Mind games

Deception is particularly effective in the cyber domain but remains underexploited by defenders. **Neil Ashdown** examines the current state of defensive deception and ongoing research that would place a greater emphasis on deception for proactive cyber defence.

Key points

- The efficacy of deception for defence in the cyber domain is well-established, with modern commercial services focused on detecting adversaries and collecting intelligence on their activities.
- Automation has the potential to reduce the resource cost of creating and monitoring defensive deceptions, although it is very likely that the most sophisticated deceptions will remain resource-intensive, human-centric operations.
- The focus of defensive deception in the cyber domain is likely to shift towards deception that shapes an adversary's understanding of the situation, and thereby alters their behaviour, underlining the focus on deceiving the human adversary, rather than technological solutions.

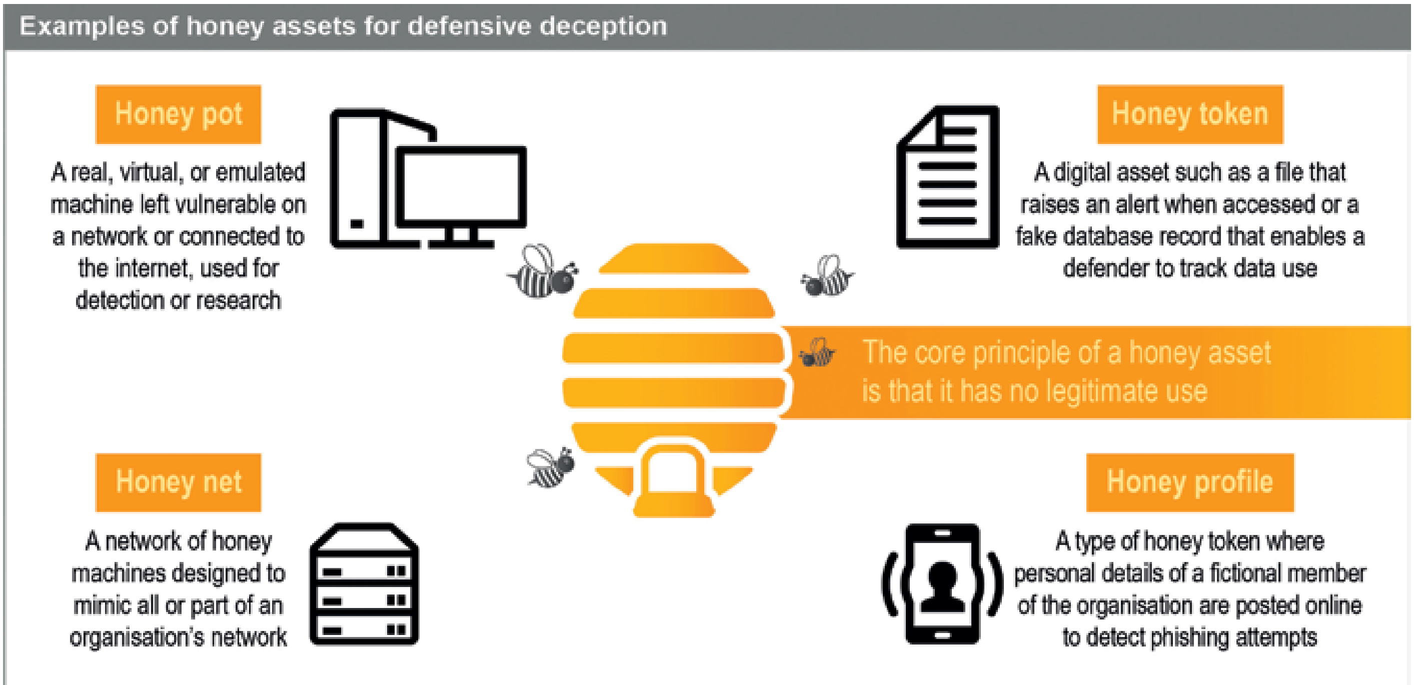
The United Kingdom established its first research unit focused solely on cyber deception in November 2019, reflecting growing awareness of the importance of deception in this domain. The National Cyber Deception Laboratory (NCDL) was established by Cranfield University and the Ministry of Defence's Cyber School. *Jane's* attended the NCDL's inaugural symposium.

The official press release for the event noted that "network defenders [could] take a proactive approach by using military deception tradecraft to effectively defend against and manipulate the activities of attackers operating within their networks". Speaking at the symposium, the head of the NCDL, Darren Lawrence, argued that the need for deception should be seen in the context of the nature of cyber conflict. Instead of comparing cyber conflict to warfare, with "discrete missile attacks between physically separated physical entities", Lawrence argued for viewing it as "a perpetual set of intimate knife fights between ... digital proxies".

