Technology in Society 59 (2019) 101177

Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc

Mutual assured destruction in information, influence and cyber warfare: Comparing, contrasting and combining relevant scenarios

Jeremy Straub

Institute for Cybersecurity Education and Research, North Dakota State University, 1320 Albrecht Blvd., Room 258, Fargo, ND, 58102, USA

ARTICLE INFO

Keywords:
Mutual assured destruction (MAD)
Cyber warfare
Information warfare
Influence warfare
Influence operations
Multi-modal warfare

ABSTRACT

Mutual assured destruction is a key deterrent against the use of the most powerful weapons. The threat of it successfully prevented the deployment of a nuclear weapon during and since the United States versus Soviet Union Cold War. It has also prevented the escalation to total warfare scenarios (where countries fully deploy their arsenals and capabilities against each other). Cyber weapons are poised to potentially create more havoc, death and destruction than a single nuclear weapon would and there has been significant contemporary use of information and influence warfare. Given the foregoing, this paper investigates whether mutual assured destruction scenarios may exist which are (or could be) responsible for keeping the use of these warfare methods in check. Further, the paper considers whether the three types of warfare might be effective in holding the others in check.

1. Introduction

The advent of nuclear weapons fundamentally changed warfare. During the United States and Soviet Union Cold War, both sides developed enough nuclear weapons to destroy each other multiple times over [1]. Each side perceived the other to be a “sensible rational opponent” whose behavior was shaped by “threats of nuclear retaliation” from the other [2]. Each relied upon the other to be concerned about its own survival and to not take an action that would lead to its own annihilation by nuclear retribution. While some secondary [3] and proxy conflicts [4] occurred, neither side could risk deploying a nuclear weapon because of the anticipated response.

The “strategic bi-polarity” model that defined the Cold War no longer represents the state of the world, in terms of physical conflict [2]. This was never an applicable model for cyber, information or influence warfare. Instead, the current status of physical world conflict is a state of “strategic multi-polarity” [2] and this same model, albeit with different players and means of warfighting, is representative of cyber, information and influence warfare. Under a the model of strategic multi-polarity, Curtis [2] contends, mutual assured destruction isn't effective. For this deterrent approach to work, each state would require the capability to assure destruction to all other states and combinations of states that might attack it. Given that not all states have nuclear capabilities, this standard would clearly not be met.

However, Curtis's conclusion is a bit extreme. As the United States and Soviets strongly influenced or controlled allies and others' actions during and subsequent to the Cold War, similar influence could be exerted in any strategic multi-polarity situation. As long as a state is willing to trust that its retaliation interests are ensured (either by its own capabilities or that of an ally), it does not need to have the capabilities to fully destroy all possible enemies. Instead of, or in addition to, a system of alliances, a ‘policeman’ (as the United States has been called [5]) might emerge that guarantees retaliatorily response against those that don't follow agreed upon rules. Under this scenario, the ‘policeman’ power removes the need for other states to have weapons, if they do not otherwise seek to, for the purposes of assured destruction self-protection.

With cyber and information warfare emerging as new mediums of conflict and enabling different approaches to influence warfare, answering the question of how to best ensure stability in the modern world, with particular focus on these types of warfare, is critical. Additional complexity is created by the fact that these forms of warfare can scale from having localized and even minimal effects to potentially causing the mass casualties of a nuclear weapon (for example, by targeting a nuclear weapon or power plant or cutting off critical infrastructure elements at key times). There may be no clear demarcation between peace and war in cyber activities, with different states having different definitions of what constitutes an act of war or a wartime activity – if they have such policies at all. Additionally, the fact that a single armament can have such widely different impacts (and may have unanticipated impact) may result in escalation to nightmare scenarios with widespread impact.

E-mail address: jeremy.straub@ndsu.edu.

https://doi.org/10.1016/j.techsoc.2019.101177
Received 31 January 2019; Received in revised form 16 July 2019; Accepted 5 August 2019
Available online 06 August 2019

This paper considers the effectiveness of a mutual assured destruction paradigm, as well as competing approaches, to the areas of cyber, information and influence warfare. It draws conclusions regarding the efficacy of these models and presents a decision-making framework for policy maker and analyst use.

2. Background

This section discusses the topics of information warfare, influence warfare and cyber warfare. A working definition for each topic is provided and relevant prior work in each area is discussed.

2.1. Information warfare

Denning [6] characterizes information warfare as "operations that target or exploit information media in order to win some objective over an adversary." This definition is enumerated to include cyber and electronic warfare, intelligence operations, hacking, sabotage, managing perception, espionage and telecommunications attacks. As this definition would encompass most if not all of influence warfare and all of cyber warfare, for the purposes of this section and this paper, more generally, refences to information warfare will be referring to information warfare techniques that do not fit within the definitions for influence or cyber warfare.

Information warfare typically serves an influence goal, but it is not the only way of conducting influence warfare. Examples of information warfare include its use by the United States against Islamic fundamentalists [7], by Al-Qaeda [8] versus the United States [9] as part of the Afghanistan [10] conflict by Patani separatists [11] as well as between Kenya and Somalia [12]. Social media [11,12] is frequently the medium of information warfare (as was demonstrated by Russian activities related to the 2016 United States presidential election [13]), but it is not the only means. Information warfare is not a new concept and is not tied to the Internet. The United States used radio stations, such as Radio Free Europe and Voice of America as part of its strategy to win the Cold War [7]. The Soviet Union also used information warfare techniques through its media. During the Vietnam war, information warfare was targeted domestically [7] to try to increase support for the war. Leaflets were used in Iraq to try to improve the perception of foreign troops and disrupt Al-Qaeda cells [7].

Information operations techniques can range from simply providing information that supports or attacks a position to more intricately planned tactics that try to target a certain side of the brain (which has been shown to increase the effectiveness of some marketing [14]), place information in long term memory and influence targets' motivations and cravings [14]. They can make use of a knowledge of psychological and marketing principles to increase their persuasiveness. For example techniques such as utilizing contrast and expecting reciprocity as well as making messages personally relevant can increase their persuasive power [14]. Social media, in particular, can facilitate the delivery of highly personalized content, while mass media can be used to generate more broadly targeted messages.

2.2. Influence warfare

Influence warfare is a broad category of techniques and tactics designed to manipulate individuals and the groups, companies and nations that they are members of. Influence can be exerted using information or cyber warfare. Conventional warfare may also be used to have influence effects. Maria [15] contends that a definition for influence warfare must be drawn from common and doctrinal use as well as the consideration of dictionary definitions. From this analysis, influence operations are defined to include both actions to persuade or dissuade with the use (or the threat of the use) of force as well as through "marketing and advertising" techniques.

