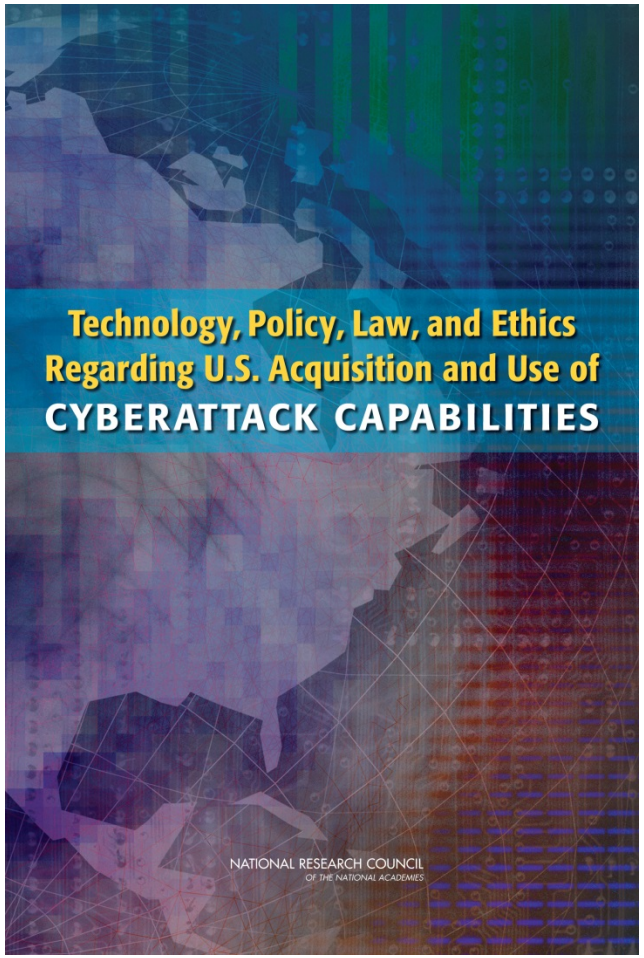
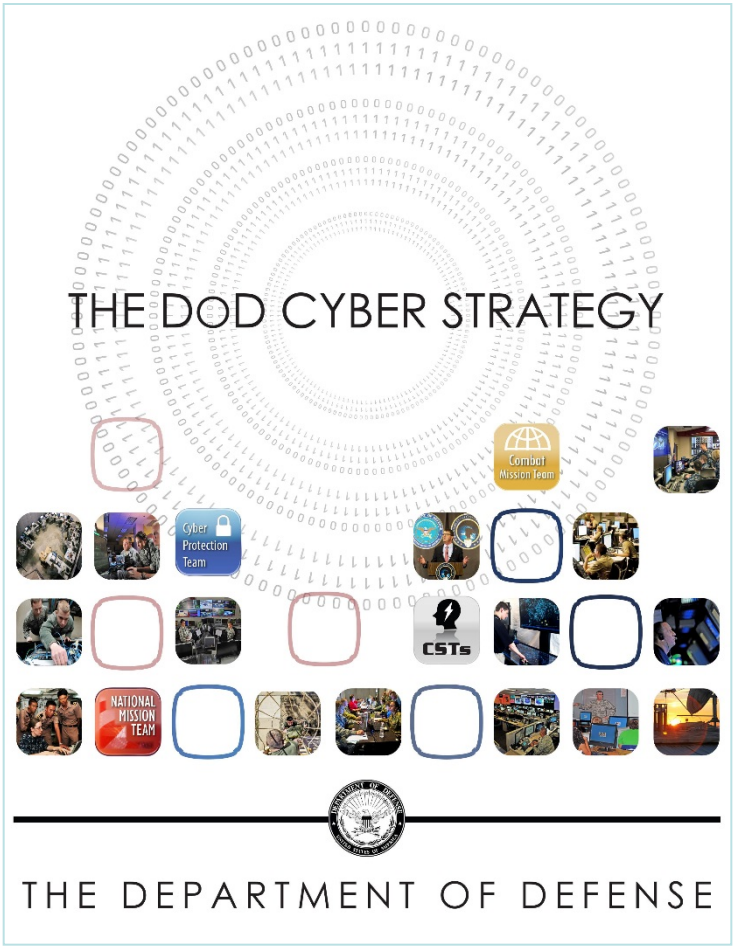
Fundamentals of Cyber Conflict

Herb Lin
Stanford University
CS-203
May 23, 2017

Some source material





Policy and Technology Framing

A fundamental distinction

- Cybersecurity vs cyber security
  - Cybersecurity: security of cyber things (computer and communications technology systems) as “proper system operation even under conditions of threat”
  - Cyber security: security of cyber domain as in “national security”: the ability to preserve the nation's integrity and territory; to maintain its economic relations with the rest of the world on reasonable terms; to preserve its nature, institution, and governance from disruption from outside; and to control its borders“ (Harold Brown, fmr SecDef)

A basic premise

- Cyber conflict and cyber security have both defensive and offensive dimensions, and comprehensive approaches require understanding both.
  - Defensive cybersecurity (very public)
    - Passive defenses
      - Anti-virus and intrusion detection software
      - Better password security
      - Greater attack resistance in software
    - More robust law enforcement mechanisms
      - e.g., Convention on Cybercrime
  - Offensive cybersecurity (usually classified)
    - Offensive operations can be used for defensive purposes.
    - Offensive operations can be used for non-defensive purposes.

