Chasing Its Tail
Nuclear Deterrence in the Information Age

Stephen J. Cimbala

Twenty-first-century nuclear arms control and deterrence will take place in a technology context that privileges the smaller, the faster, and the more agile over the larger, the slower, and the less adaptive. At the high end of conventional deterrence and war-fighting capabilities are included long-range conventional precision strike, advanced C4ISR (command, control, communications, computers, intelligence, surveillance, and reconnaissance), network-centric warfare, and the forward movement, at uncertain paces, of defense-related nanotechnology and artificial intelligence. Meanwhile, nuclear weapons remain in the arsenals of leading powers and in the aspirational tool kits of putative regional hegemons or potentially disruptive rogue states.

This present and emerging context for nuclear arms control and deterrence leads into politico-military conundrums and paradoxes. First, cyber war and nuclear deterrence may emerge as overlapping jurisdictions, bringing new complexity into the fabric of US and other military-strategic planning. Second, antimissile defenses based partly on new technologies may finally challenge the hitherto supreme status of offensive nuclear launchers. If so, then a third outcome is possible. Instead of the venerable Cold War-era triad of intercontinental land- and sea-based missiles and bombers or the post–Cold War triad of nuclear and conventional offensive forces, defenses, and supporting infrastructure, a new “triad” of cyber strategy, minimum nuclear deterrence, and antimissile defenses might merit further descriptive attention from strategic thinkers and policymakers.

The author gratefully acknowledges Jacob W. Kipp, Michael Noonan, and Timothy Thomas for helpful insights and suggestions pertinent to this research. None bears any responsibility for arguments or analysis herein.

Stephen J. Cimbala is distinguished professor of political science at Penn State Brandywine. Dr. Cimbala is the author of numerous books and articles in national security studies, nuclear arms control, and other fields and is an award-winning Penn State teacher. His current research focuses on nuclear weapons in the information age and US-Russian nuclear arms control.
Cyber and Info Wars: Concepts Aplenty

Academic and professional literature and US government agencies already offer a rich menu of definitions for important cyber-related concepts, including cyberspace and cyber power. The Department of Defense’s first formal cyber strategy, released in July 2011, anticipated that some attacks on US information systems would meet traditional definitions of war, perhaps justifying retaliatory responses that were either cyber, or kinetic, or both. Information warfare can be defined as activities by a state or nonstate actor to exploit the content or processing of information to its advantage in time of peace, crisis, or war and to deny potential or actual foes the ability to exploit the same means against it. This is an expansive, and permissive, definition, although it has an inescapable bias toward military and security-related issues. Information warfare can include both cyber war and net war.

The related concept of cyber deterrence involves degrees of uncertainty and complexity that require a leap of analytic faith beyond what we know, or think we know, about conventional or nuclear deterrence. Cyber attacks generally obscure the identity of the attackers, can be initiated from outside of or within the defender’s state territory, are frequently transmitted through third parties without their complicity or knowledge, and can sometimes be repeated almost indefinitely by skilled attackers, even against agile defenders. On the other hand, systems are vulnerable only to the extent that they have flaws unknown to the defenders that can actually be exploited by attackers. In addition, the impact of any cyber strike is relative to the time needed to recover the attacked system—of which neither attacker nor defender would have preattack knowledge. For these and other reasons, the contrast between the principles of cyber deterrence and nuclear deterrence encourages modesty in the transfer of principles from the latter to the former. As Martin Libicki summarizes,

In the Cold War nuclear realm, attribution of attack was not a problem; the prospect of battle damage was clear; the 1,000th bomb could be as powerful as the first; counterforce was possible; there were no third parties to worry about; private firms were not expected to defend themselves; any hostile nuclear use crossed an acknowledged threshold; no higher levels of war existed; and both sides always had a lot to lose.

Airpower theorist Benjamin S. Lambeth regards cyberspace as part of the third dimension of warfare that also includes air and space operations. Cyberspace, according to Lambeth, is the “principal domain” in which
US air services “exercise their command, control, communications, and ISR (intelligence, surveillance, and reconnaissance) capabilities that enable global mobility and rapid long-range strike.” In addition, US dominance—or falling behind—in cyberspace has repercussions for the nation’s success or failure in aerospace and other domains of conflict. Lambeth’s effort to conceptualize cyber power or cyber war in a larger context is supported by Colin S. Gray, who cautions against over-mystification of the problem of cyber strategy:

When you use the term cyber strategy you risk misleading people into thinking that they are entering a new and mysterious domain. Happily, we know a great deal about strategy. We should, with 2,500 years of past experience from which to learn. And we have readily to hand a good enough general theory of strategy that certainly has authority over cyber power.