In the *Art of War* Sun Tzu explains that "all warfare is based on deception" and suggests that warfighters "hold out baits to entice the enemy" and that "supreme excellence consists in breaking the enemy's resistance without fighting" [16]. Clearly, from this, influence warfare significantly pre-existed cyber warfare and while many influence strategies include an information operations component, this is not required and may be only a small part of an influence campaign.

One could argue that every armed conflict is influence warfare, given that a typical goal is to cause the adversary to capitulate (which, of course, is an influence objective). However, the focus herein will be on influence warfare where warfighters actively manage influence campaigns, use non-conventional tools and seek to achieve influence objectives (beyond, in addition to and/or in support of the general goal of adversary capitulation). Maria [15] states that, from a United States perspective, an influence operation is defined as "a deliberately planned and synchronized series of actions designed to produce desired behaviors within adversaries and affected populations." This can be achieved "through the direct or indirect, threat or actual use of all U.S. military power and capabilities in order to achieve a relative advantage or desired end state." More simply, an influence operation involves the use (or the threat of the use) of power or manipulation to shape a decision or behavior, as desired.

Influence techniques are used by state actors and non-state actors alike. Of course, commercial marketing aims to produce influence effects, which are typically mild by comparison to influence and information warfare ones. Religious institutions have demonstrated influence technique use with examples including crusades [17] and Fatwas (including cyber-specific Fatwas [18]). Information warfare techniques have been demonstrated by non-state, non-religion actors, such as Al-Qaeda [19], individuals [20] and numerous state actors. Russia, for example, demonstrated the use of influence techniques in their 2014 Crimea annexation and 2016 presidential election activities [21]. Influence operations, in Russia, have lineage back to the time of the tsars and have been used for both external and domestic goals. Many modern Russian attacks target the United States [22] and the United States has demonstrated some capabilities to mitigate, detect and respond to these attacks. Samet [22], however, contends that these techniques are not sufficient and must be updated to use the most recent technologies and to respond to the most recent Russian techniques. Influence techniques used by groups including the Vietcong, Shining Path, Hezbollah and the Jemaah Islamiyah have also been documented [23].

Influence techniques, fundamentally, utilize interactions to attempt to support or oppose an action or activity. In Ref. [24], six types of influence transactions were identified: information use support, information use oppose, information withhold support, information withhold oppose, relationship use support and relationship use oppose. Additional transaction types related to influencing through action and other means must also be considered. Multiple categories of influence relationships that can be exploited (or which must be protected) were also identified [24].

2.3. Cyber warfare

Andress and Winterfeld [25] discuss the evolving nature and difficulties related to arriving at a full definition for cyber warfare. They identify multiple mediums that could be part of cyber warfare including "cybersecurity, computer network operations, electronic warfare or anything to do with the network." They also discuss how even the term warfare is a subject requiring some interpretation. Shoaib [26] builds on this, providing a more thorough definition and conceptual framework for cybersecurity and cyber warfare. Fundamentally, though, a working definition for cyber warfare is not particularly elusive. For the purposes of this section and paper, references to cyber warfare will relate to actions to attack and defend using electronic mediums, actions to attack and defend electronic mediums themselves and non-electronic activities related to the foregoing.

Cyber warfare includes techniques that can be used to obtain information that can be used for information and influence warfare, as well as techniques that make use of information for influence purposes. Additionally, cyber operations can target adversary systems, infrastructure and those reliant on systems for the purposes of influencing by threatening to attack or actually attacking. Cyber warfare can also be used to create impact that does not have a specific influence goal. Whatever the goal, these techniques are conducted in the global commons of cyberspace [27]. Like the sea, air and outer space, cyberspace is a non-governed area and operations in it are conducted by both state and non-state actors who draw upon their history of operating in other non-governed space to formulate strategies for cyberspace operations [27].

Kosenkov [28] contends that cyber warfare is a "new global threat," noting that complicated, targeted and covert activities could easily be conducted, against nations, if others' cyber power is left unchecked. Cyber war can happen at a much more rapid pace than conventional conflict, with the potential that a conflict may be decided "in mere 'nanoseconds'" and the threat is most severe to the "technological civilizations of the west" as opposed to developing nations [29]. Cyberattacks can have profound economic consequences [30] and may result in the disablement of military, government and private sector systems, representing a vulnerability for nations that have little worry about vulnerability to conventional warfare techniques [29].

Because it is a "less-bloody" use of force technique, Arquilla [31] contends that it is "tempting" to use cyber warfare as a first attack type. This, though, is problematic, he contends, as it could lead to a "virtual arms race." Additionally problematic, to established powers, is that their conventional strength does not guarantee supremacy or even leadership in cyber warfare [32]. There is a significant potential that a small state or non-state actor could design and deliver an attack that could have dramatic impact on larger and more established players. Deterrence is needed to prevent a so-called "cyber 'Perl Harbor'" [32].

3. Mutual assured destruction and its deterrent benefit

The mutual assured destruction concept [33] was born out of the Cold War buildup of nuclear weapons, though it would later find application to other unrelated areas including fishery operations [34], geoengineering [35] and campaign finance reform [36]. During the Cold War, both sides to the conflict developed a sufficient number of weapons to more than decimate the other side [1]. Systems and processes were developed to detect launches by and inbound missiles from the other to allow missiles to be launched in response, before the response capability could be taken out by an initial strike. Because of this, there was no advantage of launching first, as the opponent's missiles would be launched in retaliation and, though they might arrive minutes after the initial strike, they would equally decimate the first-mover.

In this conflict, each side perceived the other to be a "sensible rational opponent" deterred by the "threats of nuclear retaliation" from the other party [2]. Because neither party could achieve a decisive victory (or victory at all), neither party had incentive to act first or at all. This was premised on both parties being equally concerned about their own survival and unwilling to take an action that could (and likely would) bring about their own annihilation. This doctrine was so fundamental that it even caused officers to question indications that the other party had launched [37], as no reason for doing so, without nuclear launch provocation, could be fathomed. Thus, the enormity of launching a nuclear attack – and the potential for even the use of a single missile to turn into an all-out, potentially species ending war – prevented any nuclear weapon use at all, though secondary [3] and proxy conflicts [4], which didn't run the risk of nuclear escalation, still occurred.