Governments need policy to guide use of offensive capabilities

- If police officers carry guns, policy must address:
  - Doctrine: general guidance about the circumstances in which the use of lethal force might be necessary
  - Training: how to use guns
  - Standing rules of engagement (SROE): in detail, under what circumstances to use guns
  - Command and control: exceptions to SROE re use of guns
  - Identification friend-or-foe (IFF): how to distinguish between bad guy and police/innocent bystanders
  - Liability and insurance: responsibility for mistakes
- The intent of allowing police to carry guns is defensive.
- Bad guys rarely need to worry about these issues.

What is a “gun” in cyberspace?

- Cyber weapon: instrument used to create bad effects against adversary computer
  - Destroy data/program
  - Render it inaccessible
  - Steal data
  - Harm devices attached to computers

Basic technology of cyber weapons

- Cyber weapon: instrument used for hostile or unfriendly purposes against adversary computer

- Aspects of a cyber weapon
  - Access
  - Vulnerability
  - Payload

  } Penetration

- Cyber “weapon” is not a particularly good term, but no better term available.

Access

- Technical
  - Remote (through the Internet)
  - Close-access (e.g., through chip swap, USB key, supply chain, tapped cable, clandestine WiFi, burglary, shipping)
- Social
  - Trickery, bribery, blackmail, extortion, persuasion
  - Some targets of social access
    - Users and operators
    - Vendors and service providers
- Technical and social elements often combined, e.g., phishing.

Vulnerabilities

- Software (application or system software with accidentally or deliberately introduced flaws)
- Hardware (microprocessors, graphics boards, power supplies, peripherals, storage devices, network cards).
- Communications channels (e.g., tap on fiber optic line)
- Configuration (e.g., ports improperly left open, weak passwords allowed)

Payload – determines type of offensive action

- Attack: degrade, disrupt, destroy, deny system/network or information therein
  - Integrity (data/operations are altered—includes botnet, self-destruction of computer, change data)
  - Authenticity (data/operations are forged)
  - Availability (data/operations is inaccessible)
- Exploitation (surreptitiously exfiltration of confidential information)

Note that attack and exploitation use the same access paths to take advantage of the same vulnerabilities – only payloads are different. How does victim distinguish between the two?

Cyber operations vs cyber weapons

- Offensive cyber operation requires cyber weapons, people, planning, goals, command and control, rules of engagement.
- Weapons provide offensive capabilities (enable offensive action), which can be used for offensive or defensive purposes.
  - Offensive capabilities are hostile (destroying, damaging, degrading, disruptive, denying) and act on a technological artifact
  - Defensive capabilities prevent or mitigate destruction, damage, degradation, disruption, denial
  - Offensive purpose—disadvantages adversary, advantages self
  - Defensive purpose –protects interests of self; maintains status quo.

Some cyber examples

	Offensive purpose	Defensive purpose
Offensive capability (destroy, degrade, disrupt, manipulate, deny)	• Cyber attack to degrade infrastructure for adversary nuclear capability (R&D) (Stuxnet)	• Cyber attack to take out botnet controllers (sometimes with court approval)
Defensive capability (prevent use of offensive capability)	• Encryption used to hijack data for ransom	• Encryption used to ensure confidentiality of private data

Some important characteristics of offensive cyber capabilities

- Cyber is offense-dominant in most situations
  - The only perfectly secure computer is useless.
  - The enormous complexity of modern information technology means that there are multiple places where an adversary can intervene, and often only one intervention is necessary.
  - Given enough time, the offense will always be successful.
  - When timetable is determined by external events and coordination is necessary, success is harder to achieve.

- Offensive capabilities span a very large range
  - Very destructive to nondestructive
  - Very selective to not selective
  - Immediate execution to long-delayed execution

- Offensive operations can be conducted with plausible deniability in the short term, but
  - Attribution may be possible in the long term, drawing on all sources of intelligence. Usually no smoking guns.
  - Kinetic forces also can operate with deniability.