Deception is a hallmark of military and intelligence operations. A common theme among various definitions articulated by practitioners is a focus on the impact on the target's behaviour. For example, Simon Henderson, from Artifice, an organisation which trains military and law enforcement personnel on how best to plan and conduct deception operations, defines 'deception' on the organisation's website as, "Deliberate measures to induce erroneous sense-making and subsequent behaviour within a target audience, to achieve and exploit an advantage."

Deception is particularly effective in the cyber domain, where the 'terrain' is a human construct. Speaking to *Jane's* on 24 April 2020, Frank Stech, an expert on deception at

Examples of honey assets for defensive deception



Honey pot

A real, virtual, or emulated machine left vulnerable on a network or connected to the internet, used for detection or research

Honey token

A digital asset such as a file that raises an alert when accessed or a fake database record that enables a defender to track data use

The core principle of a honey asset is that it has no legitimate use

Honey net

A network of honey machines designed to mimic all or part of an organisation's network

Honey profile

A type of honey token where personal details of a fictional member of the organisation are posted online to detect phishing attempts

Source: Neil Ashdown

© 2020 Jane's Group UK Limited/Getty Images: 1765291

the MITRE Corporation, argued that cyber deception offered a "significant asymmetric advantage" to the network defender, because they "own the terrain" and adversaries "lack [the defenders'] situational awareness".

This assessment was echoed by Robert Black, the Deputy Director of the NCDL and a lecturer on the Cyberspace Operations Masters programme at the UK Defence Academy. Speaking to *Jane's* on 27 April 2020, Black said that "the complexity of the virtual domain" could be exploited "to the defender's advantage ... The experience of the attacker can be fully manipulated by the defender, with the attacker not knowing what is real and what is false".

Black noted that in the past a deceiver needed to manipulate an aspect of the physical world in the vicinity of the target to affect the target's cognitive processes. He argued that "the virtual domain allows for cyber operations to be conducted remotely, and at scale, in a way that was not possible before, challenging traditional concepts of time and physical space". Black said that "increasingly, we find ourselves in a situation where the virtual world can effectively compete with the physical domain for dominance of an individual's cognitive processes – cyber provides a great opportunity for defenders to manipulate the mind and behaviour of the attacker".

Stanley Barr, also of MITRE, outlined to *Jane's* that MITRE was researching the use of deception for cyber security in three verticals: for detecting adversaries; for eliciting intelligence; and for adversary management.

The first examples of defensive deception for cyber security in the public domain were honeypots: computers connected to the internet and left deliberately vulnerable to attack. More advanced approaches involve creating networks of honeypot machines (honeynets) or deceptive documents and credentials (honeytokens).

The underlying principle behind 'honey' techniques is that the asset has no legitimate use. As such, an alert can be triggered any time a honey machine is accessed or a honeytoken file is opened, providing a mechanism for detection – the first vertical outlined by Barr – with a low false-positive rate. Once alerted, the defender can also observe the attacker's activity on their network, providing a source of intelligence – Barr's second vertical.

Modern commercially available defensive deception platforms generally deploy a combination of decoys and lures. Decoy assets in a defender's network complicate an adversary's efforts to map the network topology, as well as acting as detection tripwires. In a typical scenario an attacker might use a tool such as Mimikatz to search a compromised machine for credentials with higher security privileges. A commercial deception platform might automatically seed honeytoken credentials on the user's machines; the system would direct an attacker that used these credentials into a controlled environment for observation and issue an alert to the defender's security team.

The greater the presence of deceptive assets in the network, the more likely it is that an adversary will fall into the trap. However, this also translates into increased costs in terms of initial set-up and maintenance for network defenders. To tackle this resource cost, enterprise-level deception platforms already incorporate machine-learning technology to automatically populate networks with decoys and lures, as well as user dashboards or application programming interfaces (APIs) to aid the monitoring of the system or its integration into other security platforms.

The history of honeypots is a technical cat-and-mouse game between defenders and attackers. As early as 2003, a pseudonymous article released online in a fake edition of the online hacking publication Phrack Magazine detailed a number of technical tests for fingerprinting honeypots. With more interactive honeypots based on virtual machines, attackers can use techniques to identify virtualisation, similar to those used by modern malware to detect when it has been run in a sandboxed environment for analysis. Highlighting the cat-and-mouse game, modern defensive systems at times play on the attacker's preconceptions. One commercial defensive deception service claims to defeat malware by making the user's computers appear to be malware sandboxes; the malware 'detects' that it may be being analysed and remains dormant.

A 2015 report by technology consultancy Gartner on the state of commercial cyber-deception techniques and technologies described defensive deception as "still nascent", although potentially attractive "for larger organizations desiring advanced threat detection and defense solutions". Gartner assessed that "[d]eception as an automated responsive mechanism represents a sea change in the capabilities of the future of IT security". Stech told *Jane's* that "detection is the forte at this time" for modern, automated commercial defensive deception platforms.