**Attacking in the Cyber Realm**

Experts foresee that some kinds of cyber war will be part of many future military conflicts. But the term cyber war may be misleading, since attacks on computers and networks are only one means of accomplishing the critical objective of neutralizing an enemy’s critical infrastructures. The purpose of information and infrastructure operations (I2O) would not be mass destruction (although destructive secondary effects are possible), but both mass and precision disruption. According to some scholars, the purpose of an information and infrastructure operation would be to “disrupt, confuse, demoralize, distract, and ultimately diminish the capability of the other side.” This concept lends itself to candidate consideration for a nuclear responsive deterrent mission.

Under the assumption of future Russian and US strategic nuclear forces limited to 1,000 or so deployed offensive weapons with operational performance parameters comparable to present systems, each side would reasonably expect to retain some hundreds of second-strike survivable and retaliating weapons. Allocating these weapons to targets requires parsimonious retailing of weapons against targets (unlike the wholesale overkill of the high Cold War). Fighting a counterforce war against the other side’s remaining nuclear forces would rapidly deplete a force already challenged to maintain any capacity for escalation control and war termination, or for continued postwar nuclear power status. Blowing up the cities of the other side is easily accomplished but not necessarily empowering of
strategic aim or military objective. It makes sense only as an option withheld for possible future use to deter the adversary from taking a similar step.

Instead of Cold War–style counterforce or countervalue targeting (the former futile, and the latter gratuitously inhumane), US and Russian plans for retaliation might emphasize counter–information and infrastructure strikes. The cyber and industrial recuperative capabilities of a state, including electricity, transportation, refineries, depots, and military-supporting industries—together with partial disruption of warning, command-control-communications, and reconnaissance capabilities—could paralyze decision making and limit military options. Although civilian casualties would be unavoidable from widespread I2O attacks, they would not be the object. Information-infrastructure targeting could threaten to inflict decisive paralysis on the opponent’s military information systems or civil infrastructure with minimal physical damage, provided an imaginative cyber component survived the other side’s attack. Instead of a second-strike capability for mass destruction, an I2O-focused minimum deterrent would pose the credible threat of focused and mass disruption.15

One can imagine three objections to the preceding suggestions. First, increasing capabilities for I2O strikes might raise the appeal of preemption for a state. As opposed to riding out an attack and retaliating, a state might be so fearful of its cyber vulnerability that it would prefer to wager on anticipatory attacks (preemptive or preventive) instead of responsive ones. This concern is not unreasonable, especially since the identity of a cyber attacker is easier to conceal than that of a kinetic first striker. A second objection to I2O targeting for nuclear retaliatory forces is that it might not be scary enough to dissuade determined attackers. Only assured destruction of the opposed regime or its society as a functioning entity would assuredly deter in this view. However, even during the Cold War, “assured destruction” represented a mistaken view of leaders’ actual decision matrices (John F. Kennedy’s national security advisor McGeorge Bundy had the last word on this, with his equation of 10 nuclear weapons on 10 cities as a “disaster beyond history”). During the Cuban missile crisis of 1962, for example, the ExComm advisory group to President Kennedy was most anxious to avoid a war, regardless of the putative pre-war US nuclear superiority in the numbers of deployed and second-strike-survivable strategic nuclear weapons.
A third objection to an I2O-oriented second-strike capability as the basis for US-Russian nuclear deterrence is that the conditions and expectations for terminating a cyber war or a cyber component of a larger war are not well understood compared to more conventional or predigital conflicts. One aspect of this inscrutability for cyber conflicts has already been noted: the identity of the first striker or “perpetrator” might be unknown and undetectable within the time available for deciding upon retaliatory options. Another aspect is that nuclear destruction might remove reliable means of communication, including power grids, satellite links, and underground cables, between adversaries otherwise intending to negotiate for war termination. This third objection also includes the possibility that obscured identities and mistaken perceptions by one or both sides could be exploited by third parties or additional troublemakers who took the opportunity to scavenge while vultures fought over their respective carcasses.