- A given offensive cyber operation may be:
  - Known only long after penetration has occurred
  - Limited in utility: usable only once or a few times
  - Fragile: usable only as long as intelligence is valid
  - Technically fast but operationally slow; hence most suitable in nonurgent scenarios (e.g., early use); “speed of light” vs “speed of law”

- Planning is both critical and hard
  - Large range of options than most traditional military operations
  - Many possible outcome paths➔ very specialized knowledge
  - Cascading effects, collateral damage estimates, damage assessment all hard to perform

- Intelligence support for cyber operations is critical
  - Must be accurate, detailed, timely (esp. if targeted)
  - Intelligence gathering requires much advance planning (can be years!)
  - Intelligence is fragile (e.g., depends on updates being installed)
  - Adversary can take steps to invalidate intelligence if aware of need.
  - Needs for intelligence creates pressures for early use

Two different kinds of offensive cyber capability

- Cyberexploitation (surreptitiously obtain confidential information – aka espionage)

- Cyberattack (degrade, disrupt, destroy, deny system/network or information therein)
  - Integrity (data/operations are altered—includes botnet, self-destruction of computer, change data)
  - Authenticity (data/operations are forged)
  - Availability (data/operations is inaccessible)

- Hostile actions involve penetration and payload.
  - *Penetration* enables hostile action
  - *Payload* specifies what hostile action to take

Some ways to use offensive cyber capabilities

Offensive actions for defensive purposes

- Before adversary attack
  - Pre-empt offensive cyber action about to be undertaken by adversary
  - Provide early warning of adversary cyber attack (must penetrate adversary networks **before** tactical threat emerges)
- During adversary attack
  - Disrupting a cyberattack in progress by disabling the attacking computers
- After adversary attack
  - Need for conducting forensic investigation (exploitation)
  - Retaliatory attacks to discourage further attacks .

Offensive actions for offensive purposes

- Traditional military operations
  - Suppression of adversary weapons (e.g., air defenses).
  - Interference with adversary command and control
  - Disruption of logistics chains
- Nontraditional operations (aka “covert action”)
  - Influencing the outcome of a foreign election (hacking voter registration, destroying pension records, conducting information warfare)
  - Destabilizing a nation through attacks on the financial system.
  - Damaging an adversary’s nuclear weapons production facilities
  - Damaging personal reputation of adversaries, manipulating perceptions

- Attacks on critical infrastructure
  - Some elements
    - Power grid
    - Transportation (e.g., air traffic control)
    - Financial infrastructure
    - Communications
  - Most concerns expressed over large-scale damaging attacks with long-lasting effects
    - EMP burst could have such effects
    - Hard to imagine other kinds of cyberattacks with such effects
  - Attacks of concern *should* include small-scale attacks with large impacts on public confidence.

- Foreign intelligence and espionage
  - National security intelligence gathering/espionage
    - Diplomatic information
    - Negotiation positions
    - Political plans
    - Military information
    - Personnel information
  - Economic espionage
    - product development and use
    - manufacturing procedures
    - business plans,
    - policy positions and analysis
    - emails of high-ranking employees;
    - A slow bleed of large significance because of scale, but many numbers estimating economic value of loss are suspect

- Domestic intelligence
  - …. Coals to Newcastle…

- Key issues
  - Scope, scale, nature of desired intelligence gathering
  - Appropriate distinctions, if any, between foreign and domestic collection
  - Appropriate legal authorization (and enforcement)
  - Scope, scale, nature of infrastructure needed (nonconsensual surveillance requires more than consensual)

Punching or punching back?

- Punching back unlikely to work very well
  - Pre-emption require ubiquitous presence and automated decision making
  - Disruption requires instantaneous identification of attacking computers and won't harm actual perpetrators. (Also raises questions of attacks on computers belonging to US citizens.)
  - Retaliation doesn't have to be in cyberspace or against the attacking computers
  - Retrieving or destroying stolen information likely to be futile as copies will be made.

- But punching (not punching back) doesn't require tight timelines.
- Ideally suited for first use
  - May be less provocative than kinetic action
  - Plausibly deniable, at least in short run
  - Harm may be decoupled from mechanism of action and hence harder to find problem and to take countermeasures
- Ideally suited for covert action intended to weaken adversaries

US views on cyber

Excerpt from DOD Cyber Strategy: US Strategic Goals

- Build and maintain ready forces and capabilities to conduct cyberspace operations;
- Defend the DoD information network, secure DoD data, and mitigate risks to DoD missions;
- Be prepared to defend the U.S. homeland and U.S. vital interests from disruptive or destructive cyberattacks of significant consequence;
- **Build and maintain viable cyber options and plan to use those options to control conflict escalation and to shape the conflict environment at all stages;**
- Build and maintain robust international alliances and partnerships to deter shared threats and increase international security and stability.

