NORTH ATLANTIC TREATY ORGANIZATION
ORGANISATION DU TRAITÉ DE L'ATLANTIQUE NORD

AI – Friend and Foe: The Arms Race in Cyberspace

Dr. Manfred Boudreaux-Dehmer
NATO Chief Information Officer

VERSION 8 March 2024

Agenda

1. NATO
2. CIO and Cyber
3. Major Threats
4. Motivation and Sources of Threats
5. Ways to defend against AI
6. AI Threat Management
7. Adversarial AI
8. AI and Deepfakes
9. Multi-Domain Operations

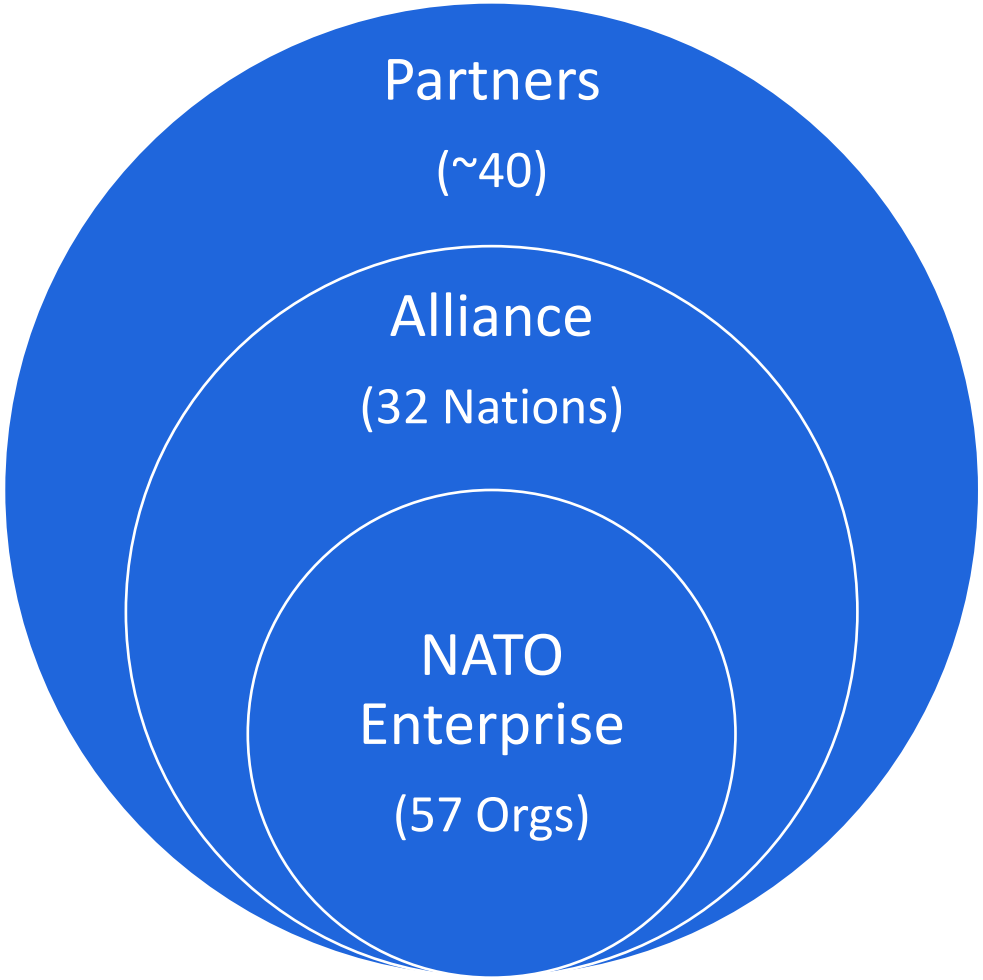
North Atlantic Treaty Organization

Washington Treaty

~~31~~ 32 Members

Guarantee freedom and security of its members through political and military means

NATO Ecosystem



Office of the CIO

Reporting to Secretary General and North Atlantic Council

Dual Mandate

IT Coherence and Cybersecurity

Cyber

Embedded in NATO's core tasks

Threats are increasing in frequency and sophistication

Cyber is a military domain

Focus on

- Protecting our networks
- Conducting operations
- Helping Allies enhance national resilience
- Providing a platform for consultation and collective action

