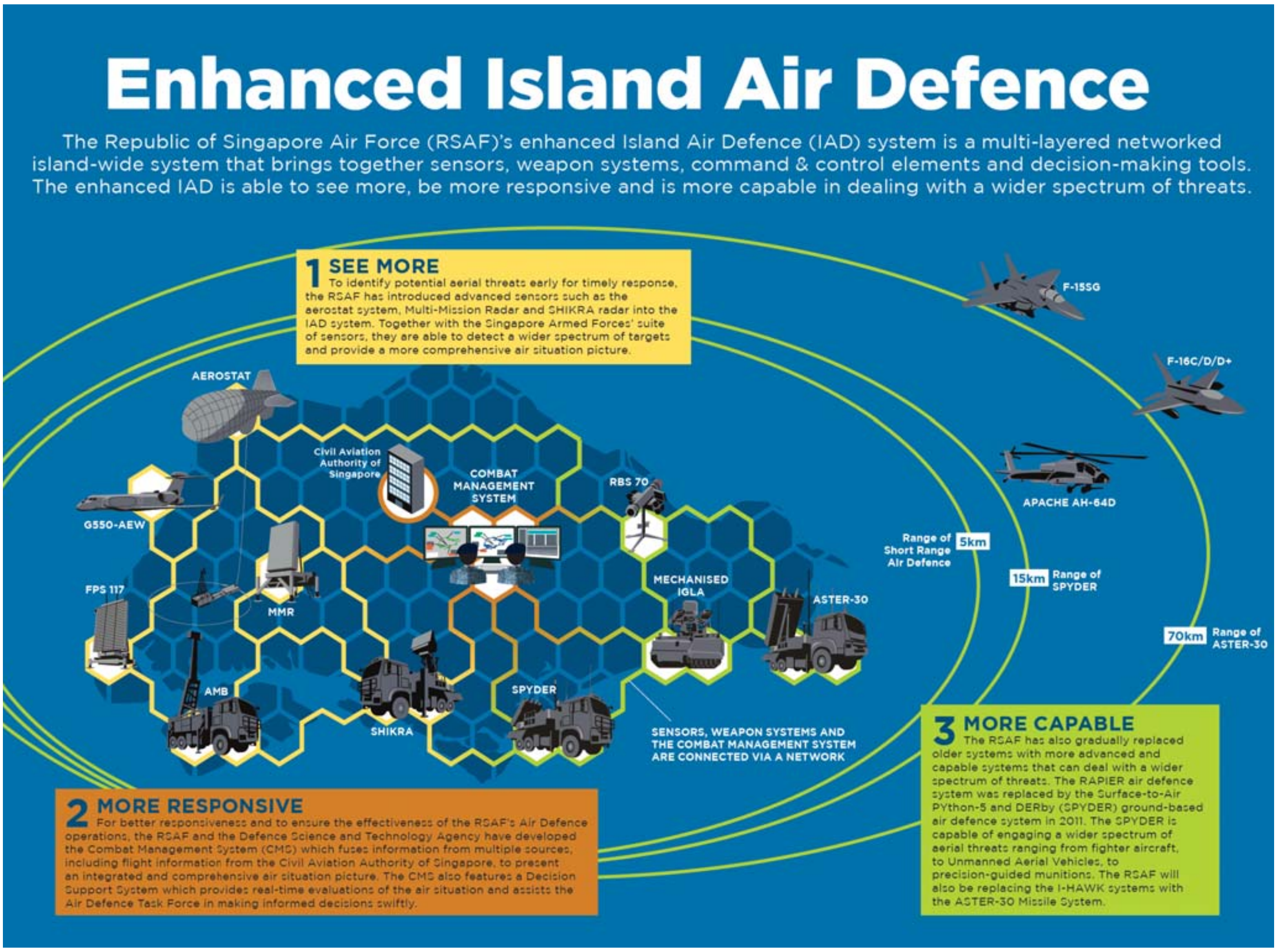

Figure 1. Island Air Defence (MINDEF, 2017)

CHALLENGES IN MANAGING A SoS

With the target of reducing workforce size despite the increasingly complex systems, there was an increased urgency to develop a tool to aid the RSAF in managing the IAD SoS while reducing the maintenance effort required.

