New Wave Cyber Attacks

Angela Mison, Gareth Davies and Peter Eden
University of South Wales, Treforest, UK
Angela.Mison@southwales.ac.uk
Gareth.Davies@southwales.ac.uk
Peter.Eden@southwales.ac.uk

Abstract: There is increasing enthusiasm for, and recognition of, the benefits that artificial intelligence (AI) can provide to society. The emphasis has been on the positive, but AI and deep learning can be used for negative purposes. Modular Neural Networks (MNN) are capable of independent learning and have been targeted at evolutionary, complex financial systems. If the goal of an MNN were to be defined as system penetration, there is no reason why an algorithm could not run in the background. There are a resource requirements, but organised crime groups, technology companies, nation states and individuals with a curious bent are all capable of such. Ordered society and security requires a degree of certainty that systems on which society depends will remain recognisable, dependable and resilient. Under current conditions, security is difficult enough. It is suggested that limitations may be required before release of certain AI systems, in the knowledge of their potential for detriment to society. An AI system capable of independent learning, permits undefined emergent behaviours. That the results of any emergent properties may be benign or malign is irrelevant. Scientific history is littered with developments whose uses were redirected away from the benign.Such concern could be interpreted as fear of the unknown, standing in the way of technological advances. Unless society wishes to become machine-driven, the power and control of systems should be defined and limited by society, not accidently sprung on humanity or based on a ruthless logic that may drive a system to an unacceptable conclusion. Currently there are sophisticated botnet forming methods ensuring botnet persistence. If combined with the concepts of AI, there is a possibility that botnets could exist in perpetuity, with no one able to predict emergence, and no time limits on evolution. Whither cyber defence in the face of the unstoppable, increasingly intelligent, goal directed systems?

Keywords: Goal directed AI, Bias, Emergent properties, Ordered society, Botnets, Persistence

1. Introduction

> *AI is an inspiring technology. It will be the most powerful tool in generations for benefitting humanity…..AI systems will also be used in the pursuit of power. We fear AI tools will be weapons of first resort in future conflicts.* (National Security Commission on Artificial Intelligence, 2021)

Whilst the above is written from a Defence perspective, it is equally applicable to all systems as technological advances and research enable more attack techniques to move from Research or well-resourced Organised Crime Groups (OCG), to the DarkMarket, Crime as a Service and zeroday exploit corporates such as NSO Group and L3 Technologies. Indeed, the Final Report [*op cit*] even goes so far as to provide guidelines on how to hack AI systems.

With the exception of some command economies, in recognition that many organisations have neither the skill nor the capacity to undertake AI research or its implementation, it is anticipated that both protagonists and defenders will avail themselves of the talent within their respective equivalents of Silicon Valley or specialised software developers. This approach disregards the necessary aspect of expert customer and thus affects the specification of systems required, control of inherent bias, integration, the system roll out, policies and governance, and maintenance. Habitually, the person who understands the nuances, implications, and potential consequences is not the one at the top of the food chain, thus does not have the full influence over requirements, a relevance when budgets are involved and such concerns are likely to result in additional cost and possible extension of timeframe.

A description of managers in charge of something of which they know not provides evidence of the frustrations arising, as conveyed within the resignation post of the Chief Software Officer of the US Air Force (Chaillan, 2021). In his experience of chasing discontinuous funding, a coherent strategy for development of systems was undermined, at the same time as being drowned in bureaucracy. He bewailed the lack of agility in system development which kept systems current, relevant, and effective. While Chaillan [*op cit*] was speaking specifically of the Defence environment, the prevailing attitudes are recognisable to anyone from an IT projects background. The solution seems radical as it requires the cooperation and continuing involvement of a team of informed, expert commissioning parties.

The new wave cyber attacks are likely to be directed at data and communication paths, finding access through legacy systems, Cloud platforms, and corrupting AI algorithms to affect learning and alter goals. Data is essential for analytics and AI. A cautionary note could be applied comparing the model of AlphaGo Zero (Silver, et al., 2017) in which a system teaches itself by playing against itself, reinforcement learning, and the WOPR (War Games, 1983) in which an AI system ultimately concludes that all war scenarios are pointless, resulting in mutually assured destruction, and that the only way to win is not to start.