- On offensive operations in the DOD Cyber Strategy
  - OCOs will be conducted in accordance with the laws of war
  - Targets of OCOs include adversary command and control networks, military-related critical infrastructure, and weapons capabilities.
  - OCOs may be conducted during periods of heightened tension (i.e., before the outbreak of outright hostilities).
  - Offensive capabilities with significant effects exercised on NCA determination for disruption of an adversary's military related networks or infrastructure so that the U.S. military can protect U.S. interests in an area of operations. Examples:
    - Conflict termination on U.S. terms
    - Disrupt adversary military systems to prevent the use of force against U.S. interests.
    - Deter or defeat strategic threats in other domains working with other USG agencies.
  - DOD concerned only with responding to attacks of highest consequence.

- PPD-20 (still classified – may not survive Trump admin)
  - Directs relevant agencies to start assembling a list of "potential targets of national importance" for OCO
  - Requires "specific presidential approval" for offensive operations with "significant consequences"
    - loss of life
    - significant responsive actions against the United States
    - significant damage to property
    - serious adverse U.S. foreign policy consequences
    - serious economic impact on the United States.
- DOD acknowledges use of cyber weapons against ISIS

The IC view of offensive operations

- Signals intelligence is “intelligence derived from electronic signals and systems used by foreign targets, such as communications systems, radars, and weapons systems.” (www.nsa.gov)
- U.S. government agencies collect and analyze SIGINT to assist U.S. policy and decision makers.
- SIGINT is governed by US law, Presidential executive order, and regulation to protect the rights of U.S. citizens (persons). Non-US persons have no (very few) rights under U.S. policy.
- EO 12333 tasks CIA with covert action, which may include offensive cyber operations.

International humanitarian law (the laws of armed conflict)

What is cyber war?

- Why is a definition of “cyber war” necessary?
  - To know when the Article 51 threshold of “armed attack” has been crossed?
  - To know when 2(4) prohibition on “use of force” has been violated?
  - To know when IHL must be invoked?

- Two cases of interest:
  - Offensive cyber operations being conducted as part of a traditional armed conflict using kinetic weapons → intended to cause kinetic-like effects.
  - Offensive cyber operations being conducted without the contemporaneous use of as part of a traditional armed conflict using kinetic weapons → intended to cause sub-threshold effects

What is not cyber war?

- A teenager defacing a DOD/MOD web site.
- A person hacking into the bank accounts of a defense contractor to steal money.
- An unfriendly nation stealing plans for a new jet fighter.
- A terrorist group using the Internet for recruiting, fund raising, propaganda, and communications.

Dividing lines between criminal acts and acts that might implicate the UN charter or IHL are unclear.

Many examples of cyberattack; few (if any) examples of cyber war.

Responses to hostile subthreshold actions are the most immediately relevant dimension of policy today.

- Worldwide Threat Assessment of the US Intelligence Community, 2015
  - "[T]he likelihood of a catastrophic [cyber] attack from any particular actor is remote at this time. Rather than a "Cyber Armageddon" scenario that debilitates the entire US infrastructure, we envision something different. We foresee an ongoing series of low-to-moderate level cyber attacks from a variety of sources over time, which will impose cumulative costs on US economic competitiveness and national security."
  - Cyber is the first-mentioned item on the list of threats faced by the United States (and hence understood to be the most significant).

Key terms in UN Charter (bolded below) not defined

- UN Charter prohibits "**threat or use of force** against the territorial integrity or political independence of any state" (Art. 2(4))
  - "Force" not defined. By practice, it
    - includes conventional weapon attacks that damage persons or property
    - excludes economic or political acts (e.g. sanctions) that damage persons or property

- UN Charter Art. 51 - "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an **armed attack** occurs against a Member of the United Nations.."
  - "Armed attack" not defined, even for kinetic force.
    - Relatively easy: If cyberattack causes effects comparable to that of a kinetic attack, it should be treated the same way (at least an effects-based analysis).
      - Damage to air traffic control, dams, nuclear reactors that cause significant death/destruction are armed attacks.

    - Relatively hard: If cyberattack causes other effects (e.g., long-term disruption to critical infrastructure or stock exchanges) without immediate large-scale death or destruction of property), legal status is unclear.

Jus Ad Bellum - UN Charter

- UN Charter prohibits "**threat or use of force** against the territorial integrity or political independence of any state" (Art. 2(4))

- UN Charter Art. 51 - "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an **armed attack** occurs against a Member of the United Nations.."

- UN Charter written in 1945, long before cyber. **Bolded terms not defined.**

Hard scenario 1: Economic damage without physical damage

Banking and financial systems are heavily dependent on computer and communications technology.

- Zendia launches an attack against the banking systems of Elbonia in which national accounts were manipulated in order to cause significant financial instability in Elbonia, resulting in panic and a loss of confidence in the Elbonian population.

Hard scenario 2: Interfering with elections

Civilian IT systems are generally not well-defended.

- A Zendian cyberattack is used to hack the electronic voting machines used in a close Elbonian election, thus tilting the election to the party favored by Zendia.

