# Chapter 11
# The Need for Trusted Autonomy in Military Cyber Security

**Andrew Dowse**

## 11.1 Introduction

Information systems in the early 21st Century have become a critical enabler of increased value to the business, or as people in Defence might call a 'force multiplier'. Clearly the converse of this logic is that in warfare any capability that provides such a competitive advantage is also a vulnerability and a focus for a potential adversary to target. In the 20th Century, this risk was mitigated through the isolation of our information systems, with closed systems inherently easier to protect. However the real value of modern information systems has been the ability to provide more accurate, complete, relevant and timely information to support the business; and this has been achieved through a trend towards openness with greater connectivity and integration of systems. The very source of value to the business also represents a risk to it, and this remains a matter of tension and deliberation in the management of information systems.

The importance the Australian Department of Defence places in protecting our information advantage is reflected in the 2016 Defence White Paper, which notes the emergence of cyber threats to the ADF's warfighting ability, given its reliance on information networks [7]. The White Paper states that national and Defence cyber security capabilities will be strengthened to protect our systems.

A simplistic response to this priority would be for Defence to put more resources towards cyber security: more people monitoring audit logs and gateways, more effort towards accreditation and assurance activities, more funding allocated to cyber security projects. However, the exponential growth in the information environment means that taking a traditional approach and providing linear increases in resources is unlikely to meet the emerging challenge.

A. Dowse (✉)
Department of Defence, Canberra, Australia
e-mail: andrew.dowse@defence.gov.au

H. A. Abbass et al. (eds.), *Foundations of Trusted Autonomy*, Studies in Systems, Decision and Control 117, https://doi.org/10.1007/978-3-319-64816-3_11

This paper will consider potential requirements for trusted autonomy in cyber security, looking firstly at the current cyber environment, including four fundamental principles of cyber security. It will then assess the emerging challenges to this mission, framed though the dimensions of Big Data and consider opportunities to apply trusted autonomy to improve cyber security. The intent of this paper is to help inform researchers of the areas in which development in trusted autonomy may provide greatest return on investment in cyber security. Whilst these areas are specifically related to the requirements for the Australian Department of Defence, they may also be relevant to many other organisations facing similar cyber challenges.

Defence's Information and Communications Technology (ICT) architecture provides a reasonably robust protection against cyber threats. The lower classification (Protected) network is connected to the Internet via a gateway that provides multiple security mechanisms, thus achieving defence-in-depth. These security mechanisms are highly effective and relatively sophisticated, but involve significant manual processes. Due to the sensitivities and the need to maintain a security advantage, this paper will not provide any details of the tools or techniques currently utilised by Defence.

Whereas cyber security threats are on the increase, incidents on Defence networks actually decreased in 2015 in comparison with the previous year. Some 50,000 events were detected in 2015, around the same as 2014, of which there were 580 incidents, which represented a 25% decrease [10]. The causality of the decrease in the number of security incidents cannot be stated with certainty, but there are strong indications that this result is through greater success of security mechanisms, especially through blocking of threats at the gateway. Notwithstanding the evidence of current success in cyber security, Defence needs to further strengthen protections to keep up with, and preferably ahead of the threat.

## 11.2 Cyber Security

Information assurance and cyber security are both concerned with the protection and defence of information and information systems by ensuring their confidentiality, integrity and availability. Information assurance accounts for the risks to information from natural, accidental and deliberate actions. Whereas cyber security tends to focus on deliberate acts, information assurance and the managers of information systems need to prepare against such acts, but also against accidents, faults, external events and human error [4].

If the overall outcome for ICT management is the preservation of confidentiality, integrity and availability of information in support of the organisation's missions and interests, in many respects it doesn't matter whether something happens due to a deliberate act or some other reason. When there is an impact on an organisation's information systems, the priority is to respond coherently and expeditiously, rather than dwelling on whether it is an attack or a fault before someone responds. Hence a principle of cyber security is that it is an integrated part of how an organisation manages its ICT environment. The organisation needs to have clear accountabilities for ICT security, from policy to accreditation to day-to-day operations. As exciting as the idea of doing cyber operations might seem, defensive operations are largely a matter of systematically reviewing candidate incidents and managing the various security mechanisms within a defence in depth approach. The ability to successfully undertake defensive operations is strongly dependent upon how well the information environment is set up and the level of discipline inherent in it. Hence the second principle is that a secure foundation is fundamental to cyber security.