One of the key hurdles faced was the limited pool of experienced maintenance personnel who could recognise and identify SoS interoperability abnormalities as well as analyse and rectify SoS faults. Without them, timely system recovery would be difficult.

SOLUTION

The IAD SoS Diagnostic Tool (see Figure 2) and the corresponding recovery process, based on manifested faults, complement the constituent systems' existing network monitoring systems (NMS). Through these NMSes, the IAD SoS Diagnostic Tool is able to collate and monitor the statuses of all hardware residing on the network that support the Simple Network Management Protocol. These include servers, workstations, network equipment and uninterrupted power supplies. DSTA then developed intelligent software to apply sensemaking on the information.

DSTA implemented the RCA engine to provide end-to-end fault diagnosis, assess the impact of each fault, and introduce a self-learning engine to fine-tune the results by incorporating past data and performing trending. With the smart intelligence built into the tool, even less experienced operators are able to diagnose complex SoS faults swiftly and call upon the right maintenance crew to recover the defective system.

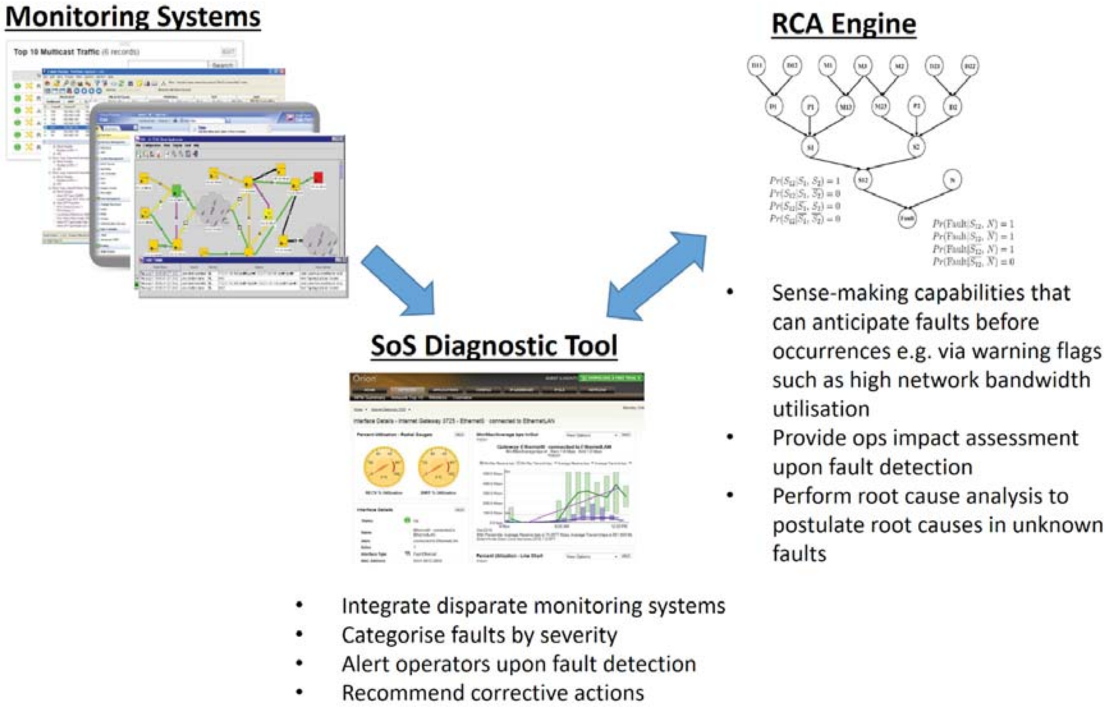

Figure 2. Functionalities of the IAD SoS Diagnostic Tool

Methodology

Traditional methods for reliability analysis, such as fault tree analysis, are not robust enough to handle complex systems with a large number of equipment. Typically, these methods also do not consider the frequency of failures and the interaction between different systems.