The objections relevant to any war with a heavy cyber component suggest that a nuclear deterrent based mainly on I2O retaliation should leave the door open for the inclusion of conventional long-range weapons (so-called PGS, or precision global strike weapons) in the responsive repertoire. Russia’s aversion to US prompt global strike systems is well known, based on the Russian military’s fear of US conventional deep-strike capabilities in the European theater of operations and globally. Russia’s wariness on this score reverts to its analysis of the US air-ground campaign against Iraq in 1991, especially the 37-day air war. Russia’s post–Cold War inferiority to NATO in conventional military capabilities, together with its allergy to NATO enlargement, creates for US and NATO—mistrusting Russians a picture of a conventional theater-strategic NATO option for a twenty-first-century Barbarossa. Even short of war, NATO enlargement and conventional deep strike, supported by US global supremacy in C4ISR and prompt global strike systems, could deter Russia from using the threat of force against former Soviet states now inside, or aspiring to join, NATO.

Granted Russia’s pessimism on this score, the United States may nevertheless choose to equip itself with retaliatory options of global reach and using conventional weapons. Launchers specifically dedicated for this mission, together with long-range and airborne hypersonic technology vehicles (HTV), could be included in any future war plan that seeks to accomplish national objectives with minimum collateral damage. The airborne element might eventually include purpose-built remotely piloted aircraft or technologically enhanced space planes. Russia’s objection, that it might
confuse the launch of a conventional PGS system with the firing of a US nuclear first strike, can be met by verifiable separation of PGS-capable and nuclear-tasked launch vehicles. As part of any US strategic retaliatory force, conventional PGS systems could deliver electromagnetic-pulse weapons, microwaves, or other devices to cripple the effectiveness of enemy computers, electronics, and other cyber assets. Conventional PGS systems, in addition to their roles in any strategic retaliatory force, could be used preemptively against terrorist storage bunkers (including bunkers storing weapons of mass destruction).

Cyber weapons used prior to or during a nuclear attack, or even during a nuclear crisis, might qualify as conventional or unconventional, depending on taste. It would be a stretch to refer to them as nuclear or even as weapons of mass destruction (although, as already argued, not as weapons of mass disruption). The issue of whether to incorporate cyber or information weapons into standing targeting plans involves complexities not addressed here. The most effective exploitation of cyber or information weapons depends on their flexibility and capacity for turning on a dime relative to the opponent's ability to complete its decision loop. On the other hand, one can imagine cyber weapons as part of preplanned attacks: viruses, Trojan horses, worms, and other corrupters of the integrity of opponents’ software systems could be planted months or years in advance of expected conflicts. Perhaps in acknowledgment of the risks of cyber dependency or digital fixation, the US Army now conducts some training exercises where units are required to turn off some of their Force XXI battle command-control systems—both to ascertain how well the troops do without them and to train troops for information-deficient environments in battle.17

Ongoing cyber attacks in peacetime to test the resiliency of competitors’ safeguards have become so routine that indignation is rare and reportage long ago lost any “gee whiz” overtones. For example, the most remarkable aspect of the reported attacks on Iran’s nuclear infrastructure by the Stuxnet worm, widely attributed to Israel and/or the United States, might be the relatively low-key manner with which the regime in Tehran reported the episode and downplayed its significance. Stuxnet raises the possibility of a growth industry for researchers in the use of cyber weapons for counter-proliferation, with the attendant difficulties of source identification and acknowledgment.18
Assured Retaliation

Suppose, for the sake of argument, that the abstract notion of basing a minimum US or Russian strategic nuclear deterrent on I2O targeting found resonance among defense planners in both states. Could it be implemented with forces at or below 1,000 operationally deployed long-range nuclear weapons? The following analysis interrogates that issue in several stages. First, we analyze hypothetical post–New START Russian and US strategic nuclear forces for their ability to provide for assured second-strike retaliation. Second, we ask whether the deployment of antimissile defenses by either or both states would preclude the effectiveness of minimum deterrence, regardless the targeting emphasis of retaliatory forces on I2O or otherwise. Third, we interrogate the model for insight into possibly combined effects of cyber and kinetic strikes.

Figure 1 summarizes the estimated numbers of surviving and retaliating second-strike warheads for US and Russian strategic nuclear forces under a deployment limit of 1,000 weapons. Each state deploys a balanced triad of launchers. The numbers of second-strike surviving and retaliating warheads are tabulated under four conditions of alertness and launch doctrine: (1) generated alert and launch on warning (Gen/LOW), (2) generated alert and riding out the attack (Gen/RO), (3) day-to-day alert and launch on warning (Day/LOW), and (4) day-to-day alert and riding out the attack (Day/RO).