The design of Defence's information environment provides requisite levels of information assurance. In addition, Defence must undertake activities that ensure that systems perform consistently as specified, that users are accountable for their actions, that risks are mitigated by monitoring the environment and reducing the impact of a failure against system or usage expectations, and that there are means available to support an effective and timely recovery from an incident. These are the foundations that need to be designed into a secure environment, and need to be continually reviewed, updated and validated as technologies and threats evolve.

Much of Defence's information assurance against cyber threats comes from application of the Australian Signals Directorate (ASD) mitigation strategies. The top 4 mitigations—application whitelisting, application patching, operating system patching and restricting privileged access—can prevent over 85% of cyber intrusions [6]. Defence also gives priority to ASD's larger list of 35 mitigations, which further reduce vulnerability to cyber threats.

Defence will further enhance cyber security mechanisms with investments through Joint Project 2068, the Cyber Security Improvement Program, and increasing the recruitment and training of cyber security specialists, with a mix of military, public servants and contractors. Defence is also supporting Whole of Government efforts through the expansion of the ASD-led Australian Cyber Security Centre.

Defence's adoption of current operating systems has been slow in the past, and this leads to vulnerabilities associated with using older systems. The Infrastructure Transformation Program, which is planned to deploy by the end of 2017, will update hardware, networks, operating systems, applications and architectures to make Defence's ICT more robust, supportable and defendable [5].

But no matter how good these systems are, cyber security can only be as good as the organisation's people make it. This outcome is not only reliant on cyber security specialists, but requires the support of all people in the organisation who use ICT. Poor discipline, stupidity, lack of awareness and deliberate acts are all critical risks to cyber security. The third principle is that everyone in the organisation contributes to cyber security.

Some aspects of Defence's information environment are already constrained to mitigate the risks of poor user behaviours. Risks could be minimised further by locking down systems, but it gets to the point that it impacts the business; in which case it is better to accept some level of risk, and perhaps mitigate through training or auditing or some other mechanism. Thus the fourth and final principle is that an organisation must balance security imperatives with business requirements for functionality and access.

There are different expectations for cyber security for different information services. For Defence's classified warfighting network, there is an expectation of a high level of availability and confidentiality, whereas Defence's financial systems on the Protected network require high integrity.

Hence Defence's classified networks utilise greater access controls and adherence to the imperatives of need to know and need to share. Their connectivity is generally only to equivalent domains, and access points are strongly monitored. The Protected network has connectivity to the Internet, but through a consolidated gateway that has multiple security mechanisms in place.

Defence faces different types of cyber threats for its different networks. For classified networks, there is a focus on ensuring availability, as well as protecting against insider threats and to guard against potential intelligence collection. On the Protected network, the most common threat is criminal and there is less of a concern with attacks against confidentiality, although there are risks with commercially sensitive information as well as the real prospect that aggregation of information is sufficiently valuable to attract sophisticated state-based intelligence collection threats.

As an example of managing the balance between functionality, access and cyber security, Defence has in the past only permitted purely unclassified emails to exit the gateway from the Protected network to the Internet, and did not allow emails with a Dissemination Limiting Marker (such as sensitive or For Official Use Only) to be passed to the Internet. While reducing risks, this practice was damaging Defence business, such as vetting processes and interaction with Defence industry. Therefore a risk based decision was made earlier this year to permit such emails to be sent to the Internet when justified for Defence business, with risks managed through user awareness, procedures and auditing.

This highlights an important point here that much of Defence's information is actually held outside controlled networks. With the Defence Industry Policy Statement intent to strengthen these industry partnership arrangements, sensitive military information held on industry's networks must be protected to the same level as on Defence's own networks. In this regard, industry is critical to Defence's cyber security and creates an additional complexity to how we might use trusted autonomy.

