国際政治学における合理的抑止理論とサイバースペースへの応用

――サイバー抑止の実現可能性を考える――

Rational Deterrence Theory and Its Applicability to Cyberspace:
Considering the Feasibility of Cyber-deterrence

周　　橋

ZHOU Qiao

愛知大学国際コミュニケーション学部　非常勤講師

Faculty of International Communication, Aichi University
E-mail: zhouqiao513@hotmail.co.jp

要　旨

国家や悪意ある非国家主体によるサイバー攻撃への危機感が高まる中、伝統的な抑止政策が果たしてサイバースペースにも対応できるのかについての議論も盛んに行われるようになっている。本稿は、合理的抑止理論の諸前提とその論理をまとめ、それらがサイバー攻撃に対する抑止にどのような示唆を与えるのかを、サイバー安全保障研究に関するこれまでに展開されてきた議論をもとに明らかにする。それにより、攻撃の帰属や比例性原則といったサイバースペースでとりわけ深刻な問題を中心に、従来の手法をサイバー抑止に当てはめることの問題点や、合理的抑止理論のサイバースペースへの適用可能性を模索する。

Abstract

With the growing sense of urgency about cyber-attacks by states and malicious non-state actors, there is a growing debate about whether traditional deterrence policies are really up to the task in cyberspace. This paper summarizes the various assumptions and logics of rational deterrence theory and identifies what implications they have for deterring cyber-attacks based on discussions that have developed to date

on cyber security research. By focusing on particularly serious issues in cyberspace, such as attribution of attacks and the proportionality principle, it explores the problems of applying conventional approaches to cyber deterrence and the applicability of rational deterrence theory to cyberspace.

Ⅰ．はじめに

2022年に始まったロシアによるウクライナ侵攻は、戦争の勃発とその予防を説明してきた安全保障における伝統的な合理的抑止理論を今一度振り返る機会となった。合理的抑止理論は国際政治学や地政学において、どのようにして国家同士が緊張を含みながらも平和を継続的に保てるか、またはどのような条件下で平和が破られるのかを示すものである。合理的抑止理論は数多くの戦争を説明・予測しており、この理論が示すように、国際社会の現状は抑止を宣言しあう国家間の駆け引きによるぎりぎりの綱渡りにより、危機下において常にシビアなバランスを保ちながらも、一方では危機を回避できない危険性を払拭できずにいる。戦争抑止のため、同盟などのメカニズムが構築されてきたが、国際政治の情勢は常に流動的であり、覇権国である米国一極体制の相対的衰退化によって不確実性は増すばかりである。

さらに、21世紀の国際社会はインターネットの時代に突入し、戦争と平和の在り方も変わりつつある。軍事行動の一環としてのサイバー攻撃や、サイバー戦に準ずるような攻撃も多発しており、大事に至らないようなサイバー攻撃やエスピオナージといった工作に至っては数えきれないほどである。現在多くの研究者と政策決定者の頭を悩ませている問題は、従来の抑止政策でサイバー攻撃を阻止できるのか、また、従来の抑止理論が果たしてサイバースペースの事象を説明できるのかという点にある。本論文は国際政治学における合理的抑止理論を再考し、そのポイントを整理したうえで、合理的抑止理論のサイバースペースへの応用の可能性や問題点を模索する。

Ⅱ．安全保障における抑止

「抑止」自体は国際政治学者と歴史家によって十分に研究されてきた概念であり、また現実に国家の安全保障戦略に深く根付いた実践でもある。相手に何かさせるように強要するのではなく、抑止とは相手に何かの（往々にして軍事）行動をさせないように説得することである（Schelling 1966）。合理的抑止理論は抑止を次のように定義している：意図的に戦争のリスクとコストを操作することにより、相手の費用便益計算に影響を与え、行動を思いとどまらせることである。抑止それ自体は非常に広い概念であり、攻撃国とその政策決定者による他国への攻撃決定を左右する様々な要因を包含する。被攻撃国の軍事力や報復措置の信憑性だけでなく、国際社会の反応といった非軍事的な要素も抑止の成否、つ

まり攻撃発生の有無に影響を与える。ミアシャイマー（Mearsheimer）が言うように、抑止とは、究極的には軍事行動により生じると「感じる」政治的利益と軍事的・非軍事的な様々なリスクとコストとの関係を表す全ての事柄と広く捉えることができる（Mearsheimer 1983）。これについては多くの合理的抑止理論家たちが、軍事行動の結果がもたらす攻撃国への代償に焦点を当ててきた。例えば、シェリング（Schelling）のような早期の理論家は信憑性ある効果的な脅しに着目し、それによる「結果への恐怖」が好戦的な行為に及ぶことを防ぐと主張した。

1．合理性について

そもそも合理的抑止理論はなぜ「合理的」とされているのだろうか。それはこの理論が合理的選択アプローチといくつかの基礎的な前提条件を共有するためである。例えば、①主体としての好みと選択肢は外因的である。これはつまり、国家が欲するもの、達成したいことやその為の選択肢が、国家を取り巻く環境によってあらかじめ決定されているということである。主体の行動はそれらに応じて期待効用の最大化を目指す。②主体は一元的である。つまり、国家の政策決定者の行動は、国内の政治環境や政策決定過程、またそれらの違いに影響されない（Achen & Snial 1989）。言い換えれば、どの政策決定者も常に外部の刺激に反応し、どの国家も見た目は違えど機能は同じビリヤード球のようなものであるとされる。

これらの前提条件により、主体の行動の動機と選択肢の評価を単純化することで、研究者たちは抽象的かつ倹約的な抑止の理論を論理的に組み立てることができるようになった。これらに加え、激しい紛争というシナリオも多くの場合で前提条件となっている（Morgan 2003）。つまり、勝利の好機が見えたら敵は躊躇わずに攻撃をしてくるという冷戦的なシナリオである。すなわち、これは危機的な緊急抑止についての理論であり、平常時の一般的な抑止については未だあまり議論されておらず、理論化やその検証もあまりなされていない（Huth 1999）。

更に、被攻撃国による報復のほうが防衛よりも注目されるということも、この理論の特徴である。防衛力は「拒否による抑止」という概念に包含され、端的には攻撃を試みても無効化されたり効果的でなかったり、防御側の激しい抵抗にあったりし、戦闘における損耗が攻撃で得られると思われる政治的利益よりも大きい場合、危機下で攻撃国が譲歩し手を引いたり、そもそも危機に繋がるような要求や脅しにうったえないといったりするものである。すなわち、抑止理論を「合理的」とするのは、戦争に繋がる可能性の高い切迫した危機下において、攻防両者ともに損得打算的であり、相互の判断はランダム性・偶発性の産物ではなく、理性に基づいたものと仮定されているためである。

2．「拒否による抑止」とその問題点

防御力を高めることにより獲得できる「拒否による抑止」は、近年のサイバー抑止の議論で再び注目されるまで、長らくあまり研究者の関心を集めなかった。その理由として、第一に戦争自体のコストが上がった点にある。今日戦われるような戦争はクラウセヴィッツの時代に戦われていたような戦争とは大きく異なる。二度の世界大戦からもわかるように、戦争での損失は大きく上がり、とりわけ核兵器の導入によりグレート・パワー間での戦争はもはや考えられないほど高価なものとなった。戦争のコストが高くなればなるほど、攻撃国は「相手を負かせて利益を得られるのか？」と考えるよりも、「もしもやり返されたら？」と考えざるを得なくなる。

したがって、焦点は「戦場でどちらがどう勝つか」から「如何に相手に最大限の打撃を与えられるか」にシフトした。つまり、継戦能力と防衛力による拒否から約束された道連れの報復へと注目が移っていったのである。理論上、報復の脅威に信憑性があり、またその報復攻撃の度合いが十分にシビアであれば、抑止側は防衛にさほど力を入れなくともよいということになる。核抑止という考え方がそれの一番鮮明な例だろう。興味深いことに、スナイダー（Snyder）のような初期の理論家は報復能力よりも拒否力に価値を見出していた。彼によれば、抑止側の意図の不透明さ、それを更に不明瞭にする価値観の違いや不合理性、または感情的行動など、そして報復手段がそもそも限られていることにより、報復攻撃の信憑性は概して低いとみなされた。したがって、地上戦力をいかに多く展開できるかが抑止側の意思の強さを反映し、軍隊が目の前にいるということ自体が攻撃国に多くを伝えられると思われた（Snyder 1961）。

また、防御兵器と攻撃兵器の境がはっきりと区切られていないことで、防衛力に支えられる「拒否による抑止」と相手にコストを科す反撃能力・攻撃能力の根拠となる「報復による抑止」の区別も難しいのが現状である。城壁と攻城兵器のように攻守の用途が明確に分別できたのははるか昔のことで、多くの兵器が攻撃にも防御にも使える[1)]ため、攻撃国は常に相手が報復攻撃用の武器を持っていることに警戒せざるを得ない。更に、技術の発展により一部の戦争領域や兵器によっては攻撃優勢、つまり同じリソースでは攻撃兵器のほうが安上がりだったり、防御が攻撃に比べて極端に高価だったりする。レヴィ（Levy 1984）は攻撃・防御どちらが優勢なのかを攻防率（Attack/Defense Ratio）、つまり特定地点にある敵陣の防御を攻略するために必要な兵数で測ることを提案した。もちろん、抑止が想定する攻撃の全てが領地の奪取に始終するわけではないため、これを更に広く解釈し

1）この点についてはジャービス（Jervis 1978: 167–214）やガルトゥング（Galtung 1984: 127–139）を参照されたい。ジャービスは戦車や対空砲などを例に防御・攻撃兵器を区別することの難しさや完璧な定義は存在しないと指摘しつつも、防御の本質は自国領の保護であるとし、他国領を脅かさずにこれを達成できる短距離兵器を防御兵器とした。

て必要な兵数ではなく必要な攻撃リソースとしよう。その場合でも核兵器の冷戦時代やサイバー攻撃、または小型化したドローン攻撃の21世紀を考えれば、それらの攻撃を凌ぎきるための防御リソースの高さが窺えるだろう。核弾頭を搭載した大陸間弾道ミサイル（ICBM）への有効的な対抗手段は未だ確立されておらず、またサイバー攻撃も攻撃に必要な資本の安さに比べてそれを全て完璧に防ぐためのコストの高さを考えれば、拒否力頼みの抑止だけでは攻撃国を説得できない可能性が高い。言い換えれば、報復による抑止に説得力を持たせるための担保である攻撃能力を保持するほうが魅力的なのである。