On the other hand, latest developments in deep learning and reinforcement learning that leverage neural networks and huge learning datasets heavily are not suited for troubleshooting or diagnosis of complex SoS failures where large historical failure datasets are not yet available.

Among the alternative approaches, the Bayesian network stands out as a useful tool for knowledge representation and inference under uncertainty. It has a wide range of practical applications, and has been implemented and deployed by the United States Army Communications-Electronics Command in the development of the Virtual Logistics Assistance Representative, an equipment diagnostic tool that enables a combat soldier to maintain critical equipment (Aebischer et al., 2017).

DSTA went a step further by integrating the use of a statistical Bayesian network model innovatively with machine learning to create an integrated knowledge-based diagnostic tool that not only improves its diagnostic capability, but can also be adapted as a teaching tool for operators and maintenance crew.

The Bayesian network model considers the conditional failure dependencies of the system design, and identifies root causes ranked according to their probability of occurrence. It adopts a top-down approach to problem solving, using pre-defined engineering knowledge of the system reliability and component dependencies.

Machine learning complements the model with a bottom-up approach. It correlates system behaviours beyond engineering factors such as human or environmental factors and configuration changes that are postulated to affect the diagnosis' outcome. It also stores a repository of knowledge where it self-learns from past diagnostic outcomes to sharpen future predictions.

This fusion of diagnosis techniques (see Figure 3) provides operators with a tool that uses both engineering knowledge and past lessons learnt to draw conclusions on the root causes, with minimal iterations and reliance on human interventions. This enhanced availability of the IAD SoS capabilities through faster insights into faults, resulting in a shortened recovery time.

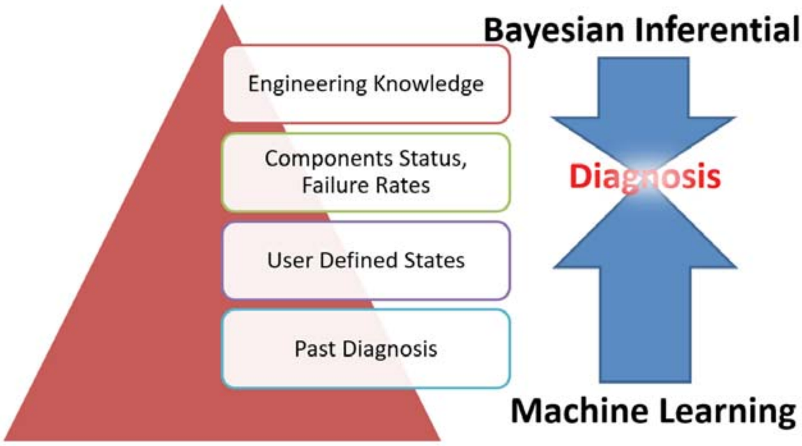

Figure 3. Fusion of inputs for diagnosis

SYSTEM ARCHITECTURE OF THE RCA ENGINE

The RCA engine comprises three key sub-engines: the Ops Impact Analysis Engine, the Bayesian Diagnostic Engine, and the Machine Learning Engine. Figure 4 illustrates its architecture view.

- The **Ops Impact Analysis Engine** employs a path-search technique based on graph theory developed using C# programming language. It draws inputs from live data of the network topology and component status that are monitored by NMSes such as SolarWinds. A change in the status of monitored components will trigger the engine to assess if the change would result in any critical ops impact. It then alerts the operators by displaying the possible component failures and associated ops impact at the entire SoS level, thus providing decision support to assess the need for any mitigating measures.