## 11.3 Challenges and the Potential Application of Trusted Autonomy

The evolution of computational techniques has taken us from automation of processes, in which a system acts in accordance with defined rules, to autonomy, in which its behaviour is governed more by an understanding of objectives combined with observation and learning. Autonomy is a characteristic of an agent in which it is aware of other entities, and interacts with them, but exercises independence in order to maintain focus on its defined interests. Key to autonomy is the interaction with the environment within a goal-directed behaviour [12]. Such independence of action is further defined as self-direction or self-governance [14].

Trust is a further characteristic in which the agent will act in a predictable and reliable manner, producing credible outcomes based upon the use of reputable sources [11]. Trust also has a connotation of the formal evaluation of systems to determine how they behave with such predictability and certainty, especially where it applies to the protection of confidentiality. This requirement in trust raises an interesting dilemma for an autonomous system, in that it may be difficult to measure predictability in such a system that does not act in an obviously deterministic manner. Trusted autonomy may in this regard be considered an oxymoron, or at least a challenge for developers and researchers.

In cyber security, trusted autonomous agents should provide reliable security outcomes that align with the interests of the organisation. Given the very nature of cyber security, there is an expectation of a significant level of trust in any agent involved in the protection of networks. The need for trusted autonomy is being driven by a number of factors or challenges in the future information environment, which will be explored in the remainder of this paper.

Fifteen years ago, the concept of Big Data was introduced, characterising the concept in terms of three 'V' dimensions [13]. Since then other authors have added more dimensions, typically continuing the alliteration, and also it has been recognised that many of these dimensions are relevant for cyber security. In this paper I'm going to examine Defence's future cyber security challenges in terms of five Vs.

The first V is volume. CIO Group manages multiple networks on behalf of Defence, the largest being the Protected network with over 100,000 users. The personnel system runs over 100 million transactions per week and the logistics system over 10 million transactions per week. Utilised storage in the Defence environment is 5.8 petabytes, with an annual growth of around 20% [9].

The Defence network architecture is designed with consolidated gateways that protect the corporate network but enable controlled access between it and the Internet.[1] This approach provides a security focus on the gateway, reducing the risks of multiple vulnerabilities, but brings a significant volume of interactions across the gateway.

Each week, Defence's High Availability Internet Gateway supports around 2 million inbound and 600 thousand outbound legitimate emails, as well as nearly 10 terabytes in web services [8]. While legitimate traffic continues to increase, it is negligible compared to blocked emails, which have gone in a period of 12 months from roughly the same quantity as legitimate in 2015 to now four times as many and increasing. While a portion of the blocked emails are due to being oversize or misaddressed, the majority are spam or potentially have malicious content.

Clearly dealing with such large and increasing volumes of data is a challenge for cyber security effectiveness. Much of this effort is focused on perimeter security at the Internet Gateway, and Defence's systems are highly capable of identifying suspicious events. However the analysis of candidate incidents involves humans in

[1] The importance of gateways in cyber security is elaborated at [3].

the loop, and increasing volumes of events will create a challenge that needs to be met through a combination of increased resources and automation, if not autonomy.

Cyber security risks however are not only concerned with incoming traffic. The increasing volumes of data exiting through the gateway to the Internet are monitored, with mechanisms to block, flag or log emails dependent on the content and other circumstances.

Additionally, the expansion of the Internet, especially with introduction of IPV6, translates to additional volume for setting rules at the gateway in respect of whitelisted and blacklisted entities.

While there are tools that support distinguishing between valid and potentially suspicious traffic, much of the actual decision making around release of traffic remains a manual process. While automation reduces the volumes of data necessitating a manual process (e.g. from events to candidate incidents), the growing volumes and sophistication of threats mean that there are growing numbers of unfiltered events that require trusted decision-making.

Hence a key future requirement for trusted autonomy is to further increase the ratio of total cyber events to those that require manual analysis. This will require increasing the trust not only in the agents that provide that filtering but also in the sources that the agents rely upon to undertake this task. Success in this endeavour to handle large volumes of data is a matter of both the defeat of cyber threats as well as the facilitation of valid business.