- A Zendian cyberattack corrupts the pension records of millions of people in Elbonia. In the next election, the ruling party in Elbonia (disfavored by Zendia) is voted out of office because of the resulting scandal.

Hard scenario 3: Ambiguities between exploitation and attack

International law does not prohibit intelligence collection by spies, and cyber exploitation is not illegal under international law.

- Zendia conducts repeated and continuing probes of Elbonian military and defense industry networks, exfiltrating classified and unclassified data on a large scale. Is this a threat of force?

- Zendia introduces Trojan horse agents into Elbonian military and infrastructure computer systems that exfiltrate data and also have a capability of being upgraded remotely. Is this a threat of force?

- Zendia introduces Trojan horse agents into Elbonian military and infrastrcture computer systems that only have a capability of being upgraded remotely. Is this a threat of force?

Jus in Bello - Some Important Principles

- Principle of Proportionality
  - Collateral damage on civilian targets acceptable if not disproportionate to the military advantage gained.
- Principle of Distinction
  - Military operations only against “military objectives” and not against civilian targets
  - Only military personnel can directly participate in hostilities

On proportionality

- Definition of “inadvertent harm to civilians” in cyberattack?
  - NO - Mere inconvenience.
  - YES - Death on a large scale.
  - MAYBE –
    - the inability to conduct financial transactions electronically,
    - periodic interruptions in electrical power,
    - major disruptions in travel and transportation schedules,
    - outages in communications capability.
  - Example case: A botnet uses compromised computers to conduct attacks. If a compromised computer is 99.99% functional for the user, is it collateral damage?

On distinction

- Targeting only “military objectives”
  - Knowing that a given computer is a valid military target may be difficult and highly uncertain. How should a cyber attacker differentiate military and civilian computers? (Does defender have responsibility to help differentiate?)
- Targeting dual-use infrastructure
  - Communications facilities
  - Electric grid
  - Financial network
  - High degree of entanglement of civilian and military networks for improved efficiency and reduced costs will only grow in future.
- Identification of military forces (e.g., uniforms on soldiers, insignias on airplanes). What “markings” must Internet-carried cyber weapons have (analogous to insignias on missiles)? How would this affect effectiveness?

On Attribution

- The conventional wisdom
  - Hostile cyber operations cannot be attributed to perpetrators with high confidence.
  - Thus,
    - Hard to mitigate ongoing harm
    - Impossible to punish
    - No disincentives to act in a hostile manner

- But conventional wisdom is wrong (or at least, it's incomplete).

A canonical example

- Distributed denial of service attack by A against Z.
  - A assumes control of many computers B, C, D, ….
  - A orders computers B, C, D,… to flood Z with phony requests for service.
  - Legitimate users of Z cannot use Z.
- If A cannot be identified, A will not be punished and Z will continue to suffer.
- Perhaps computers B, C, D … can be identified and shut off, thus mitigating the attack on Z. But A is untouched.

An illustrative scenario

- A U.S. computer is attacked in cyberspace.
- The attack traffic arrived from a computer based in Kansas owned by a 78 year old grandmother.
- The computer in Kansas was compromised using a computer in Greece.
- George sat at the keyboard in Greece.
- George is a citizen of Germany.
- George is a member of a Russian organized crime group.
- The leader of the crime group is a close personal friend of a senior leader in the FSB.
- Who is “responsible” for the attack on the U.S. computer?
  - **Only the steps in red can be addressed technically**
  - **Notice the political and policy dimensions of assignments of responsibility**.

Unpacking attribution

- Three general meanings for “attribution”
  - **M**achine or machines (technical/computer science determination)
  - **H**uman operator who initiates a hostile action (human determination)
  - **P**arty ultimately responsible for actions of human operator (political determination)

  NB – **M** does not necessarily give **H**, **H** does not necessarily give **P**
- In addition, **P** can be determined by
  - The geographical location of the machine that launched or initiated the operation (George is sitting at a keyboard located in Greece)
  - The nation under whose jurisdiction the named individual falls (George is a citizen of Germany).
  - The entity under whose auspices the individual acted (George works for an organized crime cartel with ties to the FSB).
- Appropriate meaning depends on the goal of attribution
  - Mitigate the pain as soon as possible: **M**
  - Prosecute/take actor into custody: **H**
  - Deter future acts (a primary goal for national security): **H or P**

For national security purposes, we want to attribute to P (a state)

- State-prohibited without capability to enforce prohibition against third party actions (TPA)
- State-tolerated. State does nothing to stop TPA.
- State-encouraged. State encourages or provides support (intelligence support, operational support)
- State-directed. State orders TPA.
- State-conducted. State uses military/intelligence assets to conduct offensive cyber operations, perhaps integrated with TPA

Information sources for attribution

- All-source intelligence is not just technical intelligence
- Technical forensic information (gives information about one attack)
- Technical mistakes in tradecraft (e.g., use of dating profile)
- History – use of weapons or techniques used before
- Operational security failures--discuss plans or activities on insecure communications media, receive help from careless sources
- Geopolitical context and demands

What is hard is PROMPT attribution, since it takes time to assemble and analyze clues.