Figure 1. US-Russia surviving and retaliating warheads (1,000-deployment limit)
(Source: Figures 1–6 are based on a model originally developed by James J. Tritten and subsequently modified by the author. Dr. Tritten is not responsible for its use here nor for any arguments or conclusions.)
Figure 2 replicates the analysis summarized in figure 1 but with a smaller maximum number of 500 deployed long-range weapons for each state. The results displayed in figures 1 and 2 suggest that Russia and the United States could provide for stable deterrence based on assured second-strike retaliation with numbers of deployed weapons significantly lower than those provided for in New START (or, conceivably, could not, if political relations soured and expectations of “reset” and rapprochement were replaced by expectations of a renewed nuclear arms race—politics rules!). In the present illustrations, under a deployment limit of 1,000 or 500 weapons for each state, either a balanced triad of launchers or hypothetical alternatives (interesting in case of lags in modernization, especially for Russia) provide from hundreds to many tens of thousands of surviving and retaliating weapons under every condition of alertness and launch doctrine. Although leaders in the United States and in Russia have presently ruled out any departure from triads of intercontinental launchers, future exigencies or attractive technologies might change this calculation.

![Figure 2. US-Russia surviving and retaliating warheads (500-deployment limit)](image-url)
Missile Defenses

Would missile defenses complement or conflict with the objective of minimum deterrence through reductions in offensive nuclear forces, including the option of increased emphasis on I2O targeting? In figures 3 and 4, US and Russian second-strike retaliatory forces are opposed by missile and air defenses with drawdown curves of effectiveness against penetrating ballistic missiles and aircraft-delivered weapons from 20 to 80 percent. The upper tier of defenses in this graphic provide an optimistic performance expectation for missile and antiair defenses judging by today’s standards, but it allows room for improvements in ballistic missile defense (BMD) performance that might materialize between now and 2018–2020 (the New START due date for implementation of treaty reductions and the final stage of planned European phased adaptive approach [EPAA] missile defense deployments). Figures 3 and 4 summarize the numbers of second-strike surviving and retaliating warheads for each state under the initial deployment limits of 1,000 weapons and 500 weapons, respectively. For the sake of consistency, all retaliatory forces are operating under conditions of generated alert and riding out the attack (Gen/RO), and both sides are deploying triads.

Figure 3. US-Russia surviving and retaliating warheads vs. defenses (1,000-deployment limit)
The results summarized in figures 3 and 4 offer mixed messages for US and Russian military planners and for students of nuclear arms control. On one hand, post–New START nuclear retaliatory forces, even at minimum deterrent levels, can conceivably provide for numbers of surviving and defense-penetrating warheads adequate to support a strategy of stable deterrence. On the other hand, as deployed defenses gradually improve, they make it harder to build flexibility into retaliatory targeting options. Deploying defenses that are too capable against either side’s nuclear retaliatory forces could drive military planners into launch-on-warning doctrines, increased expenditures on offensive countermeasures to defenses, or additional deployments of offensive weapons.21

Even technically improved antimissile defenses relative to offenses leave open ended the strategic and political priorities that will determine future US and Russian defense modernization. Opportunities exist for misunderstanding and misperception, creating further distance between the security agendas of Washington/Brussels and Moscow and postponing the extension of the European security community eastward to include Russia as a participant and not just as an observer.22 Russian political leaders and technical experts argue the case for participation with NATO in a European-
wide missile defense system, even as they warn of a European BMD danger to Russia’s deterrent and advocate deployment of new offensive land- and sea-based missiles equipped to defeat such antimissile systems. They are concerned about the potential threat to Russia’s deterrent and advocate deployment of new offensive land- and sea-based missiles equipped to defeat such antimissile systems.23 The Obama “reset” with Russia is also under siege in US domestic politics, adding to uncertainty with respect to future US-Russian security cooperation or lack thereof.24

US and Russian arms controllers who are attempting to detoxify the potential conflict between further offensive force reductions and missile defenses might be fighting the wrong corner. An information-infrastructure deterrent might rely less on antimissile or air defenses—or countermeasures to those defenses—than traditional models based solely on kinetic factors would suggest. Instead, an I2O first- or second-strike force might exploit the electronic spectrum and the information grid of its opponent for disruption that swept around, over, and under the sensor and shooter exchanges previously thought of as dispositive.25 Related to this possibility, Russia’s war against Georgia in August 2008 demonstrated how cyber war and information operations might be used in support of conventional military operations. The Russian cyber campaign reportedly attacked some 38 Georgian and Western websites upon the outbreak of war, including ranking Georgian government offices and the US and British embassies in Georgia, and appeared to be centrally directed and coordinated with the tempo of force operations.26