- The **Bayesian Diagnostic Engine** finds the root cause of system failures in cases where fault symptoms cannot be positively identified and resolved by the monitored components. Bayesian networks are modelled to perform causal inference, by considering both the network topology of the whole SoS and field Mean Time Between Failure rates of non-monitored components. The open-source Microsoft Research Foundation Class library is used to generate the probable root causes, which are ranked based on their probability of occurrence.

- The **Machine Learning Engine** works alongside the Bayesian Diagnostic Engine to improve the deduction of the fault. The engine implements the Random Forest approach using the open-source R library (a software environment for statistical analysis). Once the root cause has been positively confirmed by the maintenance crew, the tool will update the cause and the respective learning variables in the knowledge database to refine future diagnosis.

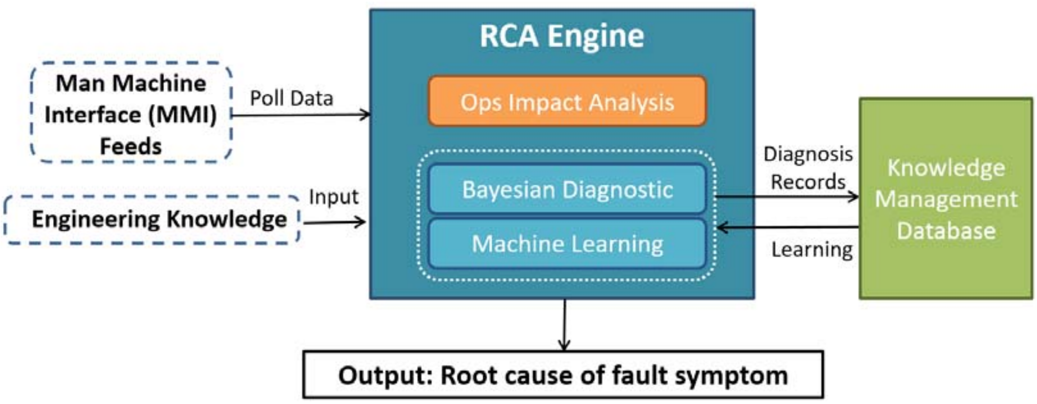

Figure 4. Architecture view of the RCA engine

Bayesian Network

The Bayesian network methodology is used to represent failure dependencies in the complex networked systems as a compact model with joint probabilistic relationships, where causal knowledge could be inferred for diagnostic assessment.

A Bayesian network is a directed acyclic graph which comprises nodes and arcs. The nodes of the graph represent variables of interest, and the arcs between pairs of nodes represent statistical or causal dependencies among the variables which are quantified by conditional probability.

The conditional probability distribution for a random variable associated to a node can be defined as the probability of each of its state conditioned on the combination of the states of its parent nodes, and the joint distribution of any set of variables $\{X_1, \ldots, X_n\}$ is given by the product

$$P\{x_1, \ldots, x_n\} = \prod_{i=1}^{n} P\{x_i \,|parent(x_i)\},$$

where x_i denotes some value of the variable X_i and the parent (x_i) denotes some set of values for the parents of X_i.

In the Bayesian network, a deductive process known as probabilistic reasoning is performed by instantiating the prior knowledge (evidence) that sets variables in known states, and propagating their effect through the network to derive the probabilities of interest, conditioned on this evidence (posterior probabilities). Bayes' theorem is used to update the probabilities:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad .$$

Bayes' theorem relates the conditional and marginal events of A and B. P(A) is the prior probability where there is no information about B, P(B|A) is the conditional probability of B given A, and P(B) is the prior probability of B which acts as a normalising constant.

For the RCA engine, the Bayesian network is synthesised from its Reliability Block Design (RBD) using a systematic mapping algorithm (Darwiche, 2010; Torres-Toledano & Sucar, 1998). Figure 5 explains how logic operators are used to translate each block in an RBD into a Bayesian network fragment with component states 0 (down) and 1 (up). Figure 6 shows a simplified reliability network and Figure 7 presents the Bayesian network after the conversion. This semantics is then used to compute the conditional probabilities of each component's availability.

