THE DIGITAL DILEMMA: EVALUATING THE THRESHOLDS OF "USE OF FORCE" AND DISTINCTION IN STATE-SPONSORED CYBER OPERATIONS

by

JOANNA BELSESS

I. INTRODUCTION

Cyberspace has become a vital arena of conflict, representing a significant challenge to existing international law. State-supported cyber activities—spanning from disruptive actions such as NotPetya to damaging acts like Stuxnet—can cause societal damage similar to that of conventional warfare. Nonetheless, the essential legal frameworks, the UN Charter (jus ad bellum) and International Humanitarian Law (IHL) (jus in bello), depend on conventional ideas of physical harm, resulting in significant legal uncertainty. This study examines two key questions that emerge from this predicament:

- Does a cyber operation sponsored by a state that leads to substantial non-physical harm (e.g., incapacitating a financial system) constitute a "use of force" according to Article 2(4) of the UN Charter?
- How should the IHL principles of Distinction and Proportionality be implemented in cyber attacks aimed at dual-use critical civilian infrastructure?

Thesis Statement: This paper contends that although the physicality doctrine restricts the categorization of non-physical assaults as a "use of force," such devastating coercion should be acknowledged as an infringement of the non-intervention principle, warranting legal responses. Additionally, the implementation of IHL proportionality should be revised to necessitate a stricter standard of foreseeability when evaluating incidental civilian damage resulting from ripple effects on interrelated digital systems

Journal of Multi-Disciplinary Legal Research | Licensed under CC BY 4.0 — https://creativecommons.org/licenses/by/4.0/ | DOI: 10.5281/zenodo.17620716 | Title: The Digital Dilemma: Evaluating the Thresholds of "use of Force" and Distinction in State-Sponsored Cyber Operations

I. The Jus ad Bellum Threshold: Non-Physical Force and the UN Charter

-The Dominance of the Physicality Principle and Its Vulnerabilities

The conventional understanding of the "use of force" in Article 2(4) of the UN Charter is still significantly shaped by the physicality doctrine.
This prevailing perspective, primarily evident in the discussions around the Tallinn Manual Rule 11, claims that a cyber operation needs to generate effects similar to kinetic force—specifically, death, injury, or physical damage/destruction—to be considered a prohibited use of force. As a result, although the Stuxnet worm (which led to the physical breakdown of centrifuges) is generally viewed as having exceeded this limit, events like the 2017 NotPetya attack—despite inflicting billions in economic losses and severely disrupting critical civilian infrastructure—are usually regarded by major nations as simple acts of coercion or intervention, falling short of warranting a military self-defense response under Article 51.
This hesitation creates a perilous legal void: a state can be incapacitated by an unseen, non-physical cyber attack without having the acknowledged right to retaliate with force, thereby compromising the principle of sovereign equality..

A, Foundational Treaties: The Primary Legal Mandate

These documents demonstrate the political and legal consequences of state-sponsored cyber operations.

a. *Jus ad Bellum* (The Law on the Resort to Force)

The principles dictating the right to use force are defined primarily by the **UN Charter**. **Article 2(4)** establishes the core prohibition against the **"threat or use of force"** aimed at compromising a state's sovereignty, while **Article 51** preserves the inherent right to **"individual or collective self-defence"** following an **"armed attack."** These two articles serve as the principal legal

instruments in this paper, against which the factual consequences of disruptive cyber incidents, such as **Stuxnet** and **NotPetya**, are measured to determine if the high legal threshold for triggering the prohibition on force or the right to self-defense has been met.

b.*Jus in Bello* (International Humanitarian Law - IHL)

The lawful conduct of hostilities is regulated by IHL, codified most notably in the **Geneva Conventions and their Additional Protocol I**. These instruments mandate the analytical structure for the second research question. Crucially, **Article 52** outlines the principle of **Distinction**, requiring combatants to target only military objectives, and **Article 51** defines **Proportionality**, which strictly limits the acceptable level of incidental civilian harm. Interpreting the language of these provisions is essential to develop a framework for regulating cyber operations directed at **dual-use civilian infrastructure.**

c. State Responsibility and Countermeasures

To account for cyber operations that do not rise to the level of armed attack, the **Articles on the Responsibility of States for Internationally Wrongful Acts (ARSIWA)** provide an authoritative, though non-treaty-based, framework. This document is central to the argument on **non-intervention**. ARSIWA allows for the conclusion that even cyber operations falling below the **"use of force"** threshold constitute an **internationally wrongful act**. This classification legally empowers the victim state to utilize **lawful countermeasures** (e.g., diplomatic responses or economic sanctions, as observed following the **NotPetya** incident) that do not involve military action, thereby lending legitimacy to non-forceful state responses.

B.Authoritative Secondary Instruments: Defining the Consensus

While binding legal obligations are derived from foundational treaties, scholarly consensus is indispensable for interpreting their application in the complex digital domain. **The Tallinn Manuals (particularly the 2.0 edition)** function as the premier non-binding resource, synthesizing the prevailing expert views on how existing international law applies to cyber

operations. These manuals are instrumental in establishing the interpretive benchmarks against which the sufficiency of current law must be assessed:

ii)Establishing the Jus ad Bellum Threshold

The Manual articulates the dominant, yet inherently restrictive, perspective on the prohibition of the use of force. Rule 11 sets forth the standard that classifying a cyber operation as a "use of force" under Jus ad Bellum generally necessitates consequences that are physically destructive or kinetic-like—a concept widely recognized as the physicality doctrine. This rule is central to the paper's framework, as it defines the high legal obstacle that operations causing profound non-physical harm, such as those witnessed in the NotPetya incident, typically fail to clear, thereby validating the necessity for a legal expansion of the term.

iii)Framing IHL Applicability

For the regulation of hostilities (Jus in Bello), the Manual confirms the direct applicability of core principles. **Citing Rules 92 through 100**, it affirms that essential International Humanitarian Law (IHL) obligations, including Distinction (the targeting of military objectives) and Proportionality (the constraint of collateral civilian harm), are fully incumbent upon actors in cyber conflict. These rules define the conventional legal requirements before the subsequent analysis critically examines their capacity to effectively manage the complex and often unpredictable reverberating effects associated with attacks on interconnected, dual-use digital infrastructure

C.State Practice and Responsibility Frameworks

3.1.,The UN GGE Reports: Bridging the Digital Legal Divide

The **UN Group of Governmental Experts (GGE) Reports** are not legally binding treaties, but function as indispensable **consensus documents** crafted by a selected panel of governmental specialists. They are pivotal in cyber diplomacy because they successfully translate the abstract

tenets of international law into a practical framework for the new domain of cyberspace, thereby establishing a shared normative foundation for state conduct .

a. Anchoring Law in Cyberspace

The most profound contribution of the GGE's work (particularly the 2013, 2015, and 2021 reports) lies in definitively settling the question of **applicability**. The reports unanimously affirm that the **entire UN Charter**—including the pivotal **Article 2(4)** prohibition on the use of force—**applies in full to state activity in cyberspace**. This critical consensus immediately pivots the legal debate away from *whether* international law governs digital acts, to the far more challenging question of **interpretation**—specifically, how to define the functional **threshold** that equates a cyber operation to an illegal "use of force."

b. Establishing Norms and Expectations

The GGE reports transcend existing treaty law by establishing a foundational set of **voluntary, non-binding norms** of responsible state behavior. These norms serve as collective expectations, promoting stability and predictability. Key among them are:

- The obligation to refrain from **knowingly allowing** one's territory to be used for internationally wrongful cyber acts.
- The duty to protect one's **critical national infrastructure** from cyber threats.

These agreed-upon norms contribute significantly to a collective sense of order and provide benchmarks against which national cyber policies are developed.

c. Affirming the Rules of Conflict

The GGE solidified the legal status of warfare in the digital realm by affirming that **International Humanitarian Law (IHL) is fully applicable** to cyber operations undertaken

during situations of armed conflict. This clarification is foundational to legal scholarship regarding the rules of engagement, providing the necessary underpinnings for interpreting and applying IHL principles like **Distinction** and **Proportionality** to the unique challenges of networked, dual-use infrastructure.

4. Directing State Practice

By achieving consensus among a diverse group of major geopolitical and cyber powers (including permanent Security Council members), the GGE reports effectively provide an **international framework that directly guides national legal and policy development**. They function as a common reference point, helping individual states integrate international legal obligations into their domestic cyber defense and offense doctrines, thereby harmonizing global state practice.

III.Applying Jus in Bello Principles: Distinction and Proportionality

This section transitions to the framework governing the conduct of hostilities (*jus in bello*) and critically assesses the efficacy of **International Humanitarian Law (IHL)** in addressing cyber operations, particularly those aimed at a state's vital infrastructure.