The National Security Commission on Artificial Intelligence [ibid] insists that whatever intelligent operational system is in use, there must always be a human in the loop. This does not necessarily hold outside the military complex, for example rights are given by the Article 22 of General Data Protection Regulation (European Parliament, 2016) to individuals to contest systems in which no human has participated. Presumably such cautions are to avoid a situation in which a learning system has corrupted goals and to counter irrationality of leadership that causes a declaration of war as in a film (Canadian Bacon, 1995), which plot bears remarkable similarities to some recent events.

2. The purpose of the AI system

It is perhaps more comprehensible to discuss purpose in a Defence context. The agility and flexibility it is thought to offer commanders in the field through the use of AI is misunderstood. For a system, any system, to be effective, it must have a defined purpose. For the military complex although on the battlefield it may seem obvious, to an AI developer, there is a need to define purpose and operational context. Without this definition, the AI developer is using his bias and knowledge to determine how to wage war and thus determining the behavioural actions and reactions of a programmed, intelligent device in the field.

The simple question, which perhaps no one wants to ask, is 'What is the purpose of this war?' Where this rises to the fore is in the use of autonomous devices. The rationale for the question is that a system will be created to wage war by evolving (AI) programmed boxes and the conduct of that war, which would normally be determined by the ethics and morality of the protagonists, is delegated and programmed into devices. What about encapsulation of United Nations ethics and morals, and alignment with international law?

Questions need answers. Is this war something that is being waged proactively as in the war on drugs, or something which is defensive, as against an attack? Does the behaviour differ in those circumstances? What are the rules of engagement? Do you choose the Rumsfeld, Bush, Powell, or any other doctrine? While it is possible to be more barbaric than your enemy and win a war, will that enable you to win the peace (Horne, 1978)? There is a concern that some nation states are in technological advance, but where they are not and it becomes a war of equals, does that lead to stagnation as in the First World War, or mutually assured destruction as in the Nuclear stand-off? In cyberspace, what are those equivalents and when will an incursion generate a retaliation?

There is much discussion of the cost savings and efficiencies surrounding repurposing AI systems. How do you repurpose AI, particularly after it has evolved in combat, learned and become something different from the original? Exactly what is it that you are repurposing? Furthermore, when repurposing, is this combat context the same as the previous context in which this system was used and evolved. Emphasis is placed on the human in the loop, but what happens when the human intervention runs counter to the behavioural choices of a device? Which one has ultimate control in an evolved system?

Diverting to civil authority purposes rather than the bellicose, surplus weaponry which has been purposed against combatants, such as the proposed taser carrying drones, can also be found in civil defence and police armouries enabling use, for example, in anti-terrorist/disaffected population control situations. Smith the inventor of the taser and body cam highlights:

> *Those groups are governed by different sets of rules and regulations;* (Tucker, 2021)

Being governed by a different set of rules indicates that the goals may be differrent and therefore the system must behave differently. As an example, while lethal use of drone carried weaponry may be appropriate in anti-terrorist situations when the prime concern of authorities is protection of citizenry, in western liberal deomcracies, the public relations catastrophe associated with any such lethal use on popular demonstrations/riots defines that admisitration as no longer belonging to the class of democracies, but to the

class of autocracies. If in doubt, consider the international reaction to the authoritarian responses to demonstations in Eastern European Democracies, Turkey, Hong Kong, Egypt etc.

In criminal hands, attaching any weaponry to drones is ripe for events such as drive by shootings, sentry duty, attacks on competitors and even attacking SWAT teams. It is much the same philosophy as use in military defensive capability, although the protagonists goals are different. OCG are exceedingly well resourced and funded through proceeds of crime. The profits from cybercrime are estimated to be at least equal to the Gross Domestic Product (GDP) of the Russian Federation.