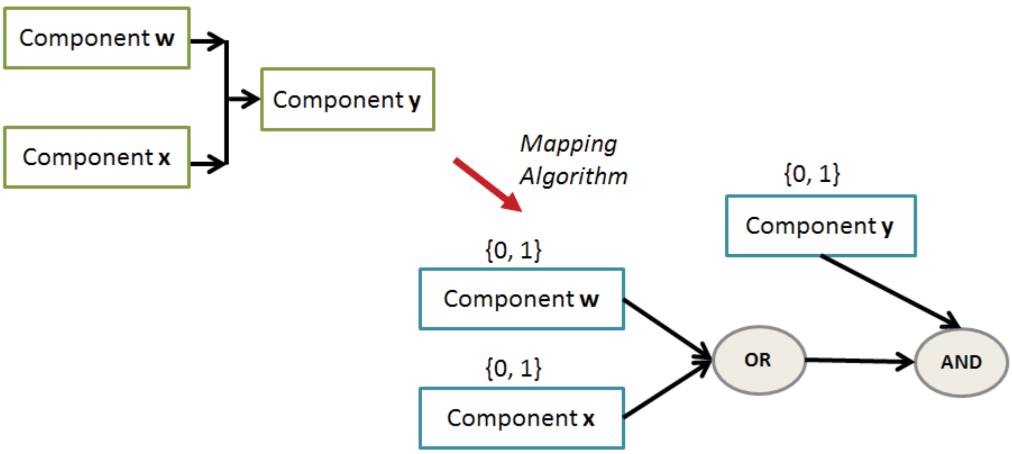

Figure 5. Mapping algorithm to translate RBD to Bayesian network

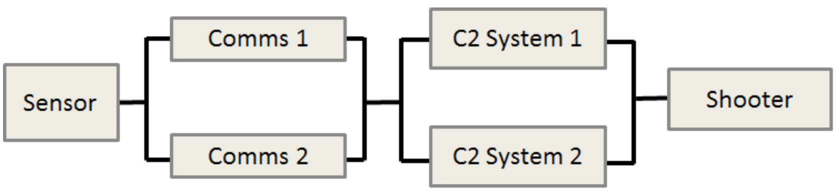

Figure 6. Simplified reliability network

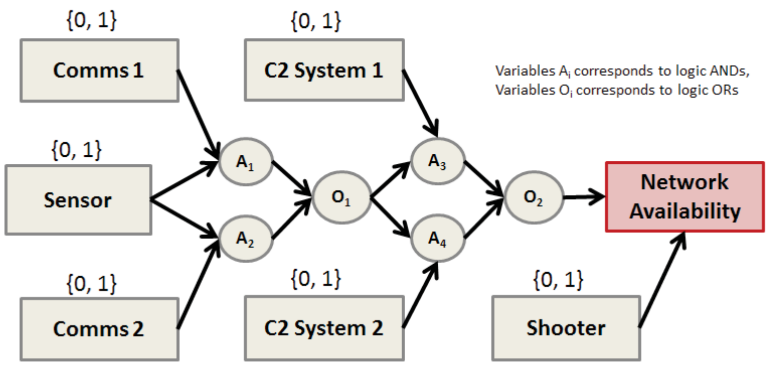

Figure 7. Bayesian network after conversion

Machine Learning

The integration of machine learning in the diagnosis brings about several advantages. Essentially, it allows the examination of system behaviours beyond engineering factors such as human or environmental factors. It also serves as a displayable repository for subject matter experts' knowledge, thus reducing the reliance on experienced maintenance personnel on the ground to perform fault diagnosis. In terms of code maintainability and reusability, it adds flexibility by allowing new possible factors for the fault diagnosis to be defined and edited, without the need to redesign the code. The machine learning engine provides a scalable solution by allowing new possible factors to be defined.