3.「報復による抑止」のメカニズム

抑止を試みる国が宣言する報復が実際に行われると信じられるかどうかという報復の約束（脅し）の「信憑性（Credibility）」は、抑止の成否に大きく関わる問題である。軍事能力を持っているだけでは攻撃国に行動を思い留まらせることは難しい。国家にはそれぞれ互いに自らの能力と意思を正確に伝えたくないという思惑と、相手がそうするのではないかと疑う余地があるためである。言い換えれば、ポーカーゲームのように自分の手札をあえて強く思わせたり、逆に弱く見せたりするブラフの動機がある。これはまさに不完全情報における交流に組み込まれたジレンマであるため、抑止側が宣言する報復には常に信憑性の問題が付きまとうことになる。

これはさらに拡大抑止にとってはなおさら致命的で、第三国が同盟の一員として抑止側を守るというコミットメントの信憑性を攻撃国はより疑うこととなる。この「信じるに足る脅し」がどのような条件、もしくは政策により得られるのかについては、その後多くの学術的関心を集めた（Fearon 1997; Powell 1990）。最後に、安定した抑止関係には抑止側の能力と意思を理解しそれらを効率よく伝達させるだけでなく、その他色々な付属的条件も存在する。例えば、大量破壊兵器の存在や軍事技術の発展は国家の先制攻撃能力を大きく上げることがある。これは緊急抑止下にある両国を先制攻撃へと誘惑する。兵器の質や種類、軍の統制、勇み足な同盟国など、安定した抑止を妨げる要素は多くある。

これらを踏まえ、合理的抑止理論は戦争に直面する危機的な状況において戦略的な交流での合理的なアクター達がどのように行動するのかを理論立てていった。ここで最も単純な交流のパターンを考えてみたい。抑止側（防御側や被攻撃側とも言われる）は自分あるいは特定の第三者が攻撃国に攻撃されるのを防ぐことを目標とする。攻撃国が先に動き、攻撃に踏み切るか否か選ぶ。続いて抑止側が動き、報復するか否かを選ぶという流れである。ただし、抑止側の能力（軍事力）と攻撃された後の反撃の意思について、攻撃国は完全情報を持っているわけではない。こうした設定のもと、合理的抑止理論の最たる主張は、抑止側の報復の脅しに信憑性がある場合、すなわち攻撃国が報復は現実的かつそれによる損失が攻撃による利益を上回ると信じる場合、攻撃は起きないというものである。同

時に合理的抑止理論はこの条件が満たされない場合には抑止の失敗を予測する。つまり、もし抑止側に報復の手段や意思がないと攻撃国が信じた場合、攻撃は発生する（Achen & Snidal 1989; Morgan 2003; Quackenbush 2011）。

実際にはどのような要素が抑止の成功につながり、攻撃国の行動を未然に防げるのだろうか。攻撃国の計算と抑止の成否を左右するとされる一般的な要素には以下のものがある。

1．抑止側の関心：争点となっているものが抑止側にとってどれほど大事なものなのか、或いはそれを失うリスクがどれほど高いのか。
2．抑止の宣言の有無：抑止側が攻撃を受けた際、何をするのかを明言しているか。
3．抑止側の拒否力・防衛力：攻撃を受けた際、被害を軽減、或いは無効化する防衛力はいかほどか。
4．抑止側の報復能力：攻撃を受けた際、報復として攻撃国に被害を与える能力はいかほどか。
5．抑止側の報復の信憑性：攻撃を受けた際、あらかじめ宣言した通りに報復を遂行する可能性はいかほどか（Huth 1999）。

簡潔的に言えば、これらの要素は全て拒否による抑止（deterrence by denial）と報復による抑止（deterrence by retaliation）の能力を左右する国家の武力とそれを行使する意思の大小である。合理的抑止理論によると、どんな潜在的な攻撃国も攻撃による期待効用を算出できる。攻撃による期待利益が攻撃に踏み切らないときのそれを下回る時、または攻撃による期待損失が上回る時、攻撃国は抑止される。そのうえ、攻撃の期待効用はある結果の効用と意思決定者が思うその結果が起こる確率との乗算で示される（Huth & Russett 1984: 500）。

これらの要素はそのほとんどが今度は攻撃国と被攻撃国の軍事力のバランスに影響される。もし抑止側が強ければ、攻撃国は戦争において勝利を掴み辛く、戦闘に際して多くのコストを支払うことになる。軍事的行動の結末に対する恐れが抑止のキモであるならば、攻撃に対する強力な防御というものは大いに役に立つ。同時に、（先制攻撃を生き延びられる）抑止側の強力な軍隊はそのまま効果的に報復攻撃に使用することができる。つまり、報復による損失は攻撃による期待利益を帳消しにしたり上回ったり出来る。言い換えれば、攻撃国が真剣に恐喝や紛争にうったえるのはそうすることによる期待効用が十分に高い場合であり、期待損失が期待利益を上回る時にこそ抑止される。同じように、抑止側の同盟のような第三者も、上記のような論理で攻撃国の行動を思い留まらせることができ

る（Bueno de Mesquita 1980）[2)]。

ミアシャイマーは更にこのバランス・オブ・パワー（勢力均衡）に基づいた説明により詳細な条件を付けくわえた。国家間のパワーの差だけが抑止の成否を決めるのではなく、戦略や戦闘技術など、攻撃国の攻撃リスクを左右する要素に影響される。もし攻撃国が軍事的優位に立っていた場合、どのようにして攻撃するかの選択肢が多く与えられ、攻撃リスクは低い（Mearsheimer 1983: 63–65）。政策決定者は通常、消耗戦、電撃戦、限定戦[3)]を選び、前者ほど攻撃リスクとそのコストは高い。つまり、後者ほどそれを用いるコストが低いため、電撃戦や限定戦により軍事目標を達成出来ると考えている相手は抑止されづらいという。一方で、攻撃国と抑止側の軍事力が拮抗している場合や攻撃国が劣る場合、低コストの作戦では目標を達成できない事態が多く、抑止されやすい。つまり、戦力に勝るドイツが電撃戦でもってポーランドへの侵攻を考えた時には、もはや抑止でそれを踏みとどまらせることはできなかったはずだというのである。

4．信憑性について

以上から、弱者は強者の行動を制限できないという初期のリアリスト達が真っ先に辿り着いたアナーキカルな国際システムにおける真理に再び帰着する。しかし、幸運な事に国際社会の歴史は強者による絶え間のない恐喝（Compellence）と弱者による抑止失敗の連続だけではない。その原因の一つに抑止側による「信憑性ある脅し」（Credible Threat）が挙げられる。報復・懲罰による抑止の論理とその現実での成否を担うこの概念と、どうしたらこれを達成できるのかは多くの研究者によって議論されてきた。

信憑性とは即ち「信じるに値する特性」である（Morgan 2003: 15）。報復の脅しは正しく伝達される必要があるが、更に大事なのはそれが攻撃国の期待を変える程度に信頼されることである。この信憑性ある脅しは、例として、核抑止の文脈において最も明確に問題となる。核兵器の与え得る損害とその後の相互破壊というシナリオを考えると、核保有国が実際に核兵器を使用するとは信じがたい。したがって、核報復が実際に行われると攻撃国が信じない場合、いくら核兵器を保有していても効果的な抑止に繋がらない恐れがある。例えば、ナラーン（Narang）は、核保有国ですら時折抑止に失敗することを指摘し、エスカレーションや全面的な戦争を防ぐのは核の保有事実ではなく、政府の抑止態勢（Deterrent Posture）であるとした（Narang 2013）。核兵器が戦争の道具として運用され、

2）ブエノ・デ・メスキータ（Bueno de Mesquita 1980）は同時に、リスク許容度の高い国が自らよりも強大な国に冒険的な攻撃を仕掛ける可能性も指摘している。これは軍事力の差だけが主体の期待効用を左右する訳ではないことを意味する。

3）Limited aim strategy の筆者訳。電撃戦と違い、その達成目標は通常抑えられており、敵軍の殲滅を伴わない戦略を指す。

なおかつ先制攻撃が可能な状態であり、更にそのことが他国に明白であるという非常に好戦的な（一方的エスカレーションの）抑止態勢でなければ核を盾とした抑止に信憑性はない。

信憑性の獲得がどのようにしてなされるのかについては、多くの抑止理論家が検証を続けてきた。例えば、シェリングは当事国の意思や戦争の結果への不確実性が抑止側に戦争へのリスクを操作する機会を与えるとした（Schelling 1966）。これはいわゆる瀬戸際外交として知られている。つまり、信憑性は精強な軍隊からでも迫真の脅迫からでもなく、双方にとって望ましくない結末へ偶発的に転げ落ちる可能性のデモンストレーションから得られるのである。また、抑止側は全面戦争や核戦争のリスクをつり上げるために限定的な戦争を仕掛けたり、そうすることを脅したりすることができる。小規模の戦闘が当事者たちの制御から離れ、急速にエスカレートすることがあるということを示す。このような「何かを天に委ねる」ような戦略は攻撃・防御両方の脅しに信憑性を与えることができるとした。シェリングに倣い、パウエル（Powell 1990: 505）は限定的報復という作戦、つまりシビアだが制御の取れた制裁により抑止の信憑性が高まると主張した。初めの脅しの信憑性が不十分であった場合、抑止側は何らかの罰を攻撃国にあえて与えることにより、将来的な罰への脅しに現実味を持たせ、攻撃国に抑止側の意思を再考する機会を与えることができるというわけである。フェアロン（Fearon 1997: 69）も同様に、いくつかの作戦により、「コストのかかる信号（Costly Signaling）」を攻撃国に送る際の信憑性を高められると述べている。例えば、「意思の固まっていない発信国にはできないであろう、自らにコストやリスクのかかる脅し」を送ることがその一つである。前述の理論家とは違い、フェアロンは政策決定者が国内において負うコスト、すなわち政策や決定が公約通りに運ばれない場合に、政治家が「事後に」払う観衆費用とバーゲニング（Bargaining）の過程で国家が支払う軍事動員などの「事前」の費用に注目した。民主主義国の政策決定者はとりわけ国内の観衆費用に敏感であるため、自ら退路を断つような「両手を縛る（Tying-hands）」作戦は相手国への脅しをより信憑性あるものにできるというわけである。