Major Threats

Russia

Israel / Hamas

China

North Korea

Russia

800% increase of attacks immediately after invasion

Massive disinformation

Israel / Hamas

Consistent operations tied to Iran and Hezbollah

China

Biggest global threat

40 "Advanced Persistent Threat" groups

North Korea

Wide-ranging financial activities (>$1B per year)


National Cyber Security Centre
Home
Information for...
Advice & guidance
NEWS
UK and allies ex
tactics of Russ
New advisory reveals evolving tac
as more organisations move to cl
CYBERSECURITY & INFRASTRUCTURE SECURITY AGENCY
AMERICA'S CYBER DEFENSE AGENC
Topics
Spotlight
Resources & Tools
News & Events
Home / Topics / Cyber Threats and Advisories / Nation-State
CSO
North Korea's state hacking program is varied, fluid, and nimble
by Cynthia Brumfield
North Korea's evolving and flexible hacking structure encompasses a wide range of malicious activity beyond stealing from cryptocurrency exchanges, fueled by a small but nimble cyber workforce.
REPORT A CYBER ISSUE
SHARE:
Related content
Opinions
Events
breadth, depth, and sophistication of the People's Republic of China (PRC) state-sponsored
Directorate. The report, which summarizes the main cyber attacks
of the last two and a half months, notes that most of these groups share
intelligence, methods, and tools with each other.

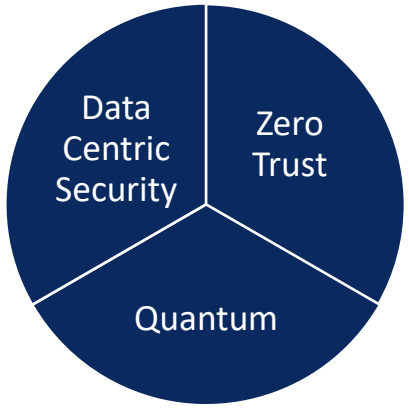
Sources of Threats

- Malware
- DDoS
- SQL-Injection/ Cross-Site-Scripting
- Insiders
- Password Brute-Force
- DNS-Spoofing
- Man-in-the-Middle
- Social Engineering
- URL-Poisoning
- Ransomware
- Botnets

Motivation of Threats

- Extortion
- Hacktivism/ Reputation
- Identity Theft
- Crypto Jacking
- Bullying
- Espionage

Ways to defend against AI



Data Centric Security

Fine-grained access at object level

Labelling of all data

Central Identity Management

Zero Trust

Assume a breach

Least-privilege access

Persistent re-authentication

Quantum-Resistant Cryptography

Q-Day: Computers will be able to break existing algorithms

Danger of "harvest now and decrypt later"

AI Threat Management

Goal: Act faster and in advance

Data Loss Prevention (DLP)

Value-add for Security Operations Center

Deal with increased complexity of threats

- Hybrid cloud / on-premise
- More network connections
- More data exchange with partners
- More teleworking

"Super charged" pattern analysis

Tracking abnormal content movement – predict!

Positive effect on incident containment / response

SOC deals with true exceptions

Key: Consistently train the model to reduce false positives

Adversarial AI

Attack automation

Adversary emulation

Generative-competitive neural networks

Attack automation

Reconnaissance, target exploitation, E2E network penetration

Increase in volume and pace (existing vectors)

New set of attack methodologies

Adversary emulation

AI-powered red team inside your network

Identifies / executes attack paths post compromise

Geared to increase knowledge-gaining and foothold

Emergent system behavior

Generative-competitive neural networks

Experimental: 2+ neural networks exploit each other – each creates training datasets in response to the other

AI and Deepfakes

Ryan Gosling outside the Westin Bayshore, looking towards the conference center, with North Vancouver in the background (generated with RunwayAI)

Personalization: Contextualization and believability (deep insights into behaviors and habits)