Effectiveness of deception

A 2013 paper by researchers at MITRE, *Active cyber defense with denial and deception*, outlined the results of a wargame intended to test the effectiveness of "a dynamic network defense cyber-security platform" called Blackjack. This platform automatically generated deceptive content to feed to an attacker, based on the real information stored on a network. However, according to the paper, the Blackjack system was ineffective; it introduced notable delays and the attacking 'red team', having penetrated the network, could see the difference between the original data and the sanitised returns.

In contrast, the defending team had more success relying on manually implemented denial and deception tactics. Knowing that the red team had compromised its network, the blue denial and deception team was able to use its own network as a conduit to pass on deceptive information. Stech – one of the authors of the 2013 paper – told *Jane's* that the blue denial and deception team had adopted an "it's all good" approach upon being told that their network was compromised; "[I]f you know that they [the adversary] think something is false you can hide something real behind it", he said.

A series of studies published in 2017, 2019, and 2020 conducted by Kimberly J Ferguson-Walter, a researcher who has worked at the US National Security Agency (NSA), described the effect of deception on cyber operators. These findings suggest that the awareness of deception does not improve the ability to accurately identify deceptive content and may lead to poorer overall performance.

Filippo Monteforte/AFP via Getty Images: 1761854

Russian soldiers without insignia patrol in Perevalnoye, Crimea, on 20 March 2014. Deception is a key component of military planning and its importance is no less significant in the cyber realm.

An article published in the *Journal of Information Warfare* in 2017 by Ferguson-Walter, D S LaFon, and T B Shade described four pilot studies intended to test the effect of deception on cyber operators. The researchers reported that in one study, 19% of the assets on a network that were decoys accounted for 83% of the exploits launched by the penetration testers. For later studies the participants were briefed on the existence of deceptive assets on the system. The participants' performance did not improve; rather, they falsely identified decoy assets as real and vice versa, avoided vulnerable assets that they viewed as traps, and generally demonstrated greater uncertainty.

In the final study the participants were briefed on how the deception technologies worked. This enabled the participants to avoid detection, but "they accomplished this feat by not sending a single packet – arguably a win for defenders", as the researchers noted. Knowledge of the possibility of deception in the environment effectively led the participants to stop attacking the network. In her doctoral dissertation, published in 2020, Ferguson-Walter observed that "the combination of the presence of deception and true information that deception is present has the greatest effect on cyber attackers".

Black, of the NCDL, told *Jane's* that Ferguson-Walter's finding provided the cyber defender "the opportunity to move on to the front foot". He argued that if attackers came to expect the presence of deceptive assets in networks, "rather than neutralise [the deceptive assets' effect], it will potentially encourage a greater likelihood of defensive success, with the attackers being deceived into adopting the behaviours we would like".

Changing behaviour

A 2016 paper by three researchers from the (US) Air Force Research Laboratory's Information Directorate, Dave Climek, Anthony Macera, and Walt Tirenin, outlined desirable outcomes from cyber defences: "forcing [adversaries] to spend more time and resources, cope with greater levels of complexity and uncertainty, and accept greater risks of exposure and detection". These outcomes align with the goals of the defensive deception practices examined above. However, Climek et al voiced concern that "thus far [deception] has been minimally employed for tactics and strategies in cyberspace to counter cyber exploitation and attack".

There are parallels with the third of MITRE's three verticals: adversary management. Barr told *Jane's* that MITRE was examining ways to use deception to "change an adversary's strategic calculus. ... We want them to choose to stop using cyber". Similarly, Black said that at the NCDL, researchers were focused on novel ways to use deception to alter adversary behaviour, putting deception at the heart of a layered defence of core networks.

As an example of deception intended to shape adversary behaviour, Professor Neil C Rowe of the US Naval Postgraduate School proposed the creation of 'fake honeypots' in a 2007 paper. Rowe's idea was to modify computers so that they would be identified as honeypots by automated tests, a form of false flagging that could lead attackers either to target other machines or to lose confidence in the diagnostic quality of their own tests. Rowe noted that "unlike most security measures, this would work best against smarter attackers".

Moreover, Ferguson-Walter et al's findings suggest that once an attacker is aware of the possibility of fake honeypots in an environment, it will be extremely difficult for an individual operator 'with fingers on the keyboard' to be able to make judgements about the reality of any particular system with a high degree of confidence.

Black outlined a range of areas for the creative deployment of deception to manipulate an attacker's behaviour, such as false flagging the presence of more aggressive attackers already in the network to 'scare off' real attackers. He argued that "the opportunities presented by having a completely artificial environment in which the attacker must operate means that we can comprehensively shape their ability to understand and make sense of what is going on and, ultimately, shape what they do next in terms of their behaviour".