> *[They] will continue to examine in an operational sense where cyber has utility, both in terms of the need to protect from it but also use it. It can be weaponised where appropriate against particular adversaries.* (British Army, 2021)

The above statement can apply to any OCG. They already have the capability to reverse engineer much of technology if they believe it can be used to their advantage and also may be major drivers of technological advance and a sink for well trained AI and cyber expertise.

3. Purpose and Cybersecurity

The answers to the above need to be programmed into devices, in advance and in the hope that your enemy will share your ethical standards. Part of the frustration expressed by Chaillan [*ibid*] was to do with recognition that in war, even a cold war, that may not be the case. It is assumed that one party to a conflict has a greater motivation to win, which leads that party to enact unethical behaviours - winning at all costs. It has been suggested that the inclusion of limiting factors in Stuxnet led to its attribution to western liberal democracy as no other states would have exhibited such an ethical approach.

Cybersecurity in its purest sense is about the security of systems. It is about protecting systems from externally induced corruption and ensuring their continuing operation in the manner specified. It is not about ethical and moral issues, nor the oversight of algorithm development. Systems are developed in accordance with requirements specifications. With intelligent systems, these are often goal directed and may contain various neural networks and be subject to machine or even deep learning capability. The way the systems adapt and grow is in response to a need to achieve their goal. There may be several behavioural steps on the way to achieving a goal, and identifying which one may be aberrant may be an impossibility for cybersecurity until malign behaviour is exhibited or a redefined goal is achieved.

It is easy to pick holes in the other people's work, but not particularly productive. A solution is required which is proactive raher than reactive. The initial risk assessment prioritises four areas: legacy systems, communications, data, and corruption of the learning mechanisms in AI systems.

4. Cybersecurity of legacy systems

Connected legacy systems are always problematic with respect to cybersecurity. They were written in the mists of time, like the 60 year old Inland Revenue Service (IRS) system holding taxation data of individuals. A time when it was believed the world was a more benign place. The USA IRS system will not complete its updating until 2030, when it will be replaced by a real time interactive system, under development since 2009 (Alms, 2021), thus already utilising obsolete systems and technology. It is difficult enough to find expertise and continuity of knowledge for maintaining the resilient operation of a legacy system, built in the times of early internet openness but this is compounded by trying to find a complete cybersecurity solution. For such systems, fortified cyber defence is possibly the only option, which provides unlimited access if breached. In an interconnected mesh, this is an undesirable characteristic.

The trial of linking legacy systems with interconnected systems is not just one of cyber security but also one of latency when data must be passed from one system to another. Without pattern of life information, under what circumstances can a change in latency be recognised as a cybersecurity failure. Too soon/early, it may be injected data, too late/slow, it may be an interface or communication failure. If it is too slow, what are the dependencies on the long awaited data which it is crucial to know for continuity purposes? If what is awaiting that data is an intelligent system, what happens?

With the passage of time, every system becomes a legacy system, requiring further patching and integration, A consideration, not necessarily directly related to pure cybersecurity, but nonetheless important, is the retention of legacy system skillsets. As a UK bank found to its cost, hiring back consultants you have just made redundant in a cost cutting exercise permits you to be held to ransom in an emergency. The moral of the UK situation is to only fire the experts after you have decommissioned the legacy system, its replacement is fully functional, and you are certain you have traced all the links to and from the legacy system and dealt with them. This latter is likely to be a growing problem in a mesh or interconnected system.

5. Cybersecurity of communications

There are many ways of disrupting communications. With trends towards utilisation of the Dew, Edge, and Fog, growing and increasing possibilities deriving from 5G and its successor 6G, evermore reliance is placed on a communications infrastructure. It is often seen to be advantageous to be at the forefront of technology - a pioneer, but as a tutor once said:

> *You can always tell the first adopters, they are the ones with the knife in their back. It is sometimes better to be a follower.* (Sumner, 1984)

This seems to run counter to military need for flexibility and agility and business need for competitiveness, but it really is better to let someone else make the first mistakes.