Although much of the cyber security emphasis is on the gateway, Defence recognises that perimeter defence is not enough for comprehensive cyber security, and endpoints around the network are monitored and analysed accordingly. Defence employs some automation to support analysis of this data, but its intent to strengthen cyber security requires further enhancement in this area, including more sophisticated and autonomous agents that can recognise anomalous behaviours and events.

This leads to a second V: visualisation. The ICT security capability in Defence has utilised pioneering visualisation technologies for some years. In order to have awareness of the health of the information environment and be able to make timely decisions, whether about cyber security or any aspect of operations, an organisation needs better visualisation. This is a challenge for Defence, especially in the future with its transformed network infrastructure managed within separate towers by outsourced service providers.

The future visualisation capability needs to provide consistent information in an operating picture that reflects the information and physical domains, and that can be shared between security operators, network operators and military command and control. Defence needs visualisation for decision support in the form of relevant and accurate information about the status of networks.

Relevant is important in the context of what information is needed to make decisions—a strategic commander will require different information from a tactical commander, and quite different again from network and security operators. Military commanders will be interested in whether systems are available, fit for purpose and are providing the required connectivity and functionality. This will demand a different approach to visualisation compared to network and security operators, who will be more concerned with the systems themselves rather than the businesses they support. In this respect, the term ‘common operating picture’ or COP is misleading as the picture is not the same across all the types of decision makers—so what is required is a consistent operating picture.

The difficulty of a cyber COP is that it is far more difficult to represent the cyber environment, and be able to be comprehended, compared to the comparatively simpler representations of the physical environment. Added to this is the critical importance of providing accurate information to a decision maker who will make potentially life or death decisions based upon the state of cyber support.[2] Therefore it is important that decision support systems provide an accurate representation of the cyber environment.

This in itself is more of a challenge than might be immediately evident. A system that provides situational awareness of the cyber environment is itself part of the cyber environment. Any system that provides such awareness must do so with credibility and reliability, having access to sources that provide this information, but having sufficient resilience so not to be vulnerable to the threats or faults of which it is providing awareness. Hence visualisation provides a second critical requirement for trusted autonomy in cyber security.

The visualisation system needs to facilitate users’ ability to drill down to get more information. It also needs to provide timely advice and support decisions to reconfigure as required in real time.

This leads to the third V: velocity. As per the case of Big Data, many business applications require real time speed for their interactions. Whereas an email or file can be taken offline for analysis, other interactions such as web services may not be so easily managed from a security perspective.

Another consideration for velocity is the critically short period between the first awareness of a new threat and the deployment of associated defence mechanisms such as detection and patching. Defence enjoys excellent relationships with partners in other CERT like organisations and in industry, and is very focused on minimising the time between identification of new threats and deployment of adjusted defences.

This may not be as helpful if the organisation’s networks are the target of a zero day attack, hence a lot of importance needs to be placed on security mechanisms that help identify new threats, limit the damage and facilitate quick recovery and support of business.

It is important to recognise the pace at which a significant cyber incident can evolve. Defence is placing greater emphasis on the coordination procedures for reacting to a cyber event and making timely and appropriate decisions on how we balance operational continuity and security. To do so demands an understanding of who the authorities for making such decisions are and to exercise realistically so that responders aren’t trying to figure it out in the middle of a real incident.

Such initiatives will help improve the ability to respond to cyber threats in a timely manner, but so long as these procedures are manual, Defence may not be able to keep

[2] This might seem overly dramatic, but Defence is now highly reliant on information systems, and a decision to sever or shut down systems due to a cyber threat may have significant consequences.

up with the tempo of cyber operations. Current procedures in Defence seek to resolve network issues in terms of hours, if not days. The consequences of cyber events can be a matter in which decisive action is needed in much shorter timeframes.

It is useful to consider the velocity issue in terms of Boyd's Observe-Orient-Decide-Act (OODA) loop [2]. Boyd identified that in modern warfare, particularly air warfare, an advantage would be gained by having a shorter decision cycle than one's adversary. Such an advantage is even more critical in cyber warfare.