Some however, don't see mutual assured destruction as the only possible outcome of a nuclear exchange and suggested, instead, that thought must be given to post-nuclear exchange warfighting [38]. These individuals proposed counter-force targets for nuclear weapons, leaving the major cities intact initially, but still threatened by a second wave of nuclear deployment. They argue that a "war fighting posture" not only serves as a better deterrent for the United States and its allies, it also is more resilient to adversary technological advances and better positions the United States for a scenario where deterrence fails [38]. Others suggest the use of flexible response doctrine where nuclear weapons can be deployed in a non-assured destruction scenario to achieve limited aims [39]. Yet others argue that winning a war in the nuclear era is not possible and thus the only value of nuclear weapons is countering nuclear weapon use [40]. Art [40] argues that many neglect the value of nuclear weapons in limiting other forms of warfare: given that a war that starts with conventional warfare aggression could potentially escalate to become a nuclear conflict, those opposing nuclear powers are more restrained in their actions for fear of nuclear escalation.

Clear superiority as well as faith in anti-missile systems potentially change the equation, allowing a decision maker to contemplate nuclear use and the potential to actually win [40]. Similarly, complete nuclear disarmament removes the deterrent value, to smaller conventional conflicts, of the potential of nuclear escalation.

An increase in nuclear powers, however, changes many of these scenarios. Weapon proliferation could result in scenarios where one power is restrained from meeting a commitment to an ally or responding to an attack against it from one party, due to a nuclear threat presented by another party [41]. Curtis [2] argues that mutual assured defense becomes problematic in a "strategic multi-polarity" situation (with multiple nuclear powers) for just this reason.

From the foregoing, the following can be concluded and applied to information, influence and cyber warfare: First, that mutual assured destruction has a deterrent effect on the use of both assured destruction (AD) weapons as well as on non-assured destruction (NAD) weapons. Second, that this deterrence value requires an unimpeded and non-impedible way of delivering the AD weapon response. Third, if there are multiple players, the AD weapon must be deployable directly against each adversary without restraint from alliances or other interference, in order to be an effective deterrent. Finally, it suggests that NAD weapons are required, in addition to AD weapons, for use to prevent the need for AD escalation, against non-AD holding adversaries and in addition to or after an AD attack.

4. Mutual assured destruction for information, influence and cyber warfare

The following three sections focus on identifying and evaluating mutual assured destruction (MAD) scenarios for information, influence and cyber warfare. While an obvious approach to this challenge would be to identify how each could be used to launch, trigger or in other way invoke nuclear weapons to create a MAD scenario, this has been largely avoided and is only a small part of the discussion. Instead, other AD armaments are discussed as are AD scenarios triggered by the use of what would typically be considered to be NAD armaments. Also considered, for each scenario, is the use of NAD approaches that may have a deterrent effect on AD use and scenarios.

4.1. Information warfare mutual assured destruction

Not much has been written, previously, regarding MAD scenarios related to information warfare. In section 2.1, information warfare was characterized as "operations that target or exploit information media in order to win some objective over an adversary" [6]. This section discusses information warfare AD methods, counter-AD methods and non-AD techniques that can be used to oppose information warfare AD techniques.

4.1.1. Assured destruction methods

Information warfare assured destruction techniques require a vector (such as a secondary attack) to create the requisite destruction for an AD scenario. Information warfare could be used to collect information to facilitate a cyber-attack that could impact infrastructure (for example, a nuclear reactor or dam) or trigger the launch of a weapon. Information warfare could also be utilized for an influence attack to persuade an individual to take an action that causes the destruction. This could be a direct action, such as an act that opens a dam, unleashing water on an area or causing a reactor to meltdown. Alternately, it could be an indirect action that, either due to an unexpected condition, occurrence or an additional concurrent action unknown to the person being influenced, causes the destruction.

4.1.2. Counter assured destruction methods

The attacks described for direct AD methods are applicable to counter AD as well. The key aspect of using information warfare as part of a counter AD scenario is that for whatever technique or techniques were selected it is necessary to ensure that the elements required for AD response are still available after the triggering event (i.e., adversary action) has occurred. Considering that there may be multiple types of adversary attacks that could trigger the information warfare AD response, this may require significant redundancy, geographical and information warfare technique diversity.

4.1.3. Non-assured destruction methods for countering assured destruction methods

Information warfare techniques can potentially be very effective for NAD activities that can be used to counter an AD action. Information warfare can be used as part of an influence or cyber campaign, to this end. In the influence warfare campaign, simply drawing international public attention to the prospective pending AD use by the adversary may be sufficient to deter it. Alternately, an influence operation could be used to persuade an individual who is critical to the AD action chain to not take action, take an incorrect action or delay action, thereby preventing the AD action. Information warfare blackmail techniques could also be utilized if the target of the pending AD action has useful information concerning the AD-planning adversary or its allies or leadership. The prospective AD target could make multiple copies (to prevent destruction of the information as part of the AD campaign) and threaten to release the documents or other information if the AD action is taken. With many of these, like with the previously described counter-AD techniques, it is critical to develop plans to allow the key individuals and information required for NAD retaliation after an AD attack to be available (i.e., survive the AD attack) to perform these functions.

4.2. Influence warfare mutual assured destruction

Similar to information warfare, there has not been significant prior literature on the influence warfare MAD topic. In section 2.2, influence warfare was defined as including actions that are designed to persuade and dissuade with the use (or the threat) of force as well as operations making use of “marketing and advertising” techniques [15]. Information warfare and cyber warfare were also discussed to have overlap with influence warfare. While the MAD concept is, arguably, an influence operation in and of itself, other types of influence have not yet been significantly considered. As will be discussed in the following section, cyber warfare influence operations have been discussed in the literature and many cyber operations are a type (or at least contain a component) of influence warfare. This section discusses influence warfare AD methods, counter-AD methods and non-AD techniques that can be used to oppose influence warfare AD techniques.

4.2.1. Assured destruction methods

Like with information warfare attacks, influence attacks require a vector to cause the required destruction. This will typically be in the form of an operation that convinces an individual or entity to either act against adversary interests, due to persuasion, or an operation that is designed to convince the individual or organization that it is in their interest to perform the required task.

To achieve these ends, persuasion can take the form of financial or other compensation (i.e., a bribe), a threat of injury to the individual, entity or an individual or entity that the target is concerned about, or the threat (or incentive) of information disclosure. A more complex influence campaign could also be designed with a goal to win over the heart and mind of the target or targets to get them to voluntarily perform the required actions. Alternately, a campaign could attempt to confuse the targets to convince them to perform (or not perform) an action without necessitating a change in their beliefs.