This challenge could lead to nation-state actors relying to a greater extent on blended human and cyber penetrations of target organisations. A human agent within the target organisation's security team could provide intelligence that would neutralise some of the organisation's defensive deceptions, in a way that could be critical if the cyber operator has been rendered ineffective by deception-in-depth. Stech told *Jane's* that in the MITRE wargame "we knew the Red adversary viewed this as a 'cyber' exercise" and that the blue team used this to their advantage, communicating through side channels and the "sneaker net" (physically walking over to talk to people), while conveying deceptive information through the penetrated network that was the red team's primary focus.

The high priority attached to scientific and technical data – for example, military and commercial research and development – by state intelligence agencies provides further challenges and opportunities for deception to shape actors' behaviour. Creating fake but credible technical data is likely to be challenging for defenders, particularly as any exfiltrated data will presumably be examined by the adversary's subject matter experts. This suggests that the defenders will need to work with their organisation's own subject matter experts to produce a credible deception.

However, this dynamic also brings advantages for a sophisticated defender. Black of the NCDL told *Jane's* that defenders could conceal deliberately inaccurate technical data on their systems. Black described this as a 'cognitive payload', drawing a comparison to the CIA counterintelligence operation in the 1980s that fed deliberately inaccurate technical data to the KGB. As Black emphasised, a key benefit of the 'cognitive payload' approach is that an attacker would be led into the position of questioning every piece of information that they discovered or exfiltrated.

The parallels with counterintelligence suggest a challenge for cyber-security professionals who seek to deploy deception at this level. In a 2010 account, a former head of CIA counterintelligence, Paul J Redmond, claimed that "[t]here exists at the human, professional, and management levels a mutual disaffinity between CI [counter-intelligence] officers and the 'computer people'". Although this account is now a decade out of date, successful deception operations will still require a diverse range of skillsets, including in computer security, psychology, the subject matter of the deceptive material, and intelligence.

Outlook

Early defensive deception technologies were mostly effective in detecting and monitoring the relatively unsophisticated self-propagating attacks that were common in the cyber domain in the 1990s and early 2000s. Modern attackers rely as much if not more on exploiting vulnerabilities in people as in software and hardware. As these attackers will almost certainly continue to use deception, if defenders do not do the same, this raises the prospect of an ongoing asymmetry in the competitive relationship.

Speaking to *Jane's*, Barr emphasised that deception was always "a supporting function, intended to accomplish a friendly action" – if that action is simply detection then there is little need for a costly and sophisticated virtual deception environment. Deceptions should be assessed based on how long they can survive scrutiny, and of what intensity, he argued. A deception that need only withstand a short, unreflective glance would require much less preparation than a deception intended to endure under prolonged, deliberative examination, Barr suggested.

Advances in automation may lead to wider use of defensive deception platforms for detection and collection, as the cost of deploying these platforms comes to align more closely with the benefits they provide in terms of detection and intelligence collection. Stech told *Jane's* that cyber deception and counter-deception were "going to be an arms race". He emphasised that awareness of this dynamic needed to shape research: "How do we innovate creatively in an arms-race situation?" Nonetheless, Stech cautioned that "cyber deception is never going to be a fire-and-forget tool ... it will always be a battle between sentient humans at some level". It is very unlikely that sophisticated, 'long con' deceptions will be capable of automation in the foreseeable future.

Moreover, measures aimed solely at detection and collection are unlikely to be sufficient for state organisations – such as government agencies, militaries, and intelligence and security services – that are among the targets of technologically advanced, well-resourced threat actors capable of operating over time and across domains. Black told *Jane's* that these organisations "cannot afford for some of their networks, data, and systems to be compromised" and that this meant that "more defensive measures will need to be taken, such as the use of cognitive payloads ... designed to exploit the attacker's decision-making and associated behaviour".

Conceiving of conflict in the cyber domain in this way would argue for a more proactive approach aimed at changing attackers' behaviour through the manipulation of their information environment. Doing this effectively would require a much broader operational remit and greater resources than many cyber-security teams are able to deploy. As Black told *Jane's*, gaining the full benefit of deception would require defenders to "radically change approaches to cyber security and, even more fundamentally, the configuration of digital networks". However, for nation-state actors, the threat to the security of their networks and data may be sufficient to motivate such a reorientation. ■

'[A] key benefit of the 'cognitive payload' approach is that an attacker would be led into ... questioning every piece of information that they discovered'

First published online: 07/05/2020

On the web janes.com/solutions

- France bolsters cyber capabilities and commitment through new doctrine
- UK to establish army cyber operational centres

Author

Neil Ashdown is a PhD researcher focused on cyber security and a former deputy editor of *Jane's Intelligence Review*.