Persistent decentralised botnets using peer to peer communication and updating whenever they encounter a newer version which can be changes at a rate that will evade the usual indicators of compromise can be an issue. These botnets can be used within the system or as an attack platform with a trusted source IP and all its concomitant implications.

Reverting to a military context, the relevance of, for example, the US Army's digital transformation program resulting in a mesh, which it is intended will connect the tactical edge to the enterprise and places all the data in the Cloud, sounds like great cost saving and efficient use of technology and infrastructure (US Army, 2021). However, it is a cybersecurity nightmare. It implies that there is a requirement for continuous communication, between and connection to all endpoints at potentially high volume.

There may well be a military equivalent of an autonomous vehicle's communication which is labelled V2X or vehicle to everything which adequately describes the risks associated with proposed IoT and Edge computing. Rather than subscribing to the secure by design paradigm, what is hinted at in (US Army, 2021) is the post hoc Line of Effort #3 which addresses cybersecurity and is exacerbated by subsequent demand for agility and flexibility in upgrading or otherwise changing the system.

The fact that the system is so ill-defined that there is a need to communicate with 'everything' means it is impossible to define the endpoints. The paradigm of endpoint security and zero trust as espoused in Secure Beyond Breach (Reiber, 2020) provides insufficient protection in an open interconnected system where endpoints cannot be defined. The adoption of microsegmentation of networks requires a system boundary definition, identification of the virtual networks operating within that context and consideration of white lists. These very acts constrain the development of systems. Any component system must contribute, in some defined way, to the higher goals identified - its purpose. If something does not contribute to the goal, it has no place in the system. Whether it is necessary for a battlefield commander to have connection and data transfer between the battlefield and the enterprise should be a subject for further discussion.

The reliance on AI and Cloud technology places an essential burden on communications which must be available 24/7/365 with no down time where there is a dependence on real time data. In business, a breakdown in cybersecurity of technological communications becomes an operational continuity problem and there is a long list of documentation necessary to deal with such a situation, contained in the Blue Team Field Manual (White & Clark, 2017).

If security of communications is breached, there should be a continuity plan. However, this plan is not the sole responsibility of the Incident Response Team. The plan identifies the critical operational plans, and in the height of battle, there may be no requirement for enterprise information. Such a distinction implies that there may be a number of continuity plans based on specific scenarios. The proposed mesh may break, but a distinction

between a real time operational platform and a less time critical enterprise platform dictates the priority of response and perhaps provides some network design input.

In simple terms, if there is building security and the building is evacuated due to some event, would it be expected that the building security team would dictate operational continuity of the organisation and that each element of the operation was of equal importance? Stangely, the operational continuity plan seems to fall into cybersecurity responsibility, and be accepted by them as part of their responsibility for incident response. It contributes to the skills shortage, distances personnel from operations and curtails the cybersecurity aware culture. Prevailing attitudes may become more entrenched as dependence on systems and distance from them increases.

6. Cybersecurity of data: the Cloud

There are the usual attacks on data in transit available, to which it may be assumed that encryption is the answer. This is not always the case in time/mission critical systems, which may include AI, where resource at the operational/functional end is limited. The processing overhead is not always acceptable or feasible.

Army CIO has said the Army does not have

> *the right policies and procedures for some recent technologies, such as internet of things devices, or ways protect them on DOD's network* (Williams, 2021).

though the National Institue of Standards and Technology (NIST) is addressing part of the issue, pursuant to an Executive Order (Biden Jr, 2021). Exactly the same sentiment could be expressed by any organisation pursuing digital transformation. The question must be asked that with finite resources what is the priority, and the answer is actually something that was not mentioned, data. If the network was intact and entire, it would be useless unless information/data ere flowing through it.