4.2.2. Counter assured destruction methods

As with information warfare, the attacks described for direct AD methods are applicable to counter AD as well. As with information warfare, for influence warfare to be an effective part of a counter AD scenario, it is required that whatever technique or techniques were selected be assured to have its/their required elements for the AD response still available after the triggering event (i.e., adversary action) has occurred. Considering that there may be multiple types of adversary attacks that could trigger the influence warfare AD response, this may necessitate significant redundancy, geographical and influence warfare technique diversity.

4.2.3. Non-assured destruction methods for countering assured destruction methods

Like with the approaches that can be used to coerce, convince or confuse individuals into performing AD and counter AD attacks, similar methods can be used to coerce, convince or confuse these individuals to not perform an AD or counter AD attack. For example, key individuals in an AD attack chain could be identified and targeted for coercion. Alternately, entities that are critical to AD attack operations could be similarly targeted for coercive activities.

4.3. Cyber warfare mutual assured destruction

The MAD concept has been applied to cyber warfare in several previous studies. In section 2.3, cyber warfare was described as warfare involving “cybersecurity, computer network operations, electronic warfare or anything to do with the network” [25]. It was defined including actions that attack and protect electronic mediums, as well as attacks and defenses using these mediums. Non-electronic activities related to the foregoing are also inherently included.

Morgan [42], Philbin [43], Nye [44] and Bendiek and Metzger [45] propose the adaptation of nuclear era deterrence approaches, based on MAD, to the cyber realm. Lonsdale [46] proposes, in particular, the use of the warfighting approach where (in nuclear deterrence) nuclear weapons were not seen as a complete deterrent solution, but rather as a part of a broader strategy designed to ensure deterrence and post-deterrence-failure capabilities. Crosston [47], alternately, proposes the concept of “mutually assured debilitation,” recognizing that cyber attacks may not destroy (in the immediate way a nuclear detonation would) but can be catastrophically debilitating for cities, nations and their economies. Ridout [48] proposes a more nuanced strategy adding defense and resilience concepts to the AD-based deterrence concept.

Others also have studied and advanced the concept. Chukwudi, Udoka and Charles [49] consider the implications of game theory to deterrence. Davis [50] considers the question of escalation and escalation ladders in the cyber domain. Geers [51] suggests that deterrence may be “an impossible task” due to issues of asymmetry and needing to determine attack attribution, while Gale [52] and Mokarram [53] discuss the use of MAD and deterrence in United States and European strategy, respectively. Huston [54] evaluates factors that may drive cyber warfare towards civilian impact, and those that may produce restraint.

This section discusses cyber warfare AD methods and counter-AD methods. It also covers non-AD techniques that can be used to oppose cyber warfare AD techniques.

4.3.1. Assured destruction methods

Cyber warfare can be used to implement several different AD methods. Cyber operations can serve as a medium for information and influence operations, as discussed in previous sections. An individual could be contacted, coerced or convinced over electronic channels to take an action that causes significant destruction. This could be through targeted contact or the implementation of a cyber-medium delivered threat or reward targeting the individual.

Cyber operations can also be utilized more directly. The could be used to compromise and electronically command a nuclear weapon or other system (such as opening a dam) that can cause significant destruction directly. Attacking or degrading electronic systems can also be used to cause immediate or long-time-scale damage by preventing communications or other processes required to sustain life (ranging from medical systems to systems required to maintain and assure the food supply).

4.3.2. Counter assured destruction methods

As has been previously discussed for information and influence warfare, the same types of attacks that are used for AD can be used for counter-AD as well. Counter-AD may also benefit from threats of longer time-scale attacks that may be used as a deterrent to attacks that may be more immediate. The ability to threaten longer-term impact may be particularly important to entities that don't have the capability to make a single large impact attack or which may lack the autonomous command capabilities needed to respond to a cyber-attack before local infrastructure is damaged or disabled, impairing the retaliation.

Previously [55], an approach was proposed based on the use of non-recallable autonomous software which would be spread (and further spread itself) onto multiple systems. Specifically, this would include systems not owned or controlled by the software's creator and operator. This would allow the system to line in wait and be activated (or self-activate) in retaliation to an AD event. Like with nuclear counter-AD, it is critical for cyber counter-AD to be able to be triggered before systems are disabled or significantly impacted by the adversary and be able to operate without the key targets of the adversary's AD attack. Alternately, systems would need to be designed to withstand likely AD attacks to guarantee retaliation. Distributing the response software onto multiple systems so that it is not destroyed or significantly degraded in the initial attack fits into this second category.

4.3.3. Non-assured destruction methods for countering assured destruction methods

Like with information and influence warfare, cyber operations can be used to counter AD methods with non-AD attacks. These can include attacks designed to blackmail or influence individuals, organizations or governments through the threatened release or withholding of information or the use of include, which are deployed over a cyber channel. Prospective, threatened or demonstrated attacks against non-AD critical infrastructure, commercial interests and other targets may also serve counter-AD goals.

5. A theory of and framework for cross domain deterrence

With only nuclear deterrence needing to be considered and two responsible actor adversaries, the MAD equation is relatively simple. Each side needed to have enough weapons to ensure that the other had to fear total or near total destruction. The weapons needed to be able to be launched before the initial strike impacted or they needed to be able to survive the initial strike. Alternately, a side might design their arsenal to have enough weapons to sustain initial losses and still be sufficient for an AD response.

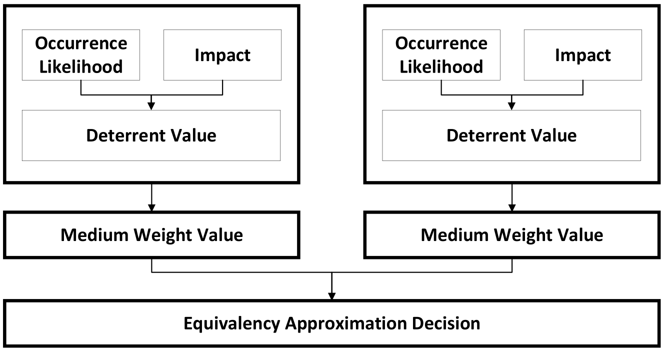

Fig. 1. Cross-domain (medium) deterrence.

However, once there are multiple adversaries [2,41] and/or multiple types of AD, significantly more thought needs to be given to what is required to deter a nation's adversaries. Considerations may include alliances (and, thus, combined attack/retaliation power), mismatched AD and counter-AD technologies as well as the actual damage that each technology can inflict and the fear and perception associated with each technology.