The RCA engine employs the Random Forest algorithm to predict the cause of failure from past failures. Random Forest is an ensemble learning method which generates a collection of tree-structured classifiers and aggregates their results (see Figure 8). Decision trees are constructed based on a pre-defined set of failure attributes (predictor variables). At every node of every tree, a subset of predictors is chosen randomly and the predictor variable that provides the best split, according to some objective function, is used to perform a binary split on that node. The prediction of new data is done by aggregating the predictions from all the trees in the ensemble. The ability of the algorithm to handle many predictor variables quickly and efficiently makes it an effective tool for prediction.

However, the disadvantages of machine learning techniques include the need for a sizable amount of data in order to self-learn, and its inability to handle rare failure events where there is no statistical significant historical data. This is where the Bayesian network, grounded on the engineering fundamental principle of the complex network, will complement the statistical inference.

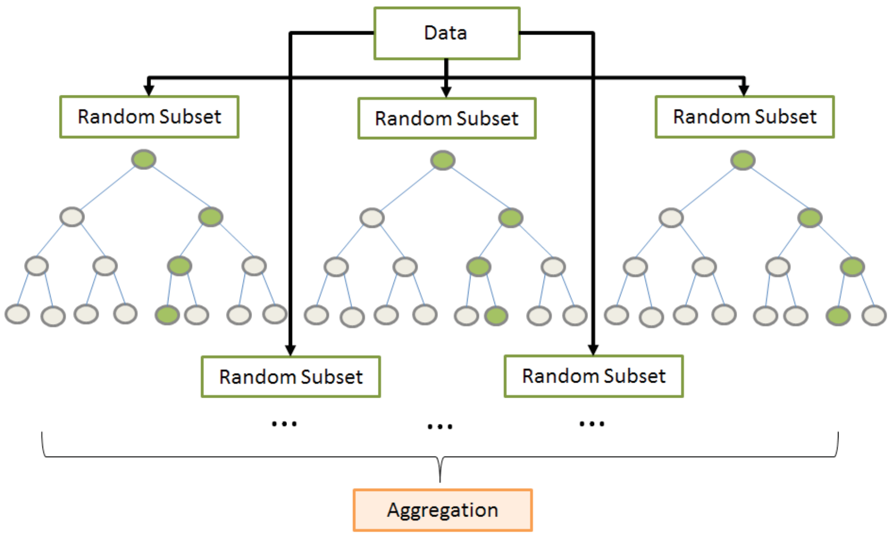

Figure 8. Demonstration of the Random Forest methodology

Benefits of Modularity

The RCA engine is scalable and can be adapted with minimal costs for other systems, ranging from individual military platforms to complex C2 networked systems. For example, other SoS in the Singapore Armed Forces (SAF) are reusing the RCA engine for their own diagnostic tools. As these systems adapt the RCA engine for their operating environment, they contribute to DSTA's experience and knowledge in this area.

FURTHER WORK

After the tool's deployment on RSAF premises in January 2018, the team expects a run-in period of up to three years to accumulate sufficient live data to further refine the RCA rule sets and enhance its machine learning capability. Upon completion of the run-in period, it may be possible to benchmark the results with other available tools for root cause analysis.

In addition, the current Bayesian network model only considers two states of the networked components: up or down. One extension of the model could be to model multi-state components with multiple failure modes (Zhou, Jin, Dong & Zhou, 2013). The consideration of varying degrees of failure will allow multiple effects of the system performance to be examined.

Beyond locating faults and short-term fixes, the systematic failure analysis may also bring about significant improvement in the aspects of manpower, machines, methods, material, measurement or environment. This could be done through rigorous analysis of the potential failure causes, removal of any non-conformance and prevention of recurring problems.