信憑性はまた、抑止側の関心にも左右される。攻撃国は通常、抑止側の関心を 2 種類の情報をもとに推測する。一つは直接的領土や第三国との軍事・経済的繋がりの損失に関する情報であり、二つ目は抑止側の国内政治より生じるバーゲニング行動に関する情報である（Huth 1999）。もし関心が高ければ、国内外の観衆費用をつくり出すことで抑止側はよりコストのかかるシグナルを送り、それが却って抑止側の政治的関心を高めることとなる（Fearon 1994）。

抑止側の関心から生じる信憑性問題は、軍事同盟のような拡大抑止のシナリオにおいて最も顕著になる。通常攻撃国は自国の領土を守ろうとする抑止側の関心の高さを疑うことはない。しかし、抑止側が本当に身を切って第三者国を守るのかという疑問は依然として

残る。歴史を辿れば、いかに明文化・組織化された協定であっても同盟を貫き通せない同盟国が存在していることがわかる（Leeds & Anac 2005; Signorino & Tarar 2006）。例えば、ハスとルセット（Huth & Russett）は抑止国とその庇護国の間の距離感、つまり両国の間にどれ程の軍事的・経済的繋がりがあるのかが攻撃国の脅威認識に影響を与えると示している（Huth & Russett 1984）。また、軍事的・政治的に関連した国々の中で拡大抑止が当てはまるペアとそうではないペア間の比較により、核保有の超大国による抑止の有効性と地域的な現状維持能力も見られることが分かった（Weede 1983）。信憑性は抑止側の脅しに対する攻撃国の認識的反応に依るため、同盟を守るというコミットメントを守らないと思われるかもしれないし、約束されていた報復が実行されないと思われるかもしれない。更には、攻撃国が戦闘や報復のリスクを負うことを厭わなかったり、無行動による国内の政治コストを背負いきれないと判断したりすることもある。以上より、防衛と報復の宣言だけでは抑止を確実なものにすることができないことがわかる。ただし、それら（宣言）無しでは抑止が成功することはない。

５．抑止の実践例

近年のアメリカを例に、抑止における軍事的・非軍事的な手段による損得勘定の操作がどのようになされてきたのかを紹介したい。冷戦終結後、真に覇権国となったアメリカはクリントン政権のもとで従来の核抑止、とりわけ警報即発射（Launch on Warning）という姿勢を和らげ、その一方で生物化学兵器に対しては核による報復を示唆するようになった。従来の抑止と異なる点として、抑止の対象がソビエト連邦のような同じ大国ではなくアメリカに比べてはるかに弱い国家であったこと、攻撃国と同じ方法や程度の反撃（生物化学兵器や通常兵器）ではなく別種あるいは釣り合いの取れていない報復（核兵器）の示唆が挙げられる。そしてこれらアシンメトリー性と不均等性を覆い隠すかのようにアメリカは意図的曖昧性ドクトリンを用いるようになった。上記のように報復をする「かもしれない」し、しない「かもしれない」とすることで、抑止側はコストを減らしつつ脅しの信憑性を確保し、攻撃国にとって最悪の万が一を想定せざるを得ない状況を作り出すのが目的であったが、意図的曖昧性ドクトリンは実際の報復の信憑性や核拡散等の観点から物議を醸した（Betts 1998; Sagan 2000）。

しかし、本当にテロリストなどの非国家組織や彼らを匿い大量破壊兵器を開発しようとするならず者国家を核兵器等で抑止できるのだろうか。抑止は攻撃国と抑止側の双方の合理性と一定のコミュニケーションを前提とする。サイコロを振るってランダムに意思決定する場合や、いわゆるマッドマンには利益とコストの計算ができないとされ、また失うべきものがない者たちには果たして報復による事後コストの強制を軸にした脅し文句はどれ程効果があるだろうか。

9.11によりパラノイアに陥ったW・ブッシュとその政権では2002年よりいわゆるブッシュ・ドクトリンというテロリストやならず者国家に対し、必要であれば、先制攻撃を辞さないという抑止よりも先制攻撃に重きを置く対外政策を敷き、翌年にはイラクへと攻め込んだ。また、逼迫した危機的状況での先制（preemptive）攻撃ではなく将来的にアメリカの安全を脅かす可能性を予防する（preventive）攻撃を正当化したことで大きな波紋を呼んだ（Gupta 2008）。

オバマ政権以降、アメリカは宇宙やサイバー空間を第4・第5の戦場と位置づけ、それらドメイン（領域）での抑止に伝統的な他ドメインの手段や対応を当てはめるクロス・ドメイン抑止を今までよりも強調するようになった[4)]。戦争行為のラインを超えないいわゆるグレー・ゾーンやハイブリッド型攻撃の脅威が増したことがきっかけとなった。また、これらの新しい脅威に対して、同ドメインでの報復にも限界があった。例えば、サイバー攻撃に対する報復のハック・バックの効果は可視化されにくく、また想定通りの損害を与える確証に欠けており、また、未だ技術的に難しい宇宙での戦闘は実践可能な報復として考慮できない。これら背景のもと、サイバー攻撃に対して異なる手法である「制裁」でコストを科すといった試みと議論が盛んになった（Lindsay & Gartzke 2019）。

ハック・バックとは、不正アクセスなどのサイバー攻撃に対して、その取り締まりや被害回復を目的に被害者から（主に被害企業が）加害者へ同様の攻撃をすることを指す。日本では「不正アクセス禁止法」で規制されており、米国では「コンピュータ詐欺・悪用法」によって規制されている。しかし、米国では2019年に、相手からのサイバー攻撃を「阻止・妨害」する目的の「積極的サイバー防御」の名目のもとで婉曲的にハック・バックを認める「積極的サイバー防御確実性法」が下院に再提出され、議論を呼んだ。こうして抑止の対象とそれを支える約束された報復が多角化するにつれ、シグナリングの鮮明度が落ち、また報復の程度も曖昧になっていると指摘できる。例えば、自国の核施設に対して他国がサイバー攻撃をした場合、どのドメインでどれ程の報復が適当なのかを判断することは難しい。報復をしない、または報復が攻撃に対して極端に弱い（攻撃国に課すコストが低い）と将来的な抑止の弱化に繋がる恐れがある。一方で、報復が過剰な場合はエスカレーションを招くばかりか第三国や国際社会から報復の正当性を疑われかねない。超大国同士の「核兵器には核兵器を」という伝統的かつ単純な抑止から、抑止の対象に小国・非国家が加わり、抑止するものに軍事行動以外の攻撃が加わり、更に報復の選択が異ドメインを含むようになることで、シグナリングが弱くなり、信憑性が損なわれ、結果として

4）クロス・ドメイン抑止は最近の政権でより強調されてきた一方で、その概念自体新しいものではなく、冷戦下より実際に度々政策として登場してきたと言える。詳しくはリンドセイとガーツキー（Lindsay & Gartzke 2019）の第2章を参考されたい。

抑止のパフォーマンスに影響を与えるのではないかと危惧される（Healey 2018）。

冷戦を経て、アメリカは「フリーサイズ（One Size Fits All）な抑止」から特定の主体に合わせた「注文仕立て（Tailored）な抑止」に変わり、戦争行為の抑止からグレー・ゾーンな攻撃行為の抑止に、そしてドメインを超えた抑止へと複雑化していった。現在のバイデン政権では、それらすべてを「統合」する抑止、すなわち統合抑止（Integrated Deterrence）が唱えられるようになった。2021年の米国防長官によるスピーチでは統合抑止の定義について、外交や制裁を含む多様な選択肢、先端技術の応用、クロス・ドメイン、同盟国との連携などが強調された（U. S. Department of Defense 2022）。しかし、侵略を抑止するための脅し文句として、最初に外交的な孤立と経済制裁を持ち出すことは意思の弱さの表れであり、結果としてロシアのウクライナ侵攻を抑止できなかったとの批判がある（Gallagher 2022）。また、戦争でのエスカレーション防止に重きを置いたバイデン政権の統合抑止は作戦になんら示唆を与えず、ただファンシーな造語で修飾した空の概念であるとの指摘もある（Jackson 2022）。

続いて最新の日本の抑止政策を見てみよう[5)]。2022年12月16日国家安全保障会議・内閣決定された日本の「国家安全保障戦略」を見てみると、上記の拒否（防衛）による抑止から報復（反撃）による抑止への若干の移行がみられる。

> 我が国への侵攻を抑止する上で鍵となるのは、スタンド・オフ防衛能力等を活用した反撃能力である。…（中略）…このため、相手からミサイルによる攻撃がなされた場合、ミサイル防衛網により、飛来するミサイルを防ぎつつ、相手からのさらなる武力攻撃を防ぐために、我が国から有効な反撃を相手に加える能力、すなわち反撃能力を保有する必要がある。この反撃能力とは、我が国に対する武力攻撃が発生し、その手段として弾道ミサイル等による攻撃が行われた場合、武力の行使の三要件に基づき、そのような攻撃を防ぐのにやむを得ない必要最小限度の自衛の措置として、相手の領域において、我が国が有効な反撃を加えることを可能とする、スタンド・オフ防衛能力等を活用した自衛隊の能力をいう。こうした有効な反撃を加える能力を持つことにより、武力攻撃そのものを抑止する（内閣官房 2022: 17–18）。

これを読む限り、アメリカのあらゆる手段（all means necessary）でのアプローチとは違い、日本は同ドメインでの対応と限定し、また「相手からのミサイル攻撃」の場合にのみ反撃をすると宣言するに留めている。更に、外交・同盟重視な姿勢はバイデン政権のそれ