Different levels of attribution certainty needed for different goals

- In criminal prosecutions
  - "beyond a reasonable doubt"
  - "clear and compelling"
  - "preponderance of the evidence"
  - Must convince an impartial jury/judge

- In national security decision making
  - Standards for taking action are much less formal.
  - "due process" and "rights of accused" have analogs that are weak at best.
  - Must convince ourselves and opinion in other nations
  - Assigning political responsibility is a political act, not a technological problem.

On deterrence

Why deterrence?

- Passive defense is inadequate and eventually will fail; law enforcement actions are too slow and uncertain in outcome.
- Attacker can choose time/place of attack, and need only succeed once, while defender must succeed everywhere all of the time.
- Thus, must persuade would-be attacker to refrain from attacking.
- Deterrence seems like the inevitable choice in an offense-dominant world.

On cyber deterrence

- **What hostile cyber action is to be deterred?**
  - Low end/high end?
- **Who is being deterred?**
  - Threatened responses must be tailored.
- **What response should be threatened to deter that action?**
  - A wide range of actions possible
- **What is needed to make the response credible?**
  - Capability
  - Attribution

What hostile cyber action is to be deterred?

- Hostile cyber actions can span a very wide range.
- Low-end actions are frequent, not particularly harmful, and may not need to be deterred.
- High-end actions are infrequent, very harmful, and need to be deterred.

Not everything bad can or should be deterred. What is the dividing line between these two kinds of action?

Examples of undesirable cyberactivity

- A teenager defacing a military web site.
- Criminals hacking bank accounts of defense contractor.
- A nation stealing plans for a new jet fighter.
- Terrorists using the Internet for recruiting, fundraising, propaganda, and communications.
- A nation stealing intellectual property stored in the computers of another nation's commercial firms.
- A military operation in which computers are used as supporting elements.
- A criminal causing massive physical damage to a steel mill.
- A nation stealing all of a government's personnel records.
- A nation causing a commercial airliner to crash.

Not all are high-end actions that must be deterred.

Who is being deterred?

- The leadership of another nation?
- Rogue elements of another nation's government?
- Subnational terrorist groups (maybe state sponsored)?
- Criminal organizations?
- Individual criminals?
- Individual criminals working in loose affiliation?
- Private citizens (e.g., students taking final exams, lone hackers seeking fame and glory)

Response must be tailored to the threat actor.

Note: Many hacking services available for hire…

High-End Attackers vs. Low-End

- High-end attackers are qualitatively different by virtue of their greater resources—
  - Money (lots more money available)
  - Talent (best talent money can buy)
  - Time (patience is a virtue)
  - Organizational support and commitment (e.g., can call on intelligence services)
  - Objectives (not just money; sometimes seek classified info, sometimes seek data alteration in pursuit of national goals)

What response should be threatened to deter that action?

- Responses against undesirable cyber activity when state not involved

- Responses against undesirable cyber activity when state is involved

Also, consider blurring of state/nonstate lines – what about national leadership engaging criminal activities?

Possible responses when state not involved

- Criminal prosecution
- Diplomatic
  - e.g., revoke travel privileges
- Counterterrorism action
  - Kill perpetrators
  - Take other (covert) actions short of killing
    - Financial (e.g., freezing bank accounts)
    - Personal (e.g., harass or intimidate family members)

EO 13694 - Blocking the Property of Certain Persons Engaging in Significant Malicious Cyber-Enabled Activities

Possible responses when state is involved

- Diplomatic
  - Develop alliances or enhance cooperation with other nations
  - Break diplomatic relations
  - Seek UN approbation
- Economic
  - Sanctions
  - Trade retaliation (e.g., tariffs?)
  - Action against the private sector (company specific action?)
  - WTO complaint?

- Intelligence
  - Use intelligence gathering for economic purposes
  - Public embarrassment, exposure of dirty laundry of leaders, target family members
  - Take covert action to sabotage adversaries
- Military
  - Redeploy military forces
  - Conduct cyber response (e.g., take down adversary Internet)
  - Increase/decrease military cooperation
  - Sell arms (including, possibly, cyber weapons)
  - Hack adversary weapons systems

What can be done without provoking escalation?

What is needed to make a response credible?

- Doing it.
- Doing it visibly (i.e., visible to the world at large).
- Doing it quickly.
- Demonstrating (or at least discussing) capability (remember Dr. Strangelove).