Part of the reason for outsourcing to the Cloud is a cybersecurity skills shortage together, a belief that a dedicated, expert cybersecurity team is available to the Cloud Service Provider (CSP), and a potentially flexible infrastructure. The CSP has the knowledge of the infrastructure and responsibility for its cybersecurity. Unfortunately the Cloud is not inviolate from attack, and there may not be restrictions on its location or movement of data between one location and another. Where security is based on pattern of life, such movement may disrupt any latency definitions, raising alarm where it is not necessary. Cybersecurity may not be able to differentiate. There may be cybersecurity failures within the Cloud itself and unless the client arrogates investigative rights, it could well remain hidden or incompletely rectified.

Moves are afoot in the US, and presumably elsewhere, regarding regulation of access and audit clauses to be contained in contracts with CSP. Such considerations apply equally to any third party attached to the interconnected network and to any third party componentry, hardware, firmware or software as each represents an opportunity to a malefactor. There is the also the dangerous belief in the invincibility of one's own cybersecurity the initial belief that any failure must derive from elsewhere in the interconnected mesh.

In conclusion, recognition of issues concerning data, communications and the Cloud do not stop with cybersecurity. As Army CIO Iyer stated:

> *We know how to patch traditional IT systems. How do you patch sensors and operational technology, how do we collect the right analytics and data, and how do we share data with our joint service partners and industry. That is a policy problem* (Williams, 2021)

And everyone knows how long policy problems take to resolve.

7. Cybersecurity of learning mechanisms

As indicated in the introduction, corruption of learning systems may be the goal of any adversary or competitor. Subversion of a system would provide an advantage to the perpetrator. It is not certain how such an intelligent system could be fixed. How is the exact point of corruption found and the system unwound? In the case of a goal directed intelligent system unleashed onto the internet, could it be retrieved?

It is not clear how aberrant behaviour could be guaranteed to be detected, let alone proactively prevented unless there exists a boundary definition, a clear set of operational parameters and a unique pattern of life in execution. Evolving systems are not amenable to pattern of life mechanisms as the system is in a state of continuous 'growth' or evolution.

Again, as with the autonomous vehicle, the learning element could be carried out on 'land based' computers defining the limits of acceptable behaviours which contribute to the goal and are in accordance with permissible and possible behaviours and a device's capability. Once the scenarios have been defined, and system learning undertaken, the resultant predictable selection could be uploaded as an immutable system to the requisite device. Obviously this lacks the flexibility of military expectations, but permits more of a plug and play capability with each system learning from downloaded previous 'experience', adapting to those experiences to produce a further predictable system which is uploaded to the device in its entirety.

Uncontained intelligent systems can lead to unpredictable emergent properties and behaviours whereby the system as a whole behaves in a way that is not foreseeable from its collection of parts. A biological example of this is the collection of neurons and synapses which effectively form an electrical circuit, which is the basis of the brain. The brain has the emergent property of situational awareness. None of the individual component pieces possess this emergent property, but in its entirety, the brain does. The time of its emergence, functionality and any further evolution is not predictable. It is not possible to predict whether the emergent property will be malign or benign, useful or not.

There are potential legal implications for AI systems. These have not yet been tested. The Courts are demanding trustworthy AI, or at least a comprehensible explanation of why and how a system has behaved in a particular manner and whether that behaviour is reasonable. Some governments are demanding audits of algorithms to determine if they are equitable or have developed from an initial bias. This audit would take specialist training and it is not certain that it is within the cybersecurity remit although it must, certainly, be part of a risk assessment. The mechanism for the justification of Facebook algorithms which recognise anger, react to it, and reinforce or amplify it will be an interesting test case addressing as it does bias and misinformation, although these are tied together with surveillance capitalism (Zuboff, 2019) and analytics.

There is some thought to declaring AI systems to be legal entities in their own right (Abbott, 2020) and at this, the role of cybersecurity becomes even less certain.

8. Conclusion

This has been a diversionary discussion of evolving technology and issues concerning that evolution, clickbait in effect. No system is immune from attack and many organisations have the capability and time to invest in developing techniques to achieve a hack.

With the increase in interconnectivity and openness, there are limits to what cybersecurity can achieve. The adage 'a system's security is only as strong as its weakest link' prevails. What has emerged is a need to consider the true role of cybersecurity professionals, the realistic limits of their tasks and the systems they are trying to secure.