5）本稿では2022年12月に決定・公開された「国家安全保障戦略」だけを見るが、更にさかのぼって全体の変移を辿る場合は千々和（2022）を参考されたい。

と似ているが、反撃能力に関するレトリックはとりわけ興味深い。上記に示した拒否と報復という抑止の大枠からすると、抑止の文脈での反撃という行為は事後（ex post）に相手にコストを科すことを事前（ex ante）に約束することである。即ち、この文言は日米同盟がもたらす拡大抑止による報復、つまりアメリカだけによる報復から方向転換をし、当事国として自らが報復を約束し抑止力を高める試みと考えられる（平和・安全保障研究所 2022: 1, 3, 12）。しかし、その一方でこれは「防衛」すなわち「拒否」、そして「必要最小限の自衛の措置」としての「報復」を強調している。そこに有事の際のエスカレーションへの警戒や専守防衛に慣れ切った国内世論の反発への懸念を窺うことができる。マッドマンが未だいない日本で、約束された報復が最も信憑性を持つのはシェリングの言うように、瀬戸際においてリスクを偶然に委ねられるかに依る。エスカレーションをも厭わず更にアクセルを踏めるかどうか、抑止の宣言に対するコミットメントが問われる。

Ⅲ．合理的抑止理論への批判

合理的抑止理論への批判は主に合理的選択の前提への不満や、主張が実証ではなく演繹より生じていることに対して向けられている。合理的抑止理論は異なる攻撃国の差異や変化、また関心を形成・再形成する政治的環境を無視し、その結果度々歴史と政治に無関心、あるいはそれらより乖離していると考えられてきた。この理論は端的には主体の認識に関するものであるにもかかわらず、主体の根本的な認識形成のプロセスを扱っていない（Jervis, Lebow & Stein 1985）。合理的抑止理論で想定されている攻撃国のリスク受容傾向は常に効用最大化・リスク受容型であるが、それにより異なる主体の在り方を考慮から除外してしまうとレボウとステイン（Lebow & Stein）は指摘している（Lebow & Stein 1989）。それによる弊害は早くから指摘されていた。例えば、主体の関心・目標、更には主体間の関係に関する固定化された仮説について、ジャービス（Jervis）は主体同士が常に敵対的な状況で脅迫しあい、妥協や褒美など他の政策への展望が存在せず、またこうした悪性関係が変わらないという前提を実世界の政策に適用した場合の危険性を指摘している（Jervis 1979）。

理性的な政策決定というこの理論の最も基礎となる前提条件が、とりわけ危機下においてしばしばなされてこなかったことをいくつかのケーススタディは明らかにしている。例えば、1969年から1973年までエジプトによるイスラエルへの 5 回にわたる軍事力行使の試みと抑止の結果をもとに、ステイン（Stein）はエジプトの首脳部は合理的抑止理論が予期するような方法では行動しなかったことを示した（Stein 1996）。彼らはイスラエルの関心を読み違えたのみならず、自他の戦力をも読み違えた。攻撃の決定は軍事力の優劣や報復の可能性によって算出されたものではなく、何もしないことへの恐れによるところが

多かったのである。たとえ攻撃以外の選択肢が机上に存在していたケース（1969年のエジプトによるイスラエルに対する消耗戦争）でも、起こり得る結果の吟味を怠った。1971年から1973年のエジプト首脳部の計算はとりわけ顕著に合理的抑止理論の想定する内容に相応しくなかった。期待効用を考慮する際に、ペイオフの大小にばかり注目し、そこに至る確率を軽視したのである。こうした選択決定におけるバイアスは抑止の失敗を生んだだけではなく、抑止が成功したケースでも見られた（Lieberman 1995）。これらにより、合理主義者らの確率論的な仮定や過度に単純化した結論の有用性の限界が指摘された。よって、リスクの操作による信憑性という合理的抑止理論の核心的な論理が攻撃国に果たしてどれ程影響を与えているのかを再度問わなければならない（Jervis 1979: 310–311）。また、ステインブルーナー（Steinbruner）は抑止行為の説明について、意思決定に際して主にネガティブなフィードバックが果たす役割を強調した合理的抑止理論に替わり、サイバネティク行動理論を提唱した（Steinbruner 1976）。これは、政策決定者は期待効用最大化のために確率論的な判断を下すのではなく、前回の決定による結果をフィードバック情報として、都度与えられた状況に役立てるというものである。

理性のもっともらしさだけでなく、抑止側が攻撃国にコミットメントの信憑性を伝えるコミュニケーションであるシグナリングの過程も実際は合理的抑止理論の想定するようにはいかないことがある。政策決定者の脳内では、様々な認知的・感情的なハンディキャップが（文化的な違いと共に）存在し、合理的な計算を阻害する。そして、それらのノイズはコミュニケーションの質に大きな影響を与える。情報の統合と処理に対する過信は、政策決定者が相手の意図を正確に読み取る能力を過大評価したり、比較的最近の抑止結果から得られる証拠に多く依存したりする原因となる。

また、政策決定者は価値のトレードオフを見誤る傾向にある。つまり、自分が支持する政策に夢中になり、それが全ての価値の次元で代替策よりも当然（しかし非合理的に）優れていると考える傾向にあり、代替策にかかるコストを正しく評価することを妨げているのである（Jervis 1982–1983, 1989）。同様に、新しい情報を自分の既存の信念に同化させる傾向も指摘される。新しい情報を誤認したり、解釈しなおしたりして、既存の信念を強化し首尾一貫した見解を維持することを優先させてしまうケースがしばしばある。その結果、相手に対するイメージが変わりにくく、行動の変化を察知できず、最終的には抑止に戦略の幅や想像力を欠くことになる。また、人々の認知プロセスのはたらきによるバイアスやショートカット以外にも、国内政治の要請などのような政策決定者のニーズによる影響も無視できない。脅威を過小・過大評価し、非合理的、あるいは自滅的な政策を採用し、正当化・合理化するこれらの意欲的・非意欲的なバイアスの数々は、合理性を前提とした理論では考慮されない。合理的抑止理論では、国家は対立と譲歩の期待コストと利益のバランスをとる必要があるが、認知分析や様々なケーススタディによれば、国家はそれ

らを暗黙のうちに不用心に行うにとどまり、またしばしばバイアスに影響されるのである。

このようにケーススタディで新たに発見されたアノマリーや理論への批判を認め、アシャンとスナイダル（Achen & Snidal）は理想型の説明として、「理論はいくつかの破綻を予想している」と反論した。「報復の脅威がなく、その脅しに信憑性がなく、または得られるものよりも報復の規模が小さいことが原因で抑止が失敗した時、この理論はそれらを完璧に予測してきた」（Achen & Snidal 1989: 152）。確かに、整合性のとれた論理で構築された理論の高度な抽象性は、他のどのアプローチよりも多くのことを説明することを可能にし、その政策的意味合いも深い。例えば、この理論は政策決定者に、防衛力のみに頼ることの危険性を教えている。ケーススタディは様々な種類の一般化できない変数や物事の注目すべき側面を明らかにするが、実際それは全ての事象を説明する統合理論とは程遠く、モデルに変数を追加していくと理論化の試みが遠のくばかりであるのも事実である。

合理的抑止理論は今日に至るまでその地位を保っている。冷戦後、核抑止の有効性についての議論は落ち着きを見せたが、国家はこの理論が想定する緊急・一般・直接・拡大といった抑止を積極的に実践している。初期の抑止理論は、抑止の成功に寄与する正確な要因や政策決定者の合理性にどれ程期待できるかについて、確信を持つほどの実証的裏付けに欠けていた。しかし、説得力のある代替理論がないため、紛争管理・危機外交・軍縮・同盟形成などの議論において、合理的抑止理論は依然として高い外的妥当性を維持している。理論の演繹的かつ簡略的な性質により、技術や兵器の進歩とともに様々な仮説が検証されるようになった。したがって、世界が情報化時代に突入し、国家が攻撃的なサイバー能力に投資するようになるにつれ、学者たちが次のような問いを立てるのは自然なことである：抑止はサイバー攻撃に対して有効か？　信頼できる報復に依存する合理的抑止理論の論理は、潜在的なサイバー戦争の文脈において通用するのか？と。

Ⅳ．サイバースペースにおける抑止理論の可能性

1．サイバー抑止の必要性

サイバースペースは、米国統合参謀本部により「情報環境の中で、情報技術（IT）インフラと常駐データの相互依存ネットワークからなる領域」と定義されており、国家の軍事活動や国民の日々の社会経済活動に不可欠な要素となっている。デジタル化された触れることのできない仮想世界、例えばインターネット、通信ネットワーク、そこで共有される情報などと、情報を蓄積し伝達するための個々の機器やプロセッサ、イントラネット、ケーブル、それらを支えるユーザーなどといった物理的基盤が含まれる。サイバースペースは無国籍でもグローバル公共財でもないが、利用者が従来の様式により制限されないと

いう意味で、大きくグローバルなものでもある。民間の大規模なインターネットサービスプロバイダーは、ほとんどが国家の直接的な支配から一定の距離を取っており、通信は国境に大きくは縛られない。例外として、プロバイダーの厳格な管理、大規模なファイアウォールや国内の監視システムを有する中国のインターネット・モデルなどがある。ウェブに接続する企業や機構、組織などの内部ネットワークを守るため、インターネットを通して外部からの不正アクセス、また、内部ネットワークから外部への許可されていない通信などを防ぐファイアウォールは国家の高度なセキュリティシステムの一部と位置付けることができる。中国にとってのサイバーセキュリティは内政の安定を最優先の目標としており、必要であれば、外部情報と遮断することも辞さないからである（蔡 2019: 42–49）。

サイバースペースの特性である「共有性」と「接続性」によって、商業やコミュニケーションが盛んになる環境が整えられた。それにしたがい、サイバースペースは今までの通信や電子商取引の領域を超え、交通システム、銀行、医療、発電などの重要インフラなど、人々の日常生活のより重要な分野を統合するようになった（Singer & Friedman 2014）。しかし、同時に、情報の共有が容易になることは、犯罪者、ハクティビスト（Hacktivist）[6]、テロリスト、あるいは第三国・勢力といった無認可の主体が獲得すべきでないものを獲得しやすくなることを意味し、「接続性」の強力な二本柱である匿名性と越境性が、サイバースペースにおける法的執行を極めて困難にしている。これにより、サイバー攻撃やCNE（Computer Network Exploitation）[7]などといった悪質な行為は、個人のインターネット・ユーザーから民間企業に至るまで、様々なレベルで深刻な問題となり、最終的には今日の国家安全保障の最大の議題の一つとなった。