Some observations about cyber deterrence in practice

- Large-scale cyberattack out of the blue that shuts down much of the U.S. economy
  - Much blowback to adversary nations (e.g., China)
  - Takes lots of skill to execute that only nations have, thus attribution likely
  - Large scale kinetic response probable
- Intermediate scale cyberattack that reduce effectiveness of US forces as part of larger crisis
  - Only under dire circumstances as seen by adversary
  - Requires lots of advance preparation (e.g., mapping, implantation of Trojan horses..)
  - Intelligence may be invalidated in preparation
  - Strong and effective coordination and high confidence in success is needed.
- Small scale cyberattack for harassment, small incremental advantages
  - Attacker selects time, place, and means of attack
  - Eventual success highly likely

On Escalation

The general story regarding escalation

- In principle, conflict has multiple stages
  1. Preparation
  2. Initiation of hostilities
  3. Escalation
  4. De-escalation
  5. Termination

- Most work to date focuses on #1 and #2 above. Little work to date on
  - #3: escalation and how to prevent/deter escalation
  - #4: de-escalation and how to facilitate/encourage
  - #5: termination and how to facilitate

Types of escalation (1)

- Deliberate escalation is carried out for specific purposes (e.g., to gain advantage, to signal intentions and motivations).
  - Will intended recipient notice a signal against ongoing background of cyberattacks?
  - What messages could one send using a cyber operation?
  - What might make a cyber operation better for signal sending than other mechanisms?
  - What means (e.g., easily reversible attacks) are available to signal intent to adversaries in cyberspace, and how might these means be used?

- Inadvertent escalation (mutual misunderstanding regarding thresholds). Communicating thresholds regarding activity in cyberspace is particularly problematic, in peacetime.
  - Active threat neutralization and exploitation can be interpreted as attack
  - How to define and communicate thresholds?
  - How do we keep tight control over lower-level personnel, who may have the ability to do things with provocative results as SOP.

Types of escalation (2)

- Accidental escalation can result from some unintended effect due to failure in weapons, operators, command, intelligence. Hard to gather adequate intelligence on cyber targets.
  - How can national authorities exercise effective command and control of cyber forces in a rapidly evolving conflict environment?
  - How will connectivity with agents be maintained? How will integrity of agents (and their environments) be ensured?
  - How will C2 be exercised when personnel can do small things with large impact, especially during crisis?

- Catalytic escalation occurs when some third party succeeds in provoking two parties to engage in conflict. Inherent anonymity of cyber operations make “false-flag” operations easier to undertake in cyberspace.
  - Misdirected retaliatory act intended to discourage further attacks is overtly offensive.

Conflict termination

- A negotiated conflict termination presumes the existence of an ongoing conflict to which the participants desire an end. A satisfactory conflict termination generally requires several elements:
  - A reliable and trustworthy mechanism that can be used by the involved parties to negotiate the terms of an agreement to terminate a conflict.
    - How to negotiate securely and privately when channels for communication may be compromised cyber channels?
  - A clear understanding about what the terms of any agreement require each side to do.
    - How to know where cyber weapons are deployed
  - Assurance that all parties to an agreement will adhere to the terms of any such agreement.
    - How much would Nation R tell Nation B about system and network penetrations it had made? Such information might be used by Nation B to prosecute an attack or defend itself more effectively against Nation R.
  - Capabilities for each party to verify compliance with the terms of a cease-fire.
    - Why would Nation B believe a claim by Nation R that R was complying with the terms of a cease-fire?
    - Overt or cooperative intelligence not likely to be believed
    - Covert cyber exploitation to gain intelligence likely to be misinterpreted if discovered
    - Patriotic hackers continuing
    - Differentiating background of ongoing “normal” hacking
    - No national identifiers on attack traffic

Some questions re escalation and termination

- What is/are enabling conditions for crisis stability in cyberspace?
  - Many incentives for striking first in cyberspace
    - Perishability of intelligence
    - Doctrine and logic calls for early use
  - Some incentives for self-restraint
    - Blowback and entanglement (may be hard to identify)
    - Mere usage of cyber is *not* inherently escalatory (as NBC might be)

- Effective C2 for cyber to manage a cyber crisis (own forces, patriotic hackers)?
- How cyber conflict might lead to kinetic conflict?
- How to verify a cyber cease-fire?
- How to do effective attack assessment (scale, nature of attack)?
- How to reassure/signal adversary regarding intentions?
  - Will anyone believe what is said?