Increasing quality in audio, imaging, video

Social Engineering / P(V)hishing

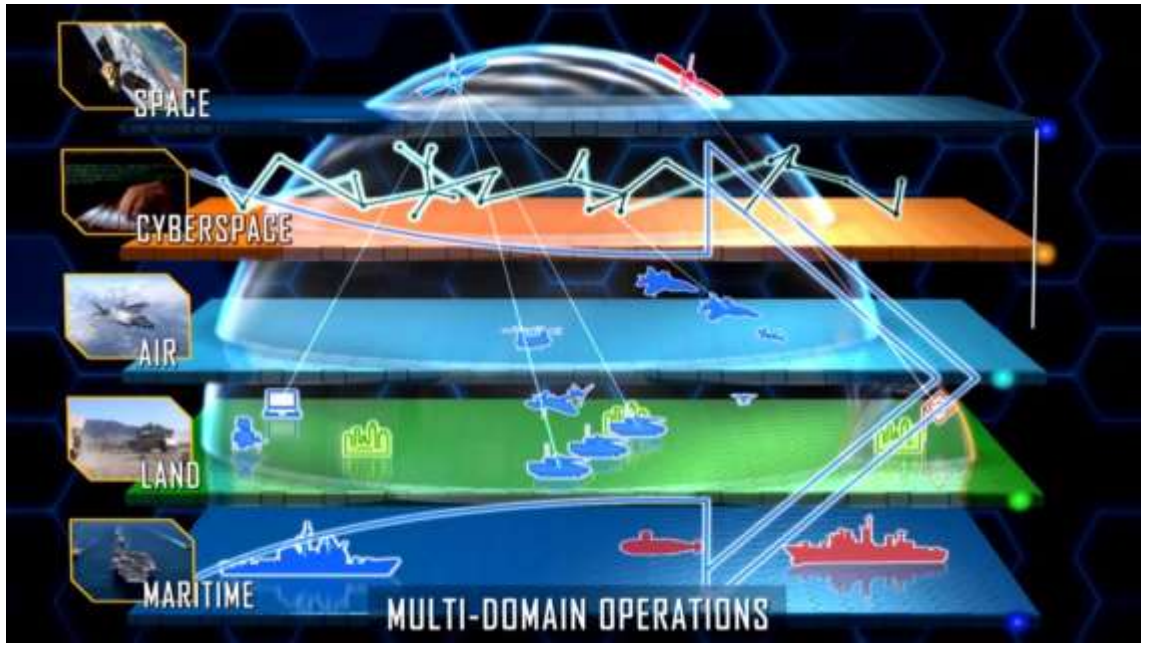
Combination of GPT-4 and AIaaS removes entry barrier to spear-and whale-phishing

Disinformation

Information battlefield: Deliberate falsehoods replace facts with fiction

Geared to weaken the Alliance – deepen divisions within and between NATO members

Response: Analyze (AI) and engage to 'right the picture'



Multi-Domain Operations

Orchestrate activities across all military domains and synchronize with non-military activities

Integrate AI, robotics, autonomous vehicles, drones – connect sensors with shooters

Strategic interplay of…

- Satellite intelligence (Space)
- Cyberattacks and defensive actions
- Kinetic domains of Air, Land, Sea

Multi-Domain Operations

Orchestrate activities across all military domains and synchronize with non-military activities

Integrate AI, robotics, autonomous vehicles, drones – connect sensors with shooters

Scenario

1

- Blue operators infiltrate red network
- Battle plans 'disappear'
- Supply orders make no sense
- Railroad signals are disrupted = traffic jam
- Blue fires Electro Magnetic Pulse missile
- Red observation drones crash
- Red forces see on social media: Leader stepped down
- Red team surrenders

Credit: Paul Nakasone and Charlie Lewis

Multi-Domain Operations

Orchestrate activities across all military domains and synchronize with non-military activities

Integrate AI, robotics, autonomous vehicles, drones – connect sensors with shooters

Scenario

2

- Red troops fight in blue territory border town
- Blue launches missile attack on decoys (spread confusion)
- Blue previously burrowed into red system
- Red side receives 'command' to surrender
- Blue deceives red with fake intelligence on troop size
- Red receives 'command' to provide grid location of every soldier
- Blue dispatches drones to every soldier, scans irises, relays positions
- Blue jams red's communication abilities
- 40 red soldiers surrender because of 4 blue cyber operators

Credit: Paul Nakasone and Charlie Lewis

Q & A

Thank You –

Dr. Manfred Boudreaux-Dehmer
NATO Chief Information Officer

LinkedIn: https://www.linkedin.com/in/manfred-boudreaux-dehmer/