サイバー攻撃、もしくは戦争・戦闘でのサイバー的手段は、1990年代初頭から議論され始めた比較的新しい概念・現象であり、とりわけ2000年代以降その数と規模が大きく増した。ここで言うサイバー攻撃とは、コンピュータを利用して標的の情報システム、ネットワーク、またはサービスを破壊、破損、制御することである（Libicki 2009）。従来の軍事攻撃とは異なり、サイバー攻撃は地理的な近接性や国境に制約されることはない。また、攻撃の主体は国家直轄の組織、国家が支援する個人やグループ、愛国心の強いハッカー、特定の政治的目的を持ったテロリストや、イラクのサイバー兵士に悪戯で扮したカリフォルニアのティーンエイジャーなど多岐にわたる。更に、攻撃の対象も国家の軍事・重要インフラから民間企業まで様々で、攻撃の規模もケースによって大きく異なる。幸いなことに、サイバー攻撃による死者の記録はまだないが、経済的損失や社会混乱を過小評

6）ハクティビスト（Hacktivist）：政治的・社会的な主張や目的などのため、ハッキングを行う個人や集団。その行動はハクティビズム（Hacktivism）と呼ばれる。

7）CNE：国家間のサイバー戦の文脈において、相手国のネットワークに侵入し弱点を探ったり、データを盗んだりする行為。広義にはスパイ行為やエスピオナージともとらえられる。

価することはできない。攻撃の曖昧性と近年ますます高まる社会の情報技術への依存から、国家は他国やならず者がサイバーな手段を用いて、いわゆるサイバー真珠湾やサイバー9.11のような突発的被害の可能性を常に危惧している。サイバー攻撃の件数とその深刻さが増し、サイバー戦争に猜疑心をいだき始めた一部の先進国は、社会のデジタル化の度合いが高いため、互いに脆弱であるのみならず、伝統的に軍事的・経済的に弱い相手に対しても脆弱であることに気づいた。それゆえ、サイバースペースは戦争領域に統合され、サイバー防衛とサイバー抑止は全体的な防衛戦略の一部となった。

1990年初頭には既に、国家間のサイバー戦争やフリーランスのハッカーによるサイバー攻撃の潜在的危険性と、それらが既存の抑止政策や社会一般に及ぼす影響について、一部のセキュリティ研究者によって議論されていた（Arquilla & Ronfeldt 1993）。シロリ（Siroli）は1990年代半ばから後半にかけて、米国が潜在的な情報戦（Information Warfare）、つまり戦争遂行に関連して敵の情報資源を拒否、破損、破壊することを意図した活動、に対する公式の反応をどのように展開してきたのかを観察した（Siroli 2006）。そして、脆弱性から重要インフラの保護を強調する過激派とそうした危険がほとんど仮説であると考える懐疑派の間の議論がいかに複雑であるのかに焦点を当てた。IT インフラがどれ程脆弱であるかを大規模に評価することは非常に複雑な作業だと指摘する一方で、産業のデジタル進行が脆弱性の源泉であり、重要インフラのネットワーク化された情報システムへの依存度が高まることで、一度攻撃を受けると連鎖反応のように混乱やカスケードが起こる可能性があると結論付けた。

サイバー攻撃への懸念と恐怖にもかかわらず、サイバー抑止を正式な防衛政策として採用するまでに、国家はかなりの時間を要した。潜在的な脅威を認識しながらも、国家は長らくサイバースペースを私的な商業領域と位置づけ扱っていたのである。国家がサイバースペースにおける抑止について真剣に検討し、実践するようになったのは2000年代末頃からである。例えば、米国の場合、サイバー抑止政策は2009年にオバマ政権が発表した「包括的国家サイバーセキュリティ構想」に遡ることができ[8)]、そこで初めてサイバースペースを新たに国家安全保障の不可欠な領域と位置づけ、サイバー攻撃に対抗する適切な抑止政策の策定を求めた。また、NATO は2008年に初のサイバー防衛政策を採択し、NATO サイバー防御センター（Cooperative Cyber Defense Centre of Excellence）を設立した。これらは、ロシアが2007年にエストニアに対し、そして2008年にジョージアに対して行った積極的なサイバー作戦に大きく刺激されたためである。中国でも2010年代初頭より、

8）一方で米国によるサイバーセキュリティの追求そのものについては90年代末に遡ることが出来る。具体的には、1998年の大統領決定指令第63号（PDD-63）を通して、クリントン政権がアメリカの重要インフラに対するサイバー攻撃からの保護を目的に、2003年までに安全な情報システムの確立を要求したことがその始まりである。

攻撃型が主流だったサイバー防衛政策に大きな変化が見られた（Jiang 2019）。実際、2010年頃のイラン核施設へのサイバー攻撃事件の刺激を受けて、2014年頃には反撃性を備えたサイバー安全保障政策に本格的に取り組むようになった（孟 2014: 46–49; 黄2015: 145–158）。攻めの思想を強調し、サイバー能力を強制的な道具として、あるいは先制的攻撃の脅しとして使うのではなく、中国軍の戦略家たちは、米国からのサイバー攻撃を抑止するための強力なサイバー部隊を確立する必要性を認識し始めたのである。

２．サイバー抑止に関する争点

国際関係論の文脈におけるサイバー抑止の学術的な議論は、1990年代後半に技術が開発・実用化されるにしたがい始まって、2000年代後半から2010年代前半にかけて盛んに行われるようになった。議論のテーマは膨大で、サイバー抑止の必要性、その内容、方法、成功と失敗を決める要素、古典的抑止理論のこの新しい国家間紛争の領域への適用可能性などが挙げられる。その抽象度の高さから、合理的抑止理論の論理がサイバー領域でも成立すると多くの研究者が同意しており、同じ理論的枠組みに基づいてサイバー抑止が成功する条件を特定することに多くの努力が支払われ、新しい理論の構築にはあまり関心が向かなかった。

一部の研究者はこれに反対し、従来の抑止力として認められている要素（キネティックな軍事衝突のシナリオ、主体の合理性、攻撃源の特定、コミュニケーションなど）がサイバー攻撃には欠けており、したがって新しいドメイン専用の理論が必要であると主張している（Taddeo 2018）。しかし、合理的抑止理論のサイバースペースへの適用に懐疑的な人々は、サイバースペース特有の要因による抑止の失敗を、理論自体の失敗と同一視する傾向にあるようである。以下で詳説する攻撃源の特定問題や、自国のサイバー能力を明らかにすることへの消極性、つまりコミュニケーションの欠如は、報復の脅しへの信憑性を阻害するサイバー抑止に特有の問題である。つまり、これら要因による信憑性問題から生じる抑止の失敗は、合理的抑止理論によって十分に予測されているところである。また、サイバースペースでは比較的小さな組織や個人が強大な国家を攻撃することは自殺行為であるため主体の合理性が疑われるという主張も、攻撃者の意思決定がランダムに行われていたり、期待効用の計算に欠けていたりすることの証明にはならない。たとえキネティックな軍事行動に基づく理論が攻撃源の特定を当然のものと扱っており（従来の軍事紛争においても攻撃源の不確実性は存在する）、攻撃源特定の成否と信憑性の関係については特に言及していなくても、信頼できる脅威による抑止、あるいは信じられないような脅威による抑止の失敗という因果関係の主張は、サイバースペースにおいても適合するはずである。

今日、サイバー抑止の最も興味深い議論は、そのような戦略の有効性についてであり、

つまり、サイバースペースでの抑止が成功するかどうかというパズルである。研究者の間においてもコンセンサスが得られておらず、そのため推奨される政策には一貫性がない。戦争のコストが高くなるにつれ、防衛（城壁や砦の建設など）よりも、大量報復の脅威に重きが置かれるようになった。サイバー兵器の登場により、国家が攻撃力と防御力の構築に努める中、サイバー攻撃を抑止するために費用と労力をかけることにどれだけの価値があるのかが問われている。

懐疑派は、サイバー領域と他の領域は根本的に違うと主張する。アクセス性や匿名性など、インターネットの性質に起因する違いは、抑止に対するユニークな課題になっているというのである。これらの課題には、サイバー攻撃への相互理解・規範の欠如、レッドラインの不可知性、安価な攻撃参入コスト、サイバースペースにおける攻撃優位性、そして最も顕著な攻撃源帰属（攻撃源・攻撃発動者の特定）の問題などがある。一般に、懐疑派からみればサイバー抑止は可能ではない、もしくは極めて困難であると結論付けられる。なぜなら、抑止側が発する脅しはこれらの課題を考慮すると、ほとんど信憑性がないからである。

一方で擁護派は、代替案である防衛・拒否力増強の非効率性を指摘し、攻撃源の特定という問題は懐疑派が唱えるほど深刻ではないと主張する。確かに既存の事例は抑止の失敗を示唆しているように見えるかもしれないが、それでも深刻な国家間のサイバー戦争が欠如していることは、全体として抑止の有効性を示しているというのである。抑止は伝統的な国家間関係においても失敗をしており、それら失敗が国家を抑止政策から遠ざけるわけではない。現在、国家が関与したサイバー攻撃の事例が比較的少ないことから、理論的・仮説的な議論が多く、どちらの主張がより説得力があり、正確に現実を予測しているのかを判断することは困難である。

今日多くの研究者が受け入れている合理的抑止理論は、合理性の仮説とゲームから生まれ、そこから導かれる命題が検証され、反証から生じた批判をされることで発展してきた。同じ理論に基づくサイバー抑止も、それと同じ道をたどっている。このパズルが実証的に検証され、因果関係が解明されれば、既存の抑止政策をサイバースペースに適用するという現在の国家の動向を遡って評価することになるだろう。

先述のように、サイバー抑止をめぐる議論は1990年代に始まった。ハークネット（Harknett）は、冷戦時代に開発された抑止モデルは、この新しい問題には不十分な指針であると論じている（Harknett 1996）。効果への信頼性が高い核兵器とは異なり、通常抑止に関連する脅威のコストは、技術的、戦術的、そして運用的に様々であるため、攻撃側にはその効果を疑う余地がある。コストの大小を疑う余地があればあるほど、攻撃側は抑止側の脅しに挑戦する可能性が高い。ネット戦争（ハークネットはこれを社会的なつながりの破壊・混乱と定義）の文脈では、個人、組織、国家レベルのいずれに対する同種報復