Fig. 1 depicts the most basic form of this consideration. It deals with two different AD/counter-AD mediums or technologies. The most basic form of the equation makes the deterrent value equal to the product of the likelihood of the successful deployment of the technology and the impact it would have if it was used. Each AD/counter-AD medium/technology has a weight value applied. This allows the model to consider perception, fear, longer term impact and other factors that are not considered in the basic calculation.

Each decision maker must then determine in their own organizational, national or personal planning as to how to compare different AD/counter-AD scenarios. The equivalency approximation decision embodies this process.

It is also important to note that for true MAD, the two sides must be at least roughly equivalent. If not, one side may see a scenario where they incur significant destruction, but their side still comes out ahead (while their opponent suffers full AD and ends up powerless). This might not result in mutual deterrence, but rather in a sense that there may be something to be gained from the conflict, actually encouraging it to occur.

In this most basic (Fig. 1) version, there are two AD/counter-AD technologies potentially being deployed in different domains. Of course, the two could also be in the same domain (with the adversaries using the same or very similar technologies for AD/counter-AD), with minimal comparative weighting (if any) being applied between them.

A more complex scenario must also be considered and is presented in Fig. 2. This figure depicts a scenario where each adversary is using only one medium (and thus requires only a single medium weight), but there are multiple elements of AD/counter-AD within each medium. This is handled relatively similarly to how the basic equation was evaluated. In this case, the different elements' deterrent values are added together and the weighting is applied. Again, parties must make a critical decision with regards to the equivalency calculation. If is not roughly equivalent, one party may have incentive to commence conflict, believing that they will fare better, even despite overwhelming losses.

6. A theory of and framework for Multi-Domain deterrence and assured destruction

Building on the foregoing, another scenario that must be considered is where adversaries have defensive and offensive capabilities that span

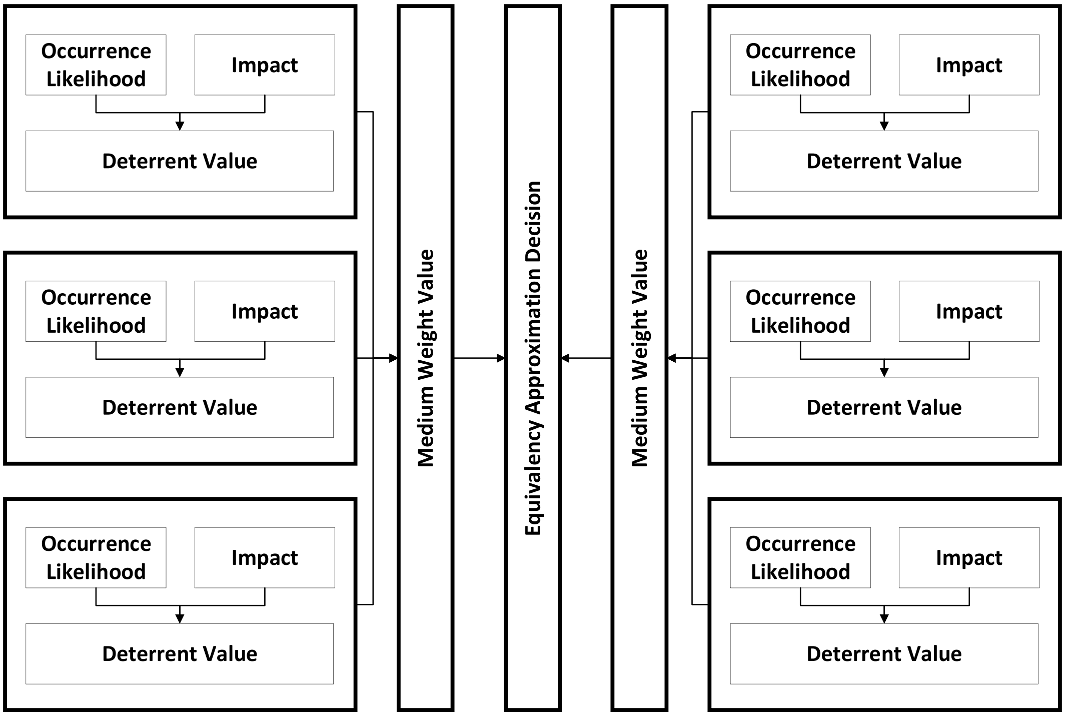

Fig. 2. Multi-factor cross-domain (medium) deterrence.