VALUE-ADD TO OPERATIONS AND SUPPORT

From the knowledge repository residing within the RCA engine, the adverse impact from the turnover of experienced maintenance crew members will be mitigated. At the same time, new operators and maintenance crew members will have a quicker pace of learning, with less reliance on tacit knowledge.

The RCA engine could potentially evolve into a training platform for maintenance technicians. By providing detailed explanations on how the results are derived, it could allow new operators to gain insights into how the SoS functions as a whole, and understand the intrinsic behaviours of the different systems that form the SoS.

Through the machine learning diagnosis, maintenance crew members can gain insights into the potentially weaker links within the complex SoS, and be able to design out or manage potential failures with forward planning, and enhance its reliability.

CONCLUSION

Development of the IAD SoS Diagnostic Tool has enhanced both DSTA and the RSAF's understanding of inter-dependency and inter-operability across each IAD constituent system and their ops impact on the overall IAD SoS capability. This knowledge can be applied when designing the architecture for future SoS systems.

DSTA has also built up competency in the development of smart diagnostic algorithms covering SoS fault inference, machine learning and data analytics. More experience could be gained through the potential application of the knowledge to other SAF systems.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to the RSAF for providing inputs and supporting the field trials, and the network experts from DSTA's InfoComm Infrastructure Programme Centre for their valuable advice for the integration of the tool unobtrusively with the operational systems. The authors would also like to thank Mr Seow Tian Beng, Ms Cheng Siew Yen and Mr Lester Seah Yong Huang for their guidance in preparing this article.

REFERENCES

Aebischer, D., Vatterott, J., Grimes, M., Vatterott, A., Jordan, R., Reinoso, C.… Hepler, A. B. (2017). Bayesian networks for combat equipment diagnostics. *Interfaces, 47*(1), 85-105. doi: 10.1287/inte.2016.0883

Conrady, S., & Jouffe, L. (2013). *Introduction to bayesian networks & BayesiaLab*. Franklin, Tennessee: Bayesia USA.

Darwiche, A. (2010). Bayesian networks. *Communications of the ACM, 53*(12), 80-90. doi: 10.1145/1859204.1859227

Ministry of Defence. (2017, June). *Factsheet: Protecting Singapore against air threats*. Retrieved from https://www.mindef.gov.sg/imindef/press_room/details.html?name=30jun17_fs5&date=2017-06-30

Torres-Toledano, J. G., & Sucar, L. E. (1998). Bayesian networks for reliability analysis of complex systems. *Proceedings of the Progress in Artificial Intelligence – IBERAMIA 98*, AI Lisbon, Portugal, 195-206. doi: 10.1007/3.540.49795.1

Zhou, Z., Jin, G., Dong, D., & Zhou, J. (2013). Reliability analysis of multistate systems based on bayesian networks. *Proceedings of the 13th Annual IEEE International Symposium and Workshop on Engineering of Computer-based Systems (ECBS'06)*, Potsdam, Germany, 344-352. doi: 10.1109/ECBS.2006.58

ENDNOTES

[1] Complex faults refer to faults manifesting in one of the constituent systems, caused by abnormal system behaviours from another constituent system.

BIOGRAPHY

LEE Shie Yen Elaina is a System Manager (Systems Management). She joined the Island Air Defence (IAD) Team as a software developer for the IAD Combat Management System (CMS) and led in the CMS safety assessments before moving on to oversee the operationalised IAD CMS. With her experience and involvement in the different phases of IAD capability development, she played a leading role in the assessments and analysis of IAD System-of-Systems behaviours. Elaina graduated with a Bachelor in Computing degree from the National University of Singapore (NUS) in 2001.

ONG Xiu Hui is a Senior Analyst (DSTA Masterplanning and Systems Architecting) and is currently the primary analyst maintaining and developing features to enhance the simulation capability of Optimised Decisions in Networks. She graduated with a Bachelor of Science degree with Honours, with a major in Mathematics and a minor in Statistics from NUS in 2008.