The third critical requirement for trusted autonomy in cyber security is to streamline decision-making to minimise the time to take action. As a minimum, this requires superior decision support, as discussed earlier under visualisation. Timeframes can be reduced if the autonomous agent provides recommendations based upon a comprehensive knowledge of the network, the cyber threats and consequences, and the supported business.

To fully comprehend the need for timely and integrated decision-making in cyber security, it is important to appreciate the nature of our vulnerabilities. The Defence network architecture focuses perimeter defences on areas of greater vulnerability. Defence is creating greater access in its networks to information and services, with deployable and mobile users soon having similar access to services as one might have in an office in Canberra. The centralising of data processing and making greater use of thin client technologies may represent a shift in vulnerabilities.

The seamless integration of the Defence ICT environment means that the organisation cannot afford for localised and independent decisions about balancing risk and reward as it pertains to cyber security. A violation in one area can conceivably proliferate throughout the network, which demands an approach to configuration management, cyber security, technical control and support to military operations that can balance overall risks. Often the risk to a mission needs to be weighed against the risk to the enterprise. This needs to be taken into account in developing the ability, including autonomous capability, to respond quickly to cyber threats.

The primary intent therefore is to ensure vulnerabilities are minimised and responses are quick, coherent and effective when facing cyber incidents. Defence leadership must recognise the potential that a sophisticated and strong cyber threat could impact its systems, and therefore must be prepared through training to fight on in situations of disrupted or degraded information services.

One might suggest the ultimate goal for trusted autonomy in cyber security is to take the human out of the loop in defensive cyber operations, making the speed of response dependent only on electronic processes, rather than also including physical and cognitive elements. Such a goal would require considerable investment in research and development, as well as a very highly refined (and continually updated and tested) understanding of the relative risks (to missions, systems and security). Right now the level of trust required for this goal does not seem within reach, but it is certainly a fertile area for future research and development. As mentioned earlier, this raises the interesting question of predictability and whether truly trusted autonomy is even possible.

The fourth V is variety. Like the substantial variety of data Defence has in operating a rather complex business, there also is a great variety of cyber threats that it needs to protect against. For externally sourced threats, this challenge of variety of threats is addressed with a variety of defences. This is a similar approach to Ashby's Law of Requisite Variety [1], in that in cyber security we use multiple tools and multiple sources to help increase the likelihood of 'catching' the different threats.

The bulk of the day to day cyber threats Defence deals with, in terms of pure numbers, are about unsophisticated criminal scams. The organisation must be on its guard to deal not only with these prevalent threats, but with less common threats. The bigger concern is about seeing and dealing with the threats that aren't so obvious and are more dangerous, such as sophisticated malicious code, or the exfiltration of information by such code or by a trusted insider.

Whilst such events may be identified through monitoring, Defence's cyber security approach tends to focus on known threats such as through signature matching. In addition to these mechanisms, better systems need to be developed that characterise the normal environment and can effectively and responsively identify anomalies. Such a capability will help protect Defence networks against the unknown threats. Thus the fourth critical requirement for trusted autonomy in cyber security is to help in the identification of potential cyber threats through monitoring of anomalous activity.

The fifth and final V challenge is variability. Here I diverge from the Big Data view and consider variability more in a macro sense of the word, and this has several dimensions.

Whereas Defence embraces the concept of a Single Information Environment, in reality there are a lot of networks within the Department that are managed by individual business units, and have variable adherence to security requirements. Defence is working to remediate and accredit these networks to reduce vulnerability. Defence is also considering the introduction of a cyber-readiness or cyber-worthiness regime, to regularly test the security of Defence's ICT networks.

Another aspect of variability arises at the application layer, in that the Single Information Environment comprises different applications, and versions of applications, that largely do similar things. The inertia in moving on from legacy applications creates a management burden and results in security risks in operating unsupported systems. Defence continues to work on the rationalisation of legacy systems through the Infrastructure Transformation Program.

Like the Internet and many organisations' systems, Defence networks also have a lot of outdated content. Such 'untidiness' of the environment can impact productivity and is also a security risk. The Enterprise Information Management initiative of Defence's First Principles Review is endeavouring to address this problem.