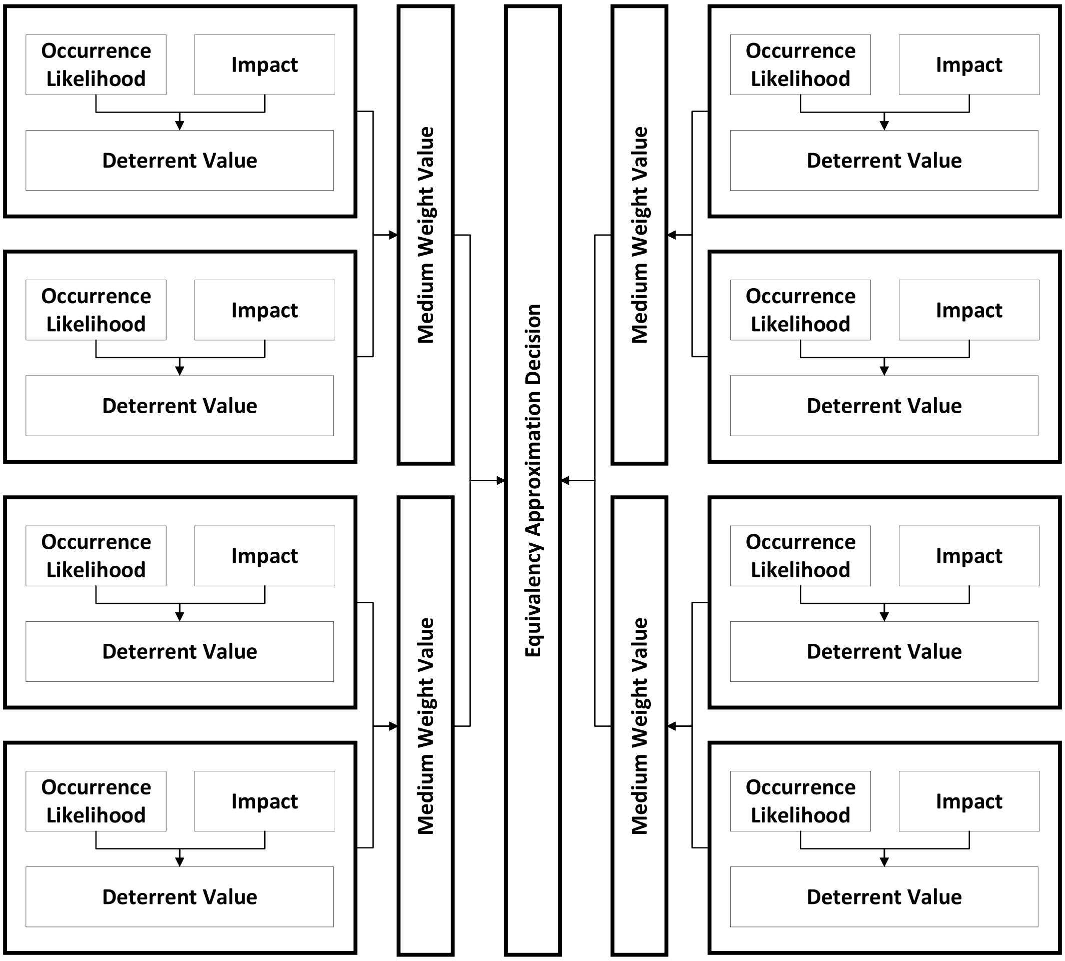

Fig. 3. Multi-domain (medium), multi-factor deterrence.

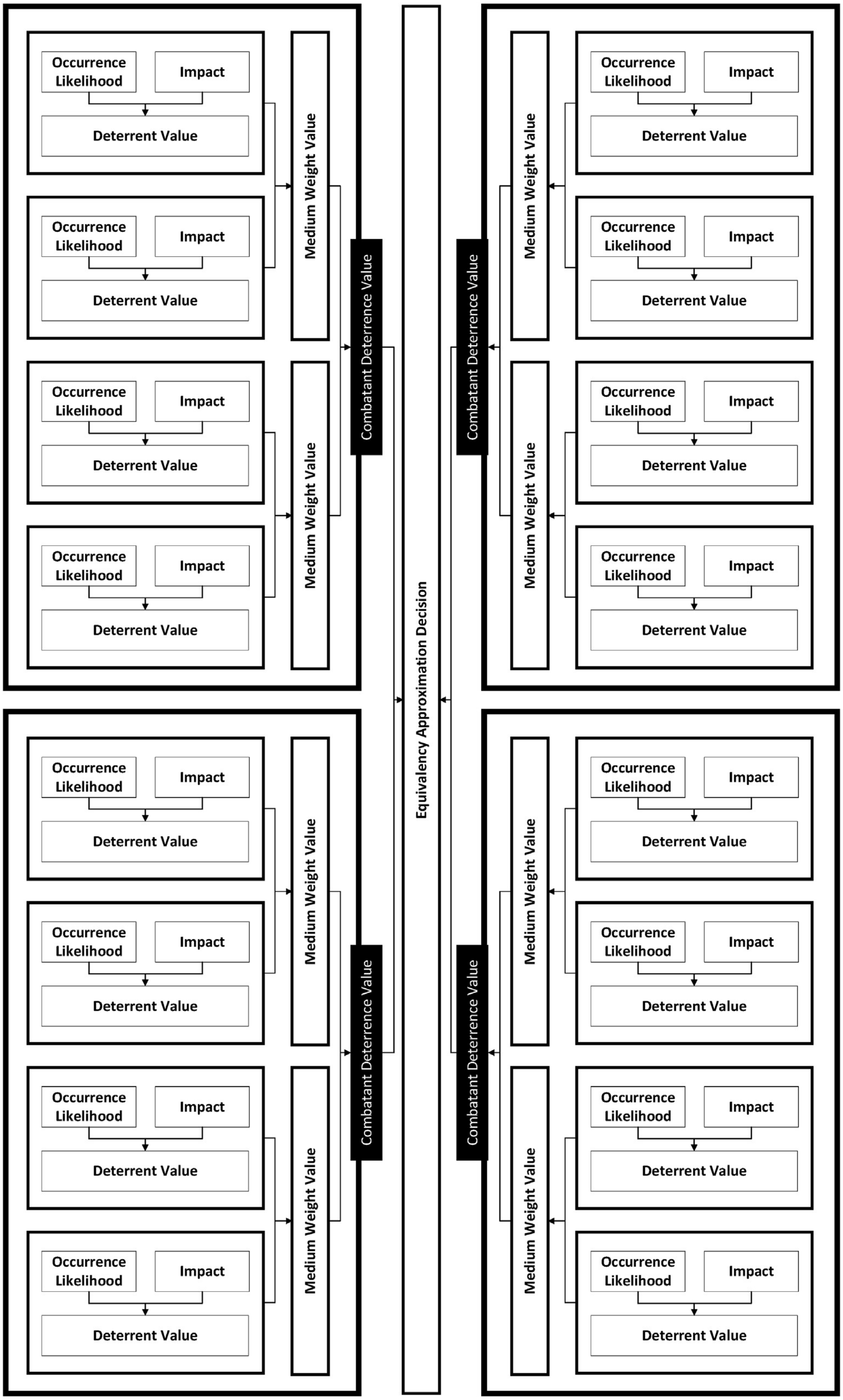

Fig. 4. Multi-combatant, multi-domain (medium), multi-factor deterrence.

multiple domains or mediums. This is, of course, a likely scenario in any but the most basic of real-world environments. Fig. 3 depicts a model for a multi medium, multiple AD/counter-AD technology per medium scenario.

While this model can be used for almost any scenario involving only two adversaries, it is not always necessary to consider the full complexity. In some cases, certain domains and mediums can be ignored, as they are insignificant when compared to the capabilities of others.

However, if one domain or medium doesn't dominate others (with orders of magnitude more deterrent value), then all will likely need to be considered. This is particularly true when one adversary's lesser capabilities are positioned close to the other side or are distributed, meaning that they wouldn't be destroyed in an initial attack. The Zanryū Nipponhei ('Japanese holdout') scenario [56,57] is one example of how an inferior fighting force (of even a single individual, in this case) that survives the use of an armament of much greater deterrent value can cause significant ongoing damage over time. The proposed non-recallable autonomous vehicles and software [55] necessitate similar consideration.