Instead of a single integrated operational plan (SIOP), however flexible, for fighting a nuclear war if deterrence failed, planners would have to devise a matrix of plans linking information strike with kinetic options. How complicated this might be is probably beyond the power of mere mortals to demonstrate with any proficiency—much is speculative as to the two-way complexity of combined cyber and nuclear or conventional kinetic attacks. On the other hand, analysts and planners must do what they can in the face of questions and demands for performance that will not go away.

A simplified approach to one aspect of a cyber-soaked SIOP might be illustrated as follows. Let us assume that both the United States and Russia were required to carry out second-strike retaliation after having absorbed both cyber and kinetic first strikes. To measure the impact of such strikes, we estimate that the cyber component directly or indirectly neutralizes as many surviving and retaliating weapons as does the kinetic portion. The second-strike surviving forces would therefore be in a position equivalent to
that of a third striker in a series of exchanges without information weapons. In effect, they would be fighting World War III-b. The additive effects of both cyber and kinetic strikes are summarized in figures 5 and 6 representing the 1,000- and 500-weapon prewar deployment limit (without defenses), respectively.

Figure 5. US-Russia surviving and retaliating warheads with information and kinetic attacks (1,000-deployment limit)

Figure 6. US-Russia surviving and retaliating warheads with information and kinetic attacks (500-deployment limit)
Figures 5 and 6 show that, in a hypothetical but not necessarily unrealistic exercise of cyber-kinetic nuclear strike plans, the United States and Russia could still retain sufficient numbers of weapons to create historically unprecedented and socially unacceptable damage in retaliation. Cyber attacks on command-control, communications, and warning systems might lead to ragged retaliations and strikes more dependent upon the most survivable launch platforms such as submarines and mobile missiles. Alternatively, two expectations about such a scenario would be mistaken. First, information operations cannot make any nuclear war between states with large arsenals into a surgical operation or an exercise in “soft” power. Second, a state’s cyber and kinetic strategies need to be carefully coordinated as to their political and military objectives, not only up to the brink of war but even beyond that threshold. Otherwise, the objectives of escalation control and conflict termination will be impossible to realize for either state when its opposite number is brain dead as well as partly but not completely disarmed.

**Conclusion**

Faced with exigent threats, states with cyber capabilities will be tempted to employ them to good effect. For example, imagine a replay of the Cuban missile crisis between a future Russia and the United States, with Russia having deployed nuclear-capable missiles and/or warheads into South Ossetia. Or, to flip the example, hypothesize a NATO missile defense installation deployed to protect Tbilisi or Kiev, supported by short- and medium-range ballistic missiles as a trip wire. One can expect that cyber operations of the information-technical type (attacking enemy systems and networks) as well as the information-psychological variety (influencing public opinion among foreign and domestic audiences, including elites and general publics) will commend themselves to peacetime and crisis political leaders and their military advisors.27

The larger context for cyber operations and nuclear deterrence also involves the possible adoption of minimum deterrence force postures and the deployment of missile defenses by the United States and NATO or perhaps others. Minimum deterrence might appeal to the United States and to Russia under very favorable political conditions, including a rethink of European and central Eurasian territory as a unified security community instead of as a fight club. In this regard, the United States and NATO phased adaptive approach to missile defenses offers the choice
between cooperative security and Cold War retro approaches to arms control. Regardless the outcome of the imbroglio over EPAA, US plans for a global missile defense system will include technology transfers and security cooperation with regional allies in Europe, the Middle East, and Asia. Prospective US opponents in those regions may therefore cultivate both nuclear deterrence and information operations as means for antiaccess and area denial (A2/AD) deterrence and defense.

Nuclear deterrence in the Cold War was something sui generis that grew from a way station for coping with new weapons and new threats into an all-purpose solvent for problems of military strategy. Nuclear weapons remain alive and menacing in the twenty-first century, but they are presently and prospectively circumscribed by new contexts. One of these contexts is the coexistence of information warfare or military cyber operations and nuclear deterrence.

Notes


8. Libicki, Cyberdