調査・分析レポート

情報操作型サイバー攻撃の脅威（1）
―ディスインフォメーションを利用した情報戦の現状と課題―

独立行政法人情報処理推進機構　サイバー情勢研究室　研究員　長迫　智子

Ⅰ．はじめに

近年、世界各国で国家支援型サイバー攻撃が激増しており、日本においても港湾や病院などの重要インフラや学術機関などが攻撃を受けている。さらには、機能破壊型や情報窃取型のサイバー攻撃にとどまらず、世論操作、社会分断を企図する情報操作型のサイバー攻撃も増加しており、既存類型のサイバー攻撃と融合して実行されるハイブリッド型のサイバー攻撃へと高度化、複雑化が進んでいる。2016年の米国大統領選を契機として、各国の選挙における情報操作型サイバー攻撃による影響工作がみられ、またウクライナ戦争やイスラエル・ハマスの武力衝突といった有事においても、サイバー空間を通じた情報戦、認知戦が繰り広げられた。

こうした情勢を受けて、2022年12月16日に閣議決定された我が国の新たな安全保障戦略および国家防衛戦略において、国家戦略上はじめて、「認知領域における情報戦」という概念が言及された[1]。国家安全保障戦略においては、「偽情報等の拡散を含め、認知領域における情報戦」、国家防衛戦略においては、「認知領域を含む情報戦」および「ハイブリッド戦や認知領域を含む情報戦」といった記述がなされている。しかし、こうした新しい概念、用語について整理が十分になされないまま、情報戦や認知戦、偽情報（ディスインフォメーション）という言葉が一人歩きしている傾向にある。実際に、情報戦や影響工作を扱った先行研究においては、「情報領域における戦略的な力の行使を構成するすべての要素の定義についてコンセンサスが欠如している」[2]といった指摘がなされている。

そのため、本稿から3回にわたる連載を通じて、情報操作型サイバー攻撃が用いられている情報戦の様相を整理するとともに、新たな脅威である認知戦や生成AIによるディスインフォメーションの問題を取り上げ、情報操作型サイバー攻撃の全体像と課題を議論する。

[1] 国家安全保障会議『国家安全保障戦略』2022年12月16日、24頁。(https://www.cas.go.jp/jp/siryou/221216anzenhoshou/nss-j.pdf)
国家安全保障会議『国家防衛戦略』2022年12月16日、10頁、20頁。
[2] P. Brangetto and M. A. Veenendaal, Influence Cyber Operations: The use of cyberattacks in support of Influence Operations, p115, 2016 8th International Conference on Cyber Conflict (CyCon) (2016), p.115.

Ⅱ．情報操作型サイバー攻撃を利用した情報戦の概観[3]

１．ハイブリッド戦の構造

上述した安保戦略等だけでなく、近年、ハイブリッド戦という用語が各種報道で改めて脚光を浴びている。特に2022年２月に始まったウクライナ戦争を契機として、現代戦の形態としてハイブリッド戦争が改めて注目されている状況である。これは決して新しい用語ではなく、2000年代には米陸軍の教範にすでに盛り込まれており、ウクライナ戦争の背景の一つである2014年のクリミア危機の際にも、ロシアの戦略についてハイブリッド戦争という概念を用いて一定の評価がなされていたことが指摘されている[4]。ハイブリッド戦争とは、軍事・非軍事的手段を組み合わせる（＝ハイブリッドさせる）ことによって敵国や非友好国に対しての攻撃が行われる戦争ということが原義である。このことから、テロやゲリラ戦、プロパガンダといった古くからある手法の組み合わせもハイブリッド戦の一つとなる。

また、米国の統合作戦に係る教範においては、情報戦は「我の情報及び情報システムを防護し、敵の情報及び情報システムに影響を与える戦い」[5]と定義されており、20世紀ごろまでは、地形や気象情報、通信情報の攪乱やこれらに係る情報システムの防護に焦点が当てられており、心理戦などもあくまでも戦闘のためにデザインされたものであった[6]。

しかし、インターネット技術の隆盛により、サイバー空間上での戦いが大きな影響力を持つようになったことで、陸、海、空といった伝統的な戦闘領域および宇宙領域に加え、サイバー領域が新たな第五の戦場として認識されるようになった。それにより、サイバー戦がハイブリッド戦争の一角を成すようになる。サイバー戦におけるサイバー攻撃は、相手の情報システムを攻撃することで機能破壊を目的とする機能破壊型サイバー攻撃、相手の情報を窃取し金銭詐取や影響工作に利用しようとする情報窃取型サイバー攻撃、そして偽の情報や歪曲された情報（＝ディスインフォメーション）を流布することで相手の社会を分断し、国家の意志決定や民主主義の価値観を害する情報操作型のサイバー攻撃などを代表的な類型として主に六つの類型[7]に大別される（次章で詳述）。このようなディスインフォメーションを用いた情報操作型のサイバー攻撃は、SNSやマイクロターゲティングによるウェブ広告といった新たなウェブサービスを利用することで、単なる偽の情報の流布だけでなく、我々の認知を攻撃し、選挙時の投票行動や政治行動に影響を与える影響工作の一つであることが認識されるようになった。すなわち、第六の戦場としての認知領域、そして認知戦の登場である。

並行して、情報戦についても、軍事上の情報システムの攻防や心理戦、欺瞞作戦から射程が推移していき、広く市民社会への世論誘導や影響工作といったアプローチが含まれるようになった。更には、こうしたディスインフォメーションの累積が攻撃者に有利なナラティブを意図的に形成し、影響工作としてさらに大きな流れを生む。このナラティブの戦いにおいては、われわれの認知に影響を与えやすいナラティブとして、陰謀論が一つの脅威にもなっている。このような複雑化するハイブリッド戦の様相を整理したものが下図１となる。

---

[3] 本章については、以下の拙稿のⅠ章をもとに加除修正、再構成を行ったものである。
長迫智子「認知戦情勢に鑑みる対日本の攻撃アプローチの検討」『戦略研究』第34号、2024年3月、3－26頁。

[4] 小泉悠「ウクライナ危機にみるロシアの介入戦略：ハイブリッド戦略とは何か」『国際問題』658号(2017年１，２月号)、2017年、38-49頁。

[5] Robert J. Bunker, “Information Operation and the Conduct of Land Warfare,” The Land Warfare Papers, No. 31 (Arlington, Association of the United States Army, 1998).

[6] 菊地茂雄「米軍における情報戦概念の展開（上）――ソ連軍「無線電子戦闘」（REC）から「情報環境における作戦」（OIE）へ」『NIDSコメンタリー』第267号（2023年7月20日）、3-4頁。

[7] 大澤淳「サイバー領域の安全保障政策の方向性」『新領域安全保障』株式会社ウェッジ、2024年1月、185頁。

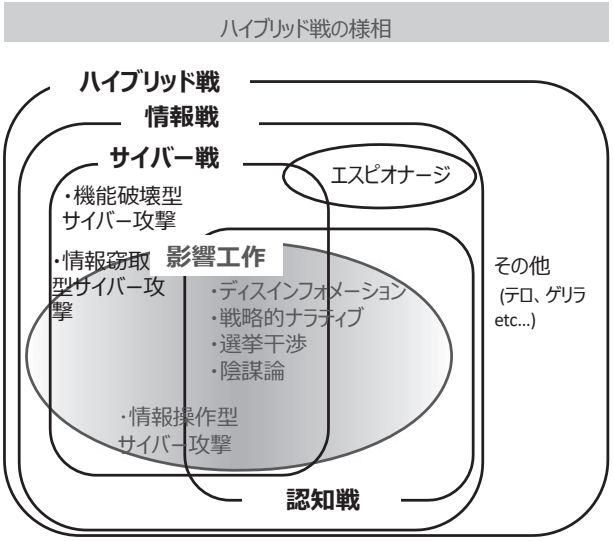

図１．ハイブリッド戦の様相（上図は右記初出図を筆者にて訳したもの。長迫智子「我が国の認知安全保障の確保を目指して」『宇宙・サイバーと先端技術研究会報告書（仮）』中曽根平和研究所，発行予定日未定.）

上図のとおり、認知戦や認知領域という語を用いても、それは単にわれわれの認識や言論空間にとどまるものではなく、サイバー領域が情報戦や認知戦の基層となって影響を与えていることに注意が必要である。これらは単独の戦いではなく、現代の安全保障においては、サイバー安全保障と認知領域安全保障を両輪として確保することが喫緊の課題であるといえる。

２．影響工作の定義

本節では、情報戦や認知戦下において主要な活動となる影響工作の定義を整理する。ここにおいては、NATO の先行研究[8]における定義が、サイバー攻撃とセットになった影響工作の実態を適切に表していると考えられ、そうした分析を補強する関連研究[9]もあることから、本稿では NATO の定義を採用する。

影響力工作は各国が敵対国に対して影響力を行使するための工作活動の一つであるが、軍事作戦に限定されるものではなく、外交の場も含め、あらゆる種類の紛争の一部となりうる。原則としては、非軍事的（非キネティック）な手段を用いて敵の意志力を削ぎ、意思決定を混乱させ、制約し、公的支持を弱めることで、発砲することなく勝利を達成する活動をいう。これには、平時であれ武力紛争中であれ、国家やその他の集団が対象となる聴衆の行動に影響を及ぼすために行う、あらゆる努力が含まれる。したがって、ソフトパワー活動を含む、情報領域におけるあらゆる活動の総称である。しかし、影響工作はソフトパワーの行使だけに限定されるものではない。武力紛争や軍事作戦の一環として行われる秘密活動や侵入活動も含まれる。これは、侵入的なサイバー能力の使用の可能性を含んでいる。よって、影響工作は、平時、危機、紛争、紛争後において、国家の外交、情報、軍事、経済、その他の能力を協調的、統合的、同期的に活用し、外国の対象者の態度、行動、意思決定を促進する活動であるといえる。

さらにこの影響工作は、Inform & Influence Operations (IIOs)（情報影響工作）、Influence Cyber Operations (ICOs)（影響工作のためのサイバー攻撃）、Information Operations (IOs)（情報工作）の三つの類型に区分できる。それぞれの定義は以下の通りである。

Inform & Influence Operations (IIOs)（情報影響工作）
…行動、発言、信号、メッセージを通じて、特定の対象者に情報を提供し、影響を与え、説得するための行動。
Influence Cyber Operations (ICOs)（影響工作のためのサイバー攻撃）
…対象者の態度、行動または意思決定に影響を与えることを意図して、サイバースペースの論理層に影響を与える作戦。
Information Operations (IOs)（情報工作）
…軍事作戦中に、情報関連能力を統合的に使用すること。他の作戦と連携して、敵および潜在的敵対者の意思決定に影響を与え、混乱させ、腐敗させ、または簒奪するために、軍事作戦中に情報関連能力を統合的に用いること。敵対者および潜在的敵対者の意思決定に影響を与え、混乱させ、腐敗させ、または簒奪し、同時に自国の意思決定を保護する。この IOs の定義については米国国防省の定義に依存しており[10]、軍事上の心理戦や欺瞞

[8] P. Brangetto and M. A. Veenendaal, Influence Cyber Operations: The use of cyberattacks in support of Influence Operations, p115, 2016 8th International Conference on Cyber Conflict (CyCon) (2016), pp.113-126.
[9] Sean Cordey, “Cyber Influence Operations: An Overview and Comparative Analysis,” Center for Security Studies (CSS) (2019).
[10] S Department of Defense, Directive 3600.01. (2 May 2013), p.12.