A recognition that the security associated with increasing use of AI systems may not be the preserve of current cybersecurity professionals forces decisions concerning definitions of purpose, ethics, morals, bias and behaviour to further professional classes (DeepMind, 2021). Currently it may not be possible to provide the full flexibility of AI systems that a military force might desire/need safely. With respect to unconstrained learning systems, even Google terminated the program for AlphaGo Zero before it had reached its full intelligent capability, so its full capability will remain unknown.

References

Abbott, R., 2020. The reasonable Robot: Artifical Intelligence and the Law. Cambridge: Cambridge University Press.

Alms, N., 2021. A 60 year-old IRS IT system won't finish modernising until 2030. [Online] Available at: https://fcw.com/articles/2021/10/20/irs-gao-cade2-delays.aspx?oly_enc_id= [Accessed 21 October 2021].

Biden Jr, J., 2021. Executive Order on Improving the Nation's Cybersecurity. [Online] Available at: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/ [Accessed 15 October 2021].

British Army, 2021. THEIA: The British Army's Digital Transformation. [Online] Available at: https://www.purestorage.com/content/dam/pdf/en/misc/british-army-digital-transformation.pdf [Accessed 15 October 2021].
Canadian Bacon. 1995. [Film] Directed by Michael Moore. USA: Metro Goldwyn Meyer.
Chaillan, N., 2021. It's time to say Goodbye!. [Online] [Accessed 12 October 2021].
Davies, G., 2021. Private Conversation [Interview] (4 June 2021).
DeepMind, 2021. Safety and Ethics. [Online] Available at: https://deepmind.com/safety-and-ethics [Accessed 15 October 2021].
European Parliament, 2016. Regulation (EU) 2016/679. [Online] Available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj [Accessed 15 October 2021].
Horne, A., 1978. A Savage War of Peace: Algeria 1954-1962. New York: Viking Press.
National Security Commission on Artificial Intelligence, 2021. Final Report, Washington DC: United States Government.
Reiber, J., 2020. Secure Beyond Breach. [Online] Available at: https://www.illumio.com/sites/default/files/2021-02/micro-segmentation-secure-beyond-breach-20eb10%20%281%29.pdf [Accessed 13 October 2021].
Silver, D. S. J. et al., 2017. Mastering the game of Go without human knowledge. [Online] Available at: https://www.nature.com/articles/nature24270.epdf?author_access_token=VJXbVjaSHxFoctQQ4p2k4tRgN0jAjWel9jnR3ZoTv0PVW4gB86EEpGqTRDtpIz-2rmo8-KG06gqVobU5NSCFeHILHcVFUeMsbvwS-lxjqQGg98faovwjxeTUgZAUMnRQ [Accessed 15 October 2021].
Sumner, F., 1984. Private converstion [Interview] 1984.
Tucker, P., 2021. The inventor of the Taser and the Body Cam Wants to put them on Drones. [Online] Available at: https://www.defenseone.com/technology/2021/10/inventor-taser-and-body-cam-wants-put-them-drones/186095/ [Accessed 15 October 2021].
US Army, 2021. The Army Unified Network Plan 2021 Enabling Multi Domain Operations. [Online] Available at: https://api.army.mil/e2/c/downloads/2021/10/07/d43180cc/army-unified-network-plan-2021.pdf [Accessed 15 October 2021].
War Games. 1983. [Film] Directed by John Badham. USA: United Artists; Sherwood Productions.
White, A. & Clark, B., 2017. Blue Team Field Manual. version 1.2 ed. Scotts Valley(California): CreateSpace Independent Publishing Platform.
Williams, L., 2021. ARMY CIO sets out to revamp IT policy. [Online] Available at: https://fcw.com/articles/2021/10/14/iyer-army-tech-policy-revamp.aspx?oly_enc_id= [Accessed 15 October 2021].
Zuboff, S., 2019. The Age of Surveillance Capitalism: The fight for a human future at the new frontier of power. 1st ed. London: Profile Books.