3.1. Distinction: The Challenge of Dual-Use Cyber Systems

The IHL principle of **Distinction** requires warring parties to strictly differentiate between **military objectives** and **civilian objects**, targeting only the former (**Additional Protocol I (API) Arts. 51 & 52**). This fundamental separation is undermined in cyberspace, where essential infrastructure—including power grids, telecommunication backbones, and financial networks—is inherently **dual-use**, simultaneously serving both indispensable civilian life and military operational needs. The ambiguity is heightened when a cyber operation causes mere functional disruption rather than physical destruction, complicating the legal classification of the target. To safeguard civilian populations and ensure IHL compliance, this paper advocates for a **rebuttable presumption of civilianity** for interconnected dual-use digital infrastructure. This proposed standard would legally obligate the attacker to demonstrate a clear, direct, and immediate military advantage, effectively increasing the protective threshold for the civilian population.

3.2. Proportionality and Unpredictable Cascading Harm

The principle of **Proportionality (API Art. 51(5)(b))** prohibits an attack if the expected incidental civilian damage or casualties are deemed excessive relative to the anticipated military gain. This assessment hinges critically on **foreseeability**, a concept severely strained by cyber conflict. Unlike conventional kinetic attacks, cyber operations targeting networked systems initiate highly complex and often unpredictable **cascading failures**—termed **"reverberating effects."** These failures can spread rapidly across disparate sectors (e.g., a network disruption leading to failures in water distribution or healthcare systems). Since accurately forecasting the full extent of this cumulative civilian harm is inherently compromised, the traditional proportionality test is rendered insufficient. To restore the principle's relevance, IHL must institute a **heightened standard of foreseeability** for cyberattacks on critical infrastructure. This standard should impose a legal duty to anticipate the **maximum plausible collateral damage**—including indirect economic and life-sustaining impacts—due to the known fragility of networked systems, ensuring that failure to meet this elevated due diligence requirement constitutes a breach of IHL.

IV. Conclusion and Recommendations

The analysis of state-sponsored cyber activity demonstrates that the current corpus of international law struggles to safeguard the objectives of the **UN Charter** and **International Humanitarian Law (IHL)** in the digital arena. The continued reliance on legal interpretations derived from kinetic, physical warfare generates severe **accountability deficits** and exposes civilian populations to significant new risks.

4.1., Synthesis of Core Findings

This research established two foundational findings. Firstly, the strict adherence to the **physicality doctrine** when determining a **"use of force"** under *Jus ad Bellum* leaves states

vulnerable to highly **coercive cyber operations** (such as the NotPetya incident) that inflict systemic damage without achieving physical destruction. This interpretive rigidity creates a critical legal vacuum, necessitating the robust application of the **non-intervention principle** to ensure legal accountability for such acts. Secondly, current IHL principles are structurally inadequate for managing conflict in cyberspace. The inherent **dual-use nature** of vital infrastructure, coupled with the complex and unpredictable spread of **"reverberating effects,"** severely compromises the required assessment of **Proportionality** as mandated by the Geneva Conventions.

4.2. Policy and Legal Reforms

Based on these conclusions, the following specific reforms are essential to bolster the regulatory capacity of international law concerning cyber warfare:

- **Reinterpreting the Prohibition of Coercion:** Member States, via forums such as the UN General Assembly or the GGE, ought to formally clarify that cyber operations causing **catastrophic socio-economic or political destabilization** constitute a clear breach of the customary **non-intervention principle**. This categorization would subject the perpetrating state to immediate and **lawful countermeasures** (under the ARSIWA framework), even if the threshold for the use of force is not met.
- **Mandating an Elevated Standard of Care in IHL:** Authoritative bodies, including the ICRC, should issue official guidance confirming that the proportionality calculus for cyberattacks on critical infrastructure must legally integrate the anticipation of **maximum plausible collateral damage**. This standard necessitates that military commanders account for **foreseeable loss of life** stemming from the disruption of essential life-sustaining services (power, communication, water) due to systemic volatility, thereby ensuring a functional duty of **due diligence**.
- **Strengthening International Attribution Mechanisms:** Recognizing that the enforcement of both *Jus ad Bellum* and IHL is contingent upon the clear identification of the attacker, greater global cooperation and information sharing are paramount. Establishing a robust, objective, and transparent technical and legal framework for

attribution is the indispensable precursor to holding any state accountable for breaches of its international obligations in the digital domain.