Perhaps the most complex scenario is one involving multiple combatants, whether allied into two or more groups or acting individually. Curtis [2] notes that this scenario presents significant challenges. Like with the basic MAD scenario, you need to have reasonable balance between opposing forces. If you do not, then one may feel that it has an advantage and choose to strike, despite the potential for high losses. They may even feel that their greater capabilities (and the fact that they will retain capabilities after the counter-strike) may dissuade their adversary from attacking at all. This balance must be considered between all permutations of opposing forces, which is what makes this scenario very complex. Not only does the current system of alliances need to be considered, potential short-term alliances and alliance changes must also be considered and should result in a similarly stable configuration (of evenly matched AD capabilities). This is obviously quite problematic.

An alternate approach is to have a 'policeman' power that holds greater capabilities than others (and at least as great as any possible combination of other powers). If this policeman power is trusted and agrees to retaliate against any other power that starts an illegitimate conflict, this can also be a stable situation. However, it relies on the policeman power to not start a conflict, despite potentially being in a position where it can win decisively against at least some other powers. In addition to a single policeman configuration, dual (or multiple) policeman configurations are possible; however, these add much greater complexity and may deteriorate into principal powers with client states, instead of a policeman configuration.

Irrespective of the exact configuration, Fig. 4 presents a model for representing the MAD evaluation. In the figure, an example including four powers is depicted. Each has multiple mediums of AD capability and multiple armaments within each medium. In addition to what is shown, a dotted line could be added, encircling two or more of the powers, to depict alliances and an alliance combined deterrent value could be calculated for the alliance members, using the same approach as it is for individual state actors.

7. Evaluation of models and their efficacy

The models presented provide a convenient way to depict and a perspective from which to approach MAD scenario evaluation and the comparison of AD capabilities between individual adversaries in a single medium and across multiple domains/mediums. They also support scenarios where there are two strong alliances (which can each be treated as a single actor) and scenarios where there are more than two adversaries (which cannot be consolidated down to only two sides), including scenarios with weak and changing alliances. Like any model, though, their greatest weakness is in their reliance on the correct population of information and planners and decision makers, who use the models, having access to all relevant information. Internal controls (such as security clearance levels and 'off book' programs) may obfuscate friendly capabilities. Allies may, similarly, fail to fully disclose their capabilities. Alternately, 'fog of war' issues may result in significant over or under estimates of adversary and adversary alliance capabilities. Adversary capability disclosure, pursuant to treaty or other obligations and facilitate facility inspections may be similarly suspect and subject to manipulation.

Given that adversaries and allies alike may have reason to provide incorrect information about capabilities (declaring the possession of non-possessed capabilities or failing to declare possessed ones), verification of these claims would be highly desirable. However, because cyber capabilities can be developed without the necessity for detectable demonstrations and testing (unlike the very notable nuclear tests) there may be significant potential for misrepresentation. Even activities that are sensed can be problematic, due to issues with attribution.

This lack of information can be both beneficial and problematic. It is beneficial because it creates a margin of error, allowing the two sides to have capabilities more divergent than might otherwise be acceptable to prevent one side from feeling that it has the upper hand and attacking. On the negative side, a significant incorrect projection of adversary capabilities may be sufficient to create the same type of scenario where one party believes (incorrectly) that it is in their best interests to attack at present.

Given the foregoing, while the models provide a framework for considering MAD scenarios and a nomenclature and system of representation for them, they cannot guarantee that the AD capability and MAD comparison calculations are correct in any particular conflict. The quality, completeness and accuracy of the information fed into the models is absolutely critical to ensuring that the answer produced is suitable for decision making.

8. Conclusions and future work

This paper has considered the dilemma presented by the existence of multiple AD technologies that have different scopes, immediacy, long term impact and methods of impact. In particular, it has considered how MAD scenarios could play out across multiple domains and mediums and how a MAD scenario could be created from AD technologies from different domains and of different capabilities that are satisfactorily paired to counterbalance the adversary's own capabilities. Further, this paper has presented models for single domain, two adversary scenarios as well as advanced scenarios where there are multiple domains involved, multiple adversaries and adversaries have multiple capabilities in some or all of the domains. It has described how these models can be used to evaluate, discuss and present work in analyzing MAD. It also discusses limitations on the models, principally due to their reliance on human input.

Future work will include the consideration of the incorporation of non-state actors who possess some AD capabilities and may factor into MAD scenarios, both at present and in the future, into models. Model development that considers issues of attribution and anonymity, including deliberate 'false flag' operations is another key area of future work. Further evaluation of the model proposed herein, through its application to relevant scenarios, is also planned.

References

[1] J. Swift, The Soviet-American arms race, Hist. Today's Hist. Rev. 63 (2009).

[2] W. Curtis, The assured vulnerability paradigm: can it provide a useful basis for deterrence in a world of strategic multi-polarity? 16 (3) (2000) 239–256.

[3] Z. Brzezinski, The Cold war and its aftermath, Foreign Aff. 71 (1991).

[4] I. Salehyan, The delegation of war to rebel organizations, J. Confl. Resolut. 54 (3) (Jun. 2010) 493–515.

[5] B. Conry, "U.S. 'Global Leadership', A Euphemism for World Policeman, (1997).

[6] D.E. Denning, Information Warfare and Security, ACM Press, New York, 1999.

[7] D.J. Schouten, U.S. Strategic Communications against Islamic Fundamentalists, " Naval Postgraduate School, 2016.

[8] P.R. Baines, N.J. O'Shaughnessy, "Al-Qaeda messaging evolution and positioning, 1998–2008: propaganda analysis revisited, Public Relations Inq. 3 (2) (May 2014) 163–191.

[9] S. Gorka, D. Kilcullen-Al-Qaida Training, Manual, "who's winning the battle for narrative? in: J. Forest (Ed.), In Influence Warfare, Praeger, Westport, CT, 2009, pp. 229–240.

[10] G.R. Dimitriu, Winning the story war: strategic communication and the conflict in Afghanistan, Public Relat. Rev. 38 (2) (Jun. 2012) 195–207.

[11] V. Andre, The janus face of new media propaganda: the case of Patani neojihadist YouTube warfare and its islamophobic effect on cyber-actors, Islam Christ.