Active Defense

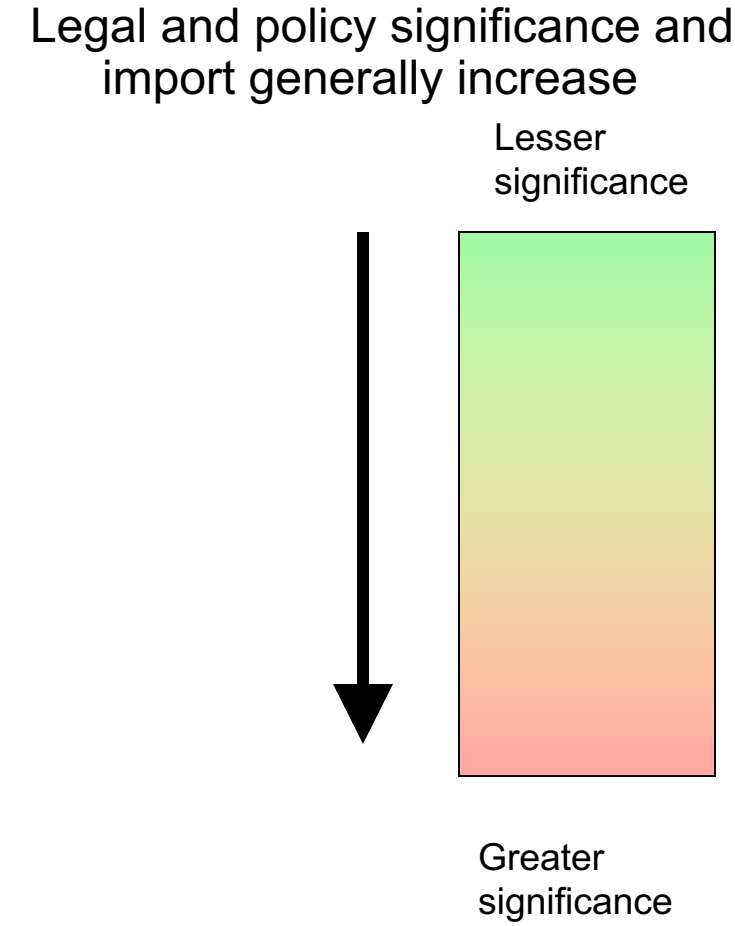
Active vs Passive Defense

- Passive defense
  - "measures taken to reduce the probability of and to minimize the effects of damage caused by hostile action without the intention of taking the initiative" (DOD definition)
  - Examples:
    - Firewalls
    - Antivirus software
    - Access control
    - Audits
    - Intrusion detection
  - General characteristics
    - Operate within domain of organizational custody
    - React to hostile intrusions
    - "Hostile" is recognizable

- Active defense…
  - Is a confused concept, used in multiple ways:
    - Anything outside the organizational domain
    - Anything non-cooperative
    - Anything harmful
    - Anything proactive
- In practice, active defense is anything that is not passive defense.
  - History of term goes back to 1980s Air Land Battle and defense of Europe (dynamic defense vs attrition defense) and missile silo defenses
- Note well: connotations of “active”, “dynamic” vs “static”, “passive”.

A hierarchy of active defenses

- within organizational domain
- in flight
- on adversary systems
  - identification and logical location
  - Exploitation/exfiltration of intelligence/forensic information
  - Damage to adversary system
    - Preemption
    - Threat neutralization in real time
    - Retaliation subsequent to threat actions



Categories of active defense

- Within organizational domain
  - Distract and delay intruder (e.g., honeypots)
  - Deceive intruder (e.g., files with misinformation)
  - Reroute/drop traffic from intruder
  - Slow responses to intruder
  - Collect forensic information on intruder to help law enforcement
  - Allow interactions only with whitelisted parties/software/computers
  - Rapid/dynamic reconfiguraton of defenses and networks

- Legal and policy issues relatively limited
  - privacy
  - possible interference with ongoing work
    - Version control (in misinformation, configuration management)
    - False positives in intrusion detection (e.g., legitimate remote user)

Three questions on active defense

- If prompt damaging response is required, what is the significance of automating the process?
  - Human in the loop and mistaken escalation?
  - How much “in advance” preparation is possible (e.g., pre-installation of trojan horses)?
- What should be the targets of a prompt damaging active defense response?
- If aggressive active cyber defense is good for defending military assets, why isn’t it good for defending non-military assets in government and in the private sector?

A depressing future

But I want to be wrong!

A condundrum and a prediction

- The condundrum of defense and deterrence
  - Can't do good defense, hence need to rely more on deterrence.
  - Can't do good deterrence, hence need to rely on better defense.
  - Effective damage limitation in cyberspace unlikely.
- A predicted consequence of the fundamental supremacy of the offense in information technology
  - No good defense
  - No good deterrence
  - No good counterforce for damage limitation
  - What is left for taking advantage of cyberspace?
- A prediction: low-level cyber intrusions (espionage and attacks) as far as the eye can see as a routine tool of statecraft.
  - Most likely scenarios: continued espionage, targeted Stuxnet-like attacks on individual facilities, nuisance attacks on a large scale. All low-level, sub-threshold.
  - Most important effects may be second-order, e.g., loss of confidence.
    - Consider Anonymous hacking into FBI/Scotland Yard conference call(

For more information…

Herb Lin

Center for International Security and Cooperation

Hoover Institution

Stanford University

650-497-8600

herblin@stanford.edu