も、限定的な損害しか与えられないため、攻撃側が感じる報復による損害、つまり報復の効果に対する疑念の問題が抑止の有用性を更に制限する。

一方では、攻撃側は報復サイバー攻撃を受けることによる損失が自身のサイバー攻撃より得られる利益よりも小さいと見積もる可能性があり、他方では、サイバースペースでは戦略的対称性を感じ取りにくいのも事実である（Blank 2001）。そもそも敵のサイバー能力を正確に推定することは難しい。これは、使用される武器の性質によるところが大きい。攻撃的なサイバー兵器の成功は、奇襲的な要素に懸かっているためである。サイバー兵器はコードの集合体であるため、一度使用され公開されれば、リバースエンジニアリングと複製が比較的容易にでき、キネティックな兵器のように希少資源を必要とせず、コンピュータと潜入要因だけで製造が可能である。そのため、国家は自らのサイバー能力を秘密兵器のように扱い、その暴露を極力防ぐ傾向にある。以上により、攻撃側は抑止側が持つ真の能力を、報復の意思とは別に、確実に評価することができないのである。つまり、サイバースペースでは抑止側が自分の能力を確実に示すことができないため、コミュニケーションの質が低下し、攻撃側の報復コストを計算する能力に影響を与えている。

上記の「争点性の問題」を克服するどころか、物理的な被害に欠けるサイバー攻撃への同種報復が果たしてどれ程攻撃側に被害を与えるのかについても不確実性が大きく、またどの程度の異種報復であれば適当なのか判断することも難しい。この「比例性の問題」は、サイバー抑止の文脈において特に顕著である。低い攻撃参入コストは、社会が情報システムに十分接続していない後進国内に潜在的な攻撃者を増やすが、サイバー攻撃を行うこれらのデジタル化されていない国家への同種報復は、期待するような打撃にならない可能性が高い（Libicki 2009: 26）。他方で、過剰な報復は正当性のない攻撃的行為と見られかねないため、抑止側は難しい立場に立たされる。また、攻撃側に対する過度に厳しい報復は、双方が望んでいないエスカレーションを招くこともある。更に、攻撃主体の多様化が進むと、特定の主体に合わせたテイラードな抑止が難しくなる（Kugler 2009）。脅威のコストが攻撃側に感じられるようにするためには、抑止側はどのような種類の主体が自分の抑止に挑戦し、彼らが科せられたコストをどのように評価するのかを考慮しなければならない。例えば、領土の完全性への脅威は国家には効く一方、個々のテロリストにはさほど響かない場合がある。

また、この比較的新しい戦場において、何が攻撃とみなされるのか、どのような行動が正当化できないとみなされるのかについて、主体間の共通認識や共通言語がないことが、抑止の失敗を助長していると指摘する学者もいる（Todd 2009）。例えば、ルポビッチ（Lupovici）は、帰属問題が慣習や国際規範によって作られたものであり、アプリオリに存在するものではないため、報復に先立って攻撃側を確実に特定する必要性を強調する規範が今後の主体間の行動次第で変化することがあるとコンストラクティビズム

（constructivism）[9]的な主張を展開した（Lupovici 2016）。更に、サイバー戦争に関する世界的な条約がないため、サイバースペースにおける暴力行為とは何か、攻撃の手段はどのようなものかという解釈の負担が残り、それらが抑止側の判断を迷わせ、抑止の結果を左右することも考えられる。国連、NATO、サイバースペースに関するロンドン会議[10]などの国際組織が、国家による様々な制約的サイバー規範候補を提案するプラットフォームとして機能するようになってきた（Mazanec 2015）。とは言え、サイバー能力の非対称な発展、様々な程度の脆弱性、サイバー兵器の潜在的有用性に関する理解の欠如は、主導して規範を作り従わせるような主体の登場を妨げている。

3．攻撃源の帰属問題

争点性、比例性、規範の欠如とは別に、サイバー抑止の議論の中心となった問題がある。それはアトリビューション、つまり攻撃源の特定と帰属に関する問題である。報復による抑止における信憑性と抑止の結果の直接的な因果関係を示唆する合理的抑止理論が展開する方程式に帰属という変数を加え、サイバースペースでの変化を多くの研究者が見つけようと試みてきた。限られたケースの分析と理論から示唆される予測は、以下のような結論を導き出した。

サイバー抑止の懐疑論者は、抑止側がサイバー攻撃を正しい攻撃源に帰属させることが困難であると攻撃側が認識し、それにより抑止側の報復の脅しを信じる能力を妨げていると主張する。比較的簡単に攻撃源を特定できるミサイルのようなキネティックな攻撃とは違い、サイバー攻撃は追跡が確かに困難である。ボットネット（あらかじめ侵入・感染を受けたコンピュータ）や VPN などの利用により、攻撃者は自分の本当の「住所」を偽り、別の場所から攻撃しているように装うことができる。こうした技術によって、情報が破損したり、受け取る情報が不完全であったりするため、抑止側は誰が自分を攻撃したのか結論が出ないままになってしまう（Blank 2001）。つまり、サイバースペースにおける攻撃の起源、攻撃者の正体、攻撃の動機を特定することは極めて困難なのである。更に悪いことに、通常兵器の入手に比べて、攻撃側はサイバースペースでの武器のアクセスに困るこ

9）コンストラクティビズム（constructivism）：国際関係論の基礎理論の一つ。人々のアイデア・認識・アイデンティティ・規範などの世界観や認識は社会的に構成されており、それらの制度や歴史的な文化により、国家間の関係も構成されていると考える。詳しくはウェント（Wendt 1992）を参考にされたい。

10）2012年11月1–2日、英国のロンドンにおいて、サイバー空間に関するロンドン会議が開催された。60カ国の政府機関他、国際機関、民間セクター、NGO 代表など約700名が参加した。議長声明では、サイバー空間での活動における 5 つの側面、すなわち、（1）経済成長と発展、（2）社会的便益、（3）安全かつ信頼できるアクセス、（4）国際安全保障、（5）サイバー犯罪について、それぞれ議論された内容が発表されるとともに、サイバー空間における政府機関、民間セクター及び個人の行動に関する提唱が行われた（外務省 2011）。

とは少なく、他者が使った戦術を簡単に再現できるため、抑止側を欺くための偽旗を立てやすい。攻撃者を正確に特定することは不可能ではないにしても、高度な技術的捜査、労力、時間が必要である。

これら不確実性のために、国家は迅速かつ効果的に、そして繰り返し対応することが難しい（Iasiello 2014）。では、この問題は抑止にどう影響を与えるのだろうか。アトリビューション問題の主な効果は、報復の実現可能性やそうすることの意思・望ましさについての疑念を生じさせることにある（Lindsay 2015）。攻撃側の能力、動機や関心がはっきりしない場合、抑止側は報復を躊躇する。なぜなら、無実の第三者を誤って非難し、不必要な紛争を引き起こす可能性があるからである。また、攻撃側の真の能力を誤って評価し、強者と弱者を間違えてコストのかかる反撃や比例性原則に則っていない罰を与えることになりかねない。これら抑止側が抱える葛藤の存在を知って、攻撃側は自然と抑止側の報復の脅しの信憑性を疑うことになる。つまり、サイバースペースを特徴づける帰属問題という新たな不確実性がある以上、抑止側が攻撃側にコストを科すことを事前に信頼できるかたちで脅すことは困難であるというのが、多くの研究者たちの共通認識となっている（Libicki 2018; Brantly 2018; Tor 2015; Elliott 2011; Schulze 2019; Jasper 2017; Denning & Strawser 2017; Singer & Friedman 2014; 土屋 2020など）。

リビッキー（Libicki）はこの問題、ひいては同種報復によるサイバー抑止の問題を端的にまとめている。まず、帰属の難しさは、サイバー攻撃が事実上どこからでも、あらゆる個人やグループによって、物理的な痕跡がほとんど残らない状態で行われることを可能にする技術に大きく起因する。「誰がやったのか？」という疑問は、捕まる確率が低いことを意味しており、この低い特定率を打ち消して相手を抑止するためには、高いペナルティを科してバランスをとる必要がある（Libicki 2009, 2012）。これは比例性の問題につながるだけでなく、第三者国の目に報復がどう映るのかという問題にもなる。そうでなければ、抑止によって約束された報復は、第三者の目には正当性のない侵略と映るかもしれない。つまり、抑止側は正しい帰属が必要であるだけでなく、攻撃側を含む他者に、こうした攻撃で特定され捕まり、そして報復を受けたのであって、単に攻撃側が嫌いだから報復と見せかけて攻撃をしたのではないと納得させる必要がある。

一方、攻撃に関する証拠を公開することで、第三者を納得させることができる一方、攻撃側が使ったのと同じ方法で報復することで、「そちらがやったことは知っている」という明確なメッセージを送ることができる。しかし、捜査の方法を見せてしまうことは将来的な防御力の低下に繋がり、それを見せずに報復することは自分が攻撃的であると思われることになる。このように、抑止側はこれらを同時にカバーすることが不可能と思われる難しい状況に置かれる。したがって、抑止は不可能ではないにせよ、困難であると言える。

４．サイバー抑止に替わるもの

この難問を踏まえ、学者たちはサイバー攻撃と戦うために、異なるマインドセットを開発し、国家が適応すべきドメイン特有のアプローチを提案した。大多数の学者は、抑止力を維持したまま防衛力の強化を図ることが、安全保障に更なる層（レベル・次元）を提供できると同意する傾向にある。例えば、ヤシエッロ（Iasiello）は、「サイバー領域には、抑止手段を開発するために必要な透明性と行為者の可視性が欠けている」と指摘し、悪質な行為者は帰属問題による報復を恐れることなく、高い接続性という環境が彼らの策略を助長するため、サイバースペースでの破壊活動を活発にさせると述べ、その結果、拒否による抑止の優位性を説いた（Iasiello 2014: 54）。エリオット（Elliott）は、攻撃後のフォレンジックにおいて、攻撃発生時の国際的安全保障状況の考察に頼るだけでは報復するための強力な根拠が得られないと指摘し、したがって、強力な防衛が脅威に対する主な回答であると述べた（Elliott 2011）。