Relations 25 (3) (Jul. 2014) 335–356.
[12] T. Molony, "Social media warfare and Kenya's conflict with Al Shabaab in Somalia: a right to know? Afr. Aff. 118 (471) (Sep. 2018).
[13] N. Inkster, Information warfare and the US presidential election, Survival 58 (5) (Sep. 2016) 23–32.
[14] K. Cole, Turning Cyberpower into Idea Power: the Role of Social Media in US Strategic Communications, (2011).
[15] S.D. Santa Maria, Improving Influence Operations by Defining Influence and Influence Operations, (2013).
[16] S. Tzu and L. Giles (Translator), "The Art of war.".
[17] M.C. Horowitz, Long time going: religion and the duration of crusading, Int. Secur. 34 (2) (Oct. 2009) 162–193.
[18] G. Weimann, Cyber- *Fatwas* and terrorism, Stud. Confl. Terror. 34 (10) (Oct. 2011) 765–781.
[19] J. Forest, Perception challenges faced by Al-qaeda on the battlefield of influence warfare on JSTOR, Perspect. Terror. 6 (1) (2012) 8–22.
[20] G. Blanquart, D. Cook, Twitter influence and cumulative perceptions of extremist support: a case study of geert wilders, Australian Counter Terrorism Conference, 2013.
[21] J. Kennedy, The Russian Battlespace of the Mind, " Carlisle, PA, 2017.
[22] J.S. Samet, "The U.S. Needs, Trolls: Strategic Concepts to Command the Gray Zone Struggle Technology and Doctrine to Revolutionize Influence Warfare, Newport, RI, 2018.
[23] E.M. Lopacienski, W.M. Grieshaber, B.M. Carr, C.S. Hoke, Influence Operations: Redefining the Indirect Approach, Naval Postgraduate School, 2011.
[24] J. Straub, T. Traylor, Towards an influence model for cybersecurity and information warfare, Proceedings of the 2018 International Conference on Computational Science and Computational Intelligence, 2018.
[25] J. Andress, S. Winterfeld, Cyber Warfare, second ed., Syngress, Waltham, MA, 2014.
[26] M. Shoaib, Conceptualising cyber-security: warfare and deterrence in cyberspace, J. Strateg. Aff. 2 (1) (2017).
[27] W.A. Vacca, Military culture and cyber security, Survival 53 (6) (Dec. 2011) 159–176.
[28] A. Kosenkov, Kosenkov, Alexander, Cyber conflicts as a new global threat, Future Internet 8 (3) (Sep. 2016) 45.
[29] E. Gartzke, The myth of cyberwar: bringing war in cyberspace back down to earth, Int. Secur. 38 (2) (Oct. 2013) 41–73.
[30] C. Whyte, "Developed states' vulnerability to economic disruption online, Orbis 60 (3) (Jan. 2016) 417–432.
[31] J. Arquilla, From blitzkrieg to bitskrieg: the military encounter with computers, Commun. ACM 54 (10) (Oct. 2011) 58–65.
[32] R.M. Rustici, Cyberweapons: leveling the international playing field, Parameters 41 (3) (Sep. 2011) 32–43.
[33] R. Jervis, Mutual assured destruction, Foreign Pol. 133 (Nov. 2002) 40.
[34] R. Hannesson, Does threat of mutually assured destruction produce quasi-cooperation in the mackerel fishery? Mar. Policy 44 (Feb. 2014) 342–350.
[35] H.J. Schellnhuber, Geoengineering: the good, the MAD, and the sensible, Proc. Natl. Acad. Sci. U. S. A 108 (51) (Dec. 2011) 20277–20278.
[36] N. Warshaw, Forget congress: reforming campaign finance through mutually assured destruction, UCLA Law Rev. 63 (2016).
[37] M.I. Lotan, Strategic dilemmas of WMD operators, Comp. Strat. 34 (4) (Aug. 2015) 345–366.
[38] E.S. Boylan, D.G. Brennan, H. Kahn, Alternatives to Assured Destruction, Croton-on-Hudson, New York, 1972.
[39] R. Jervis, "Why nuclear superiority doesn't matter, Political Sci. Q. 94 (4) (1979) 617.
[40] R.J., Art, "between assured destruction and nuclear victory: the case for the 'Mad-Plus' posture, Ethics 95 (3) (Apr. 1985) 497–516.
[41] R. Powell, Nuclear deterrence theory, nuclear proliferation, and national missile defense, Int. Secur. 27 (4) (Apr. 2003) 86–118.
[42] P.M. Morgan, Applicability of traditional deterrence concepts and theory to the cyber realm, Proceedings of a Workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S. Policy, 2010, pp. 55–76.
[43] M.J. Philbin, Cyber Deterrence: an Old Concept in a New Domain, Carlisle Barracks, PA, 2013.
[44] J.S. Nye Jr., Nuclear lessons for cyber security, Strateg. Stud. Q. Winter 201 (2011) 18–38.
[45] A. Bendiek, T. Metzger, Deterrence theory in the cyber-century, Lect. Notes Informatics (2015) 553–570.
[46] D.J. Lonsdale, Warfighting for cyber deterrence: a strategic and moral imperative, Philos. Technol. 31 (3) (Sep. 2018) 409–429.
[47] M.D. Crosston, World gone cyber MAD, Strateg. Stud. Q. 5 (1) (2011) 100–116.
[48] T. Ridout, Building a comprehensive strategy of cyber defense, deterrence, and resilience, Fletcher Forum World Aff. 40 (2016).
[49] A.E. Chukwudi, E. Udoka, C. Ikerionwu, Game theory basics and its application in cyber security, Adv. Wirel. Commun. Networks 3 (4) (2017) 45–49.
[50] P.K. Davis, Deterrence, influence, cyber attack, and cyberwar, New York Univ. J. Int. Law Pract. 47 (2015) 327–355.
[51] K. Geers, The challenge of cyber attack deterrence, Comput. Law Secur. Rep. 26 (3) (May 2010) 298–303.
[52] D.A. Gale, CYBERMAD: Should the United States Adopt a Mutually Assured Destruction Policy for Cyberspace, Maxwell Air Force Base, AL, 2009.
[53] A. Mokarram, European Cyber Security: a Cyber Deterrence Approach, Enschede, 2013.
[54] N.R. Huston, Mutually Assured Deletion: the Uncertain Future of Mass Destruction in Cyberspace, Air University, 2013.
[55] J. Straub, Consideration of the use of autonomous, non-recallable unmanned vehicles and programs as a deterrent or threat by state actors and others, Technol. Soc. 44 (2016).
[56] J. Mullen, Y. Wakatsuki, C. Narayan, Hiroo Onoda, Japanese Soldier Who Long Refused to Surrender, Dies at 91, *CNN Website*, 2014.
[57] T. Thornhill, Dead at 91, the Japanese WW2 Soldier Who Refused to Surrender for 30 Years while Hiding in Philippines Jungle, Dly. Mail, 2014.