There are arguments for and against whether variability within the environment contributes to or detracts from cyber security. Some might suggest that variability of systems reduces the impact of an exploit against a particular system. However, my belief is that a more consistent, tidier and disciplined environment is easier to support and to defend.

Another aspect of variability is how critical each application and information service is to Defence business. This then translates into variability for the redundancies, disaster recovery and incident resolution priorities that apply to each information service. Such requirements have been established for all the services and systems Defence supports, with a view to their continuity in the case of a fault. Defence leadership will have to consider in future whether these priorities are right, particularly in respect of recovering from a substantial cyber event. This will come from engagement with Defence's Groups and Services, as well as through exercising and wargaming of cyber events.

One last consideration for variability, and this to me is the most important, is the variability of the organisation's people when it comes to cyber security. Defence requires that its people have a standard of behaviour and awareness that adds to our defence in depth, rather than being a weakness. Despite having standard training for cyber security, practically a range of behaviours can be observed, from cautious to reckless.

So what implication does variability have for the need for trusted autonomy? As per previous discussions about anomalous behaviour, there is a need for improved systems that identify when actions, activities or attributes of the system are unexpected, and potentially to take action to mitigate risk. Additionally, there is a need for a sophisticated understanding of Defence's business, specifically an appreciation that the criticality of services varies across the organisation and thus affect the balance of risks in undertaking cyber defence. Trusted autonomy could contribute to information management and cyber security as a compliance agent, by monitoring the environment, and identifying and analysing variance—thus helping maintain our cyber readiness.

## 11.4 Conclusion

Right now is a very interesting time to be involved in cyber security, especially in Defence. The Department has a clear direction to strengthen cyber security, whilst at the same time needing to improve the functionality and the accessibility of information services. This demands creation of a solid foundation for information assurance and then cyber security and operations personnel must manage the balancing act of risks and value.

I have outlined the principles of cyber security and future challenges, with a bit of alliteration borrowing from the Big Data V concepts. It is important to recognise that traditional approaches to address these challenges will not be enough. Specifically, future cyber security will need to deal with exponentially growing volumes of information, need to have better awareness of the cyber environment, need to respond quickly to cyber events, need to identify unknown threats and have the ability to understand our complex environment.

Success in meeting these challenges requires a competitive advantage, which most likely will only come with assistance from trusted autonomy. Given the development of trusted autonomy has not to date met expectations, greater investment into research and development in these areas is important in order to keep ahead of cyber threats in future.

## References

1. W. Ashby, *Ross: An Introduction to Cybernetics* (Champman & Hall, London, 1956)
2. J.R. Boyd, A discourse on winning and losing. *Air University Document MU43947, Briefing*, 1, 1987
3. DoD, Information Security Manual (2015)
4. DoD, *Australian Defence Force Publication 6.0.3 Information Assurance*, 2nd edn. (2016)
5. DoD, CIOG presentations at DefenceWatch Briefing, 8 Sept 2016, Commonwealth Club Canberra, 2016
6. DoD (2016) http://www.asd.gov.au/infosec/mitigationstrategies.htm
7. DoD (2016) http://www.defence.gov.au/whitepaper/docs/2016-defence-white-paper.pdf, 20 Dec 2016
8. DoD, Internal Defence Statistics, sourced from Defence Strategic Communications Branch (2016)
9. DoD, Internal Defence Statistics, sourced from Enterprise Technology Operations Branch (2016)
10. DoD, Internal Defence Statistics, sourced from ICT Security Branch (2016)
11. M.N. Huhns, D.A. Buell, Trusted autonomy. IEEE Internet Comput. **6**(3), 92 (2002)
12. U. Krogmann, From automation to autonomy-trends towards autonomous combat systems. Technical report, DTIC Document, 2000
13. D. Laney, 3d data management: controlling data volume, velocity and variety. META Group Res. Note **6**, 70 (2001)
14. W. Truszkowski, H. Hallock, C. Rouff, J. Karlin, J. Rash, M. Hinchey, R. Sterritt, *Autonomous and Autonomic Systems: With Applications to NASA Intelligent Spacecraft Operations and Exploration Systems* (Springer Science & Business Media, Berlin, 2009)