サイバー攻撃に対する防衛は難しいとはいえ、核攻撃に対する防衛等のシナリオに比べれば、不可能ではないようである。サイバー防衛が高度化するにつれ、攻撃の準備段階で発見・妨害されるリスクが高まり、攻撃成功に必要なコストが冷や水のごとく高まる可能性を考えると、防衛の有効性も攻撃の脅威と共に割合増加することが言える。同様に、ブラントリー（Brantly）は、サイバースペースは実際には報復よりも拒否による抑止を優遇していると論じた（Brantly 2018）。ドメインそのものをコントロールする能力の確立に莫大なコストがかかる従来の防衛とは異なり、サイバースペースのすべて―そのネットワーク構造、ハードウェア、ソフトウェア、ネットワーク内外の個人のアクセス性など―は、防御側の意思・判断で比較的容易に操作することができるというのである。このことは、毎日米国に向けられる攻撃の成功率の低さからもわかると著者は言う。ジェスパー（Jesper）も同様に、強化された防御（ネットワークセキュリティ侵害のリアルタイム検知、被害分析、緩和など）と積極的な介入を組み合わせることで、抑止力の問題点を軽減できる可能性のある「能動的サイバー防衛」を提案している（Jesper 2017）。

これの主な焦点は、タイムリーな検出制御と修復を提供することにより、ネットワークセキュリティをより強固なものにすることである。そうすることで、サイバー攻撃は成功しにくくなり、また防御側にも対策案の選択肢が広がる。最後に、ライアン（Ryan）は、サイバー空間における敵対者の行動を阻止するには、防衛と罰則だけでは不十分であると主張し、規範やタブーなどを用いた名指し批判により攻撃側を特定の行動から遠ざけるような異なるアプローチが、罰や防衛の比較的弱い成果を強化できることを示唆した（Ryan 2018）。これらの戦略をまとめて実施することで、防御側は攻撃に対して社会的コストを大幅に上乗せすることができ、攻撃側の行動を抑止できるのである。増加するサイバー攻撃の数と現実の抑止の凡庸な成果を考えると、防衛をもっと重視すべきだという主張には

一定の妥当性がある。パスワードの厳格化、二段階認証（2FA）、ソフトウェアのアップデートなどの簡単な対策や、ネットワークの分離・遮断、適切な教育や指導などの長期的な方策でも、攻撃の成功率をある程度下げることが期待できる。

しかし、報復の仕組みがないまま防衛策に重きを置きすぎることも、深刻な問題を引き起こす可能性がある。最も注目すべき事実は、防衛は攻撃者の試み自体を止めるものではないという点にある。サイバー攻撃に対する処罰のメカニズムがない場合、攻撃者は成功するまで何度も攻撃を試みることができ、また恐らくそうするだろうが、一方で抑止側はそれら全ての試みを撃退し続けなければならない。サイバースペースは攻撃者にとって有利な性質があり、またサイバー兵器の製造・再生産が比較的低コストで済むため攻撃側にもメリットがあり、防衛側は個人レベルでも組織レベルでも穴のない高度な壁を築かなければならない（Sanger 2018）。たとえサイバースペースを防衛の敗北が構造的に不可避である攻撃優位の環境とは捉えず、ただ攻撃が持続するだけの環境とより楽観的に捉えても、状況はさほど好転しない。「攻撃持続とは、しかしながら、防衛が戦略的に単独で勝つことはなく、精々引き分けに終始するということである。防衛は戦術的・作戦的な成功を収められるが、攻撃は続き、敵との接触頻度は不変であり、防衛は守るべき地形や攻撃のベクトルが進化するに合わせて絶えず調整する必要がある」とハークネットとゴールドマン（Harknett & Goldman）はこう指摘し、「他の何か」の必要性を示唆した（Harknett & Goldman 2016: 86）。

では、報復による抑止が成り立たず、拒否による抑止が国家に必要な安全保障を十分に提供できないとしたら、我々は一体何をすれば良いのだろうか。報復による抑止の有用性を主張する研究者たちはこう答える、「やはり報復による抑止だ」と。こうした研究者たちは、帰属の難航は必ずしも抑止力を弱めるものではないと考えている。一般に、帰属問題は克服できる、あるいは帰属問題によって付与される不確実性は抑止のメカニズムをさほど妨げないという主張である。

前述のように、匿名性、したがって帰属の問題は、抑止力の有効性に悪影響を与え、より多くのサイバー紛争を引き起こす新しいレベルの不確実性を生む。しかし、リンドセイのように、帰属への理解を技術的なフォレンジックだけに限定しないようにと警告する研究者もいる（Lindsay 2015）。更に重要なことは、十分な能力と覚悟とを必要とする大規模なサイバー攻撃を有益と考える攻撃的主体が少ないこと、また、大規模なサイバー攻撃は結果的に手掛かりとなる潜在的な根拠や足がかりも多く残すことから、攻撃のリスクも高い。そのような重大な攻撃に対する抑止が帰属問題によって損なわれるとは限らないという主張である。同様に、クラークとランドウ（Clark & Landau）も、帰属問題は単に技術的な限界ということはなく、抑止は絶望的だと結論付けるのは過度に悲観的であると批判している（Clark & Landau 2010）。確かに多段階の攻撃はその性質上、法的管轄を跨ぐ傾

向があるため、帰属の試みは大きく妨げられるが、セキュリティ意識を高め、リターン・アドレスをよりよく国家に割り当てるための合意や条約といった外交手段は、この一見技術的に避けられない問題を大きく軽減できるという。

グッドマン（Goodman）は、2007年と2008年のエストニアとジョージアに対する国家主導のサイバー攻撃の事例を通して、帰属の特定は、我々が考えているほど難しいことではないと指摘した（Goodman 2010）。むろん、抑止側が攻撃をある行為者に帰属させることができたとしても、両者間の物理的・サイバー的能力の非対称性によって、抑止の脅威の信頼性が損なわれ、結果として報復の実行に至らないこともある。しかし、地政学的な状況に応じて責任を明確にさせることで、完全な帰属を求める必要はほとんどなく、結局のところ、弱者は強者を抑止できないという現実だけが残る。

「サイバー攻撃の帰属」（“Attributing Cyber Attacks”）という2015年に発表された論文の中で、リッドとブキャナン（Rid & Buchanan）は帰属を単純な技術的問題として扱うことを批判し、その認識は国家が作り出したものであると示した（Rid & Buchanan 2015）。著者らは帰属の流れを詳細に説明し、研究者の唯一の焦点となっている帰属の技術的なレベルとは別に、作戦レベルと戦略レベルの 2 つの不可欠な構成要素があることを指摘した。作戦レベルでは、事件を説明するための仮説を立てる技術的な段階からの手掛かりをもとに、技術的情報と非技術的情報を統合・分析する。前の段階での答えを集約し検討することにより、最終的に戦略レベルでの結論が導き出される。各段階間のコミュニケーションに気づくことができず、技術レベルだけを見ていては、帰属の試みの有用性を誤って判断してしまう。このように、サイバースペースにおける帰属は、もともと失敗する運命というわけではなく、技術情報をどのように集め、どのように処理し、トップレベルでどのように精査するかにかかっているのである。もちろん、それぞれの段階でできることに限界はあるが、それでもコミュニケーションが円滑になり、透明性が高まれば、帰属、ひいては抑止力を高めることができるとした。

Ⅴ．終わりに

サイバースペースにおける攻撃・あるいはサイバー戦争に準じる行動が問題になる前に、合理的抑止理論はいくつかの批判を受けながらも成り立っているというのが国際政治学界の通説だった。冷戦時代に勃興したこの理論は戦争抑止の諸条件を細かく網羅的、かつ論理的に示した。諸例外を除けば、条件付きとは言え、戦争発動者・国の暴走を未然に防ぐメカニズムとそのロジックを解明したと言ってもよいだろう。ただ、古典的な抑止理論にも落とし穴があると思われる。例えば、攻撃側と抑止側との関係性を固定化する傾向があり、国家の妥協や成長に示唆を与えない。また、現実の抑止における他の現象にも触

れていない。抑止側を自称する大国が、戦争抑止を名目に攻撃的な防衛政策を講じるといったケースが思い浮かばれる。更に、この理論は戦争に至る、もしくは戦争を寸前で回避するまでの過程に囚われすぎており、戦争が勃発した後のことをあまり気にかけない点も改めるべきだと考える。今回のウクライナ戦争からもわかるように、戦争が始まった後にも報復の脅しの応酬は続いた。主体によるこうした戦時中の行為も抑止理論をアップデートする示唆に富んでいるかもしれない。

インターネット時代になり、安全保障の概念がサイバースペースにまで拡大して解釈されるようになった。こうした新しい安全保障問題について合理的抑止理論を当てはめて説明し予測しようとする試みは非常に価値のあるものだと言える。上述した通り、サイバースペースをめぐる国際関係は当事国同士でさえ共通認識に欠けるほど未熟であるため、合理的抑止理論が与えてくれる識見は示唆に富んでおり、現実世界のサイバースペースにおける紛争の緩和につながる。更に、攻撃帰属や比例性など、このドメインにおいて合理的抑止理論が苦戦している問題が明らかになったことで、研究者・政策決定者共に焦点を当てるべき箇所が浮き彫りになり、解決に向けた方策を模索できるようになった。これらの洗い出された問題の現実での解決およびこれらの変数・要素を取り込んだ合理的抑止理論の刷新もしくは新しい理論の提唱が待たれる。

参考文献

Achen, C. & Snidal, D. 1989. “Rational Deterrence Theory and Comparative Case Studies.” *World Politics*, Vol. 41, No. 2: 143–169

Arquilla, J. & Ronfeldt, D. 1993. *Cyberwar is Coming!* Rand Corporation: Santa Monica, CA

Betts, R. 1998. “The New Threat of Mass Destruction.” *Foreign Affairs*, Vol. 77, No. 1: 26–41

Blank, S. 2001. “Can Information Warfare be Deterred?” *Defense Analysis*, Vol. 17, No. 2: 121−138

Brantly, A. 2018. “The Cyber Deterrence Problem.” *10th International Conference on Cyber Conflict*: 31–54

Bueno de Mesquita, B. 1980. “An Expected Utility Theory of International Conflict.” *The American Political Science Review*, Vol. 74, No. 4: 917−931

Clark, D. & Landau, S. 2010. “Untangling Attribution.” In *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy*, National Research Council: 25–40

Denning, D. & Strawser, B. 2017. “Active Cyber Defense: Applying Air Defense to the Cyber Domain.” in *Understanding Cyber Conflict: 14 Analogies*, (Eds.). Perkovich, G. & Levite, A. Washington, DC: Georgetown University Press

Elliott, D. 2011. “Deterring Strategic Cyberattack.” *IEEE Computer and Reliability Societies*: 36–40

Fearon, J. 1994. “Domestic Political Audiences and the Escalation of International Disputes.” *The American Political Science Review*, Vol. 88, No. 3: 577−592

———— 1997. “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs.” *The Journal of Conflict Resolution*, Vol. 41, No. 1: 68–90

Gallagher, M. 2022. “Biden’s ‘Integrated Deterrence’ Fails in Ukraine: The buzzy term is being used to justify cuts to conventional hard power that please progressives.” *The Wall Street Journal*, https://www.wsj.com/articles/biden-integrated-deterrence-fails-ukraine-russia-invasion-taiwan-xi-china-diplomacy-sanctions-hard-power-

defense-spending-budget-negotiations-11648569487

Galtung, J. 1984. “Transarmament: from Offensive to Defensive Defense.” *Journal of Peace Research*, Vol. 21, No. 2: 127–139

Goodman, W. 2010. “Cyber Deterrence: Tougher in Theory than in Practice?” *Strategic Studies Quarterly*, Vol. 4, No. 3: 102–135

Gupta, S. 2008. “The Doctrine of Pre-Emptive Strike: Application and Implications during the Administration of President George W. Bush.” *International Political Science Review*, Vol. 29, No. 2: 181–196

Harknett, R. 1996. “Information Warfare and Deterrence.” *The US Army War College Quarterly*, Vol. 26, No. 3: 93–107

Harknett, R. & Goldman, E. 2016. “The Search for Cyber Fundamentals.” *Journal of Information Warfare*, Vol. 15, No. 2: 81–88

Healey, J. 2018. “Not The Cyber Deterrence the United States Wants.” *Council on Foreign Relations*, https://www.cfr.org/blog/not-cyber-deterrence-united-states-wants

Huth, P. 1999. “Deterrence and International Conflict: Empirical Findings and Theoretical Debates.” *Annual Review of Political Science*, Vol. 2: 25–48

Huth, P. & Russett, B. 1984. “What Makes Deterrence Work? Cases from 1900 to 1980.” *World Politics*, Vol. 36, No. 4: 496–526

Iasiello, E. 2014. “Is Cyber Deterrence an Illusory Course of Action?” *Journal of Strategic Security*, Vol. 7, No. 1: 54–67

Jackson, V. 2022. “What is Integrated Deterrence? A Gap between US and Australian Strategic Thought.” *Australian Journal of Defense and Strategic Studies*, Vol. 4, No. 2: 263–274

Jasper, S. 2017. *Strategic Cyber Deterrence: The Active Cyber Defense Option*. Rowman & Littlefield: Lanham, Maryland

Jervis, R. 1978. “Cooperation Under the Security Dilemma.” *World Politics*, Vol. 30, No. 2: 167–214

———— 1979. “Review: Deterrence Theory Revisited.” *World Politics*, Vol. 31, No. 2: 289–324

———— 1982–1983. “Deterrence and Perception.” *International Security*, Vol. 7, No. 3: 3–30

———— 1989. “Rational Deterrence: Theory and Evidence.” *World Politics*, Vol. 41, No. 2: 183–207

Jervis, R., Lebow, R. & Stein, J. 1985. *Psychology and Deterrence*. The John Hopkins University Press: Baltimore, Maryland

Jiang, T. 2019. “From Offensive Dominance to Deterrence: China’s Evolving Strategic Thinking on Cyberwar.” *Chinese Journal of International Review*, Vol. 1, No. 2: 1–23

Kugler, R. 2009. “Deterrence of Cyber Attacks.” In Kramer, F., Starr, S. & Wentz, L. (Eds.). *Cyberpower and National Security*. Washington DC: National Defense University Press

Lebow, R. & Stein, J. 1989. “Rational Deterrence Theory: I Think, Therefore I Deter.” *World Politics*, Vol. 41, No. 2: 208–224

Leeds, B. & Anac, S. 2005. “Alliance Institutionalization and Alliance Performance.” *International Interactions*, Vol. 31, No. 3: 183–202

Levy, J. 1984. “The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis.” *International Studies Quarterly*, Vol. 28, No. 2: 219–238

Libicki, M. 2009. *Cyberdeterrence and Cyberwar*. RAND Corporation: Santa Monica, CA

———— 2012. *Crisis and Escalation in Cyberspace*. RAND Corporation: Santa Monica, CA

———— 2018. “Expectations of Cyber Deterrence.” *Strategic Studies Quarterly*, Vol. 12, No. 4: 44–57

Lieberman, E. 1995. “What Makes Deterrence Work?: Lessons from the Egyptian-Israeli Enduring Rivalry.” *Security Studies*, Vol. 4, No. 4: 851–910

Lindsay, J. 2015. “Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyberattack.” *Journal of Cybersecurity*, Vol. 1, No. 1: 53–67

Lindsay, J. & Gartzke, E. 2019. *Cross-Domain Deterrence: Strategy in an Era of Complexity*. Oxford University Press: New York

Lupovici, A. 2016. “The Attribution Problem and the Social Construction of Violence: Taking Cyber Deterrence Literature a Step Forward.” *International Studies Perspectives*, Vol. 17: 322−342

Mazanec, B. 2015. *The Evolution of Cyber War: International Norms for Emerging-Technology Weapons*. University of Nebraska Press

Mearsheimer, J. 1983. *Conventional Deterrence*. Cornell University Press: Ithaca, New York

Morgan, P. 2003. *Deterrence Now*. Cambridge University Press: Cambridge

Narang, V. 2013. “What Does It Take to Deter? Regional Power Nuclear Postures and International conflict.” *The Journal of Conflict Resolution*, Vol. 57, No. 3: 478−508

Powell, R. 1990. “Nuclear Deterrence and the Strategy of Limited Retaliation.” *The American Political Science Review*, Vol. 83, No. 2: 503−519

Quackenbush, S. 2011. “Deterrence Theory: Where do We Stand?” *Review of International Studies*, Vol. 37, No. 2: 741−762

Rid, T. & Buchanan, B. 2015. “Attributing Cyber Attacks.” *The Journal of Strategic Studies*, Vol. 38, No. 1−2: 4−37

Ryan, N. 2018. “Five Kinds of Cyber Deterrence.” *Philosophy and Technology*, Vol. 31: 331−338

Sagan, S. 2000. “The Commitment Trap: Why the United States Should Not Use Nuclear Threats to Deter Biological and Chemical Weapons Attacks.” *International Security*, Vol. 24, No. 4: 85−115

Sanger, D. 2018. *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age*. Crown Publishing Group: New York

Schelling, T. 1966. *Arms and Influence*. Yale University Press: London

Schulze, M. 2019. “Cyber Deterrence is Overrated: Analysis of the Deterrent Potential of the New US Cyber Doctrine and Lessons from Germany’s Active Cyber Defense.” *German Institute for International and Security Affairs*, Vol. 34

Singer, P. & Friedman, A. 2014. *Cybersecurity and Cyberwar: What Everyone Needs to Know*. Oxford University Press: New York

Signorino, C. & Tarar, A. 2006. “A Unified Theory and Test of Extended Immediate Deterrence.” *American Journal of Political Science*, Vol. 50, No. 3: 586−605

Siroli, G. 2006. “Strategic Information Warfare: An Introduction.” In Halpin, E., Trevorrow, P., Webb, D. & Wright, S. (Eds.). 2006. *Cyberwar, Netwar and the Revolution in Military Affairs*. Palgrave Mcmillan: New York

Stein, J. 1996. “Deterrence and Learning in an Enduring Rivalry.” *Security Studies*, Vol. 6: 104−152

Snyder, G. 1961. *Deterrence by Denial and Punishment*. Woodrow Wilson School of Public and International Affairs, Center of International Studies, Princeton University

Steinbruner, J. 1976. “Beyond Rational Deterrence: The Struggle for New Conceptions.” *World Politics*, Vol. 28, No. 2: 223−245

Taddeo, M. 2018. “The Limits of Deterrence Theory in cyberspace.” *Philosophy and Technology*, Vol. 31: 339−355

Todd, G. 2009. “Armed Attack in Cyberspace: Deterring Asymmetric Warfare with an Asymmetric Definition.” *Air Force Law Review*, Vol. 64: 65−102

Tor, U. 2015. “Cumulative Deterrence as a New Paradigm for Cyber Deterrence.” *Journal of Strategic Studies,* Vol. 40, No. 1: 92−117

U. S. Department of Defense. 2022. *2022 National Defense Strategy of The United States of America: Including the 2022 Nuclear Posture Review and the 2022 Missile Defense Review*. https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.PDF

Weede, E. 1983. “Extended Deterrence by Superpower Alliance.” *The Journal of Conflict Resolution*, Vol. 27, No. 2: 231−254

Wendt, A. 1992. “Anarchy is what States Make of it: The Social Construction of Power Politics.” *International Organization*, Vol. 46, No. 2: 391−425

外務省 2011年「サイバー空間に関するロンドン会議について」https://www.mofa.go.jp/mofaj/annai/honsho/fuku/yamane/cyber_1111.html

千々和泰明 2022年『戦後日本の安全保障―日米同盟、憲法 9 条から NSC まで―』中公新書
土屋大洋 2020年『サイバーグレートゲーム―政治・経済・技術とデータをめぐる地政学―』千倉書房
内閣官房 2022年「国家安全保障戦略」（令和 4 年12月16日　国家安全保障会議・閣議決定）https://www.cas.go.jp/jp/siryou/221216anzenhoshou.html
平和・安全保障研究所 2022年「日本の新しい「国家安全保障戦略」等について―分析と評価―」https://www.rips.or.jp/research/5904/

蔡翠紅 2019年「中美网络空间战略比较: 目标、手段与模式」『当代世界与社会主义』2019年第 1 期
孟威 2014年「网络安全: 国家戦略与国際治理」『当代世界』2014年第 2 期
黄志雄 2015年「国际法视角下的“网络战”及中国的対策――以诉诸武力权为中心」『现代法学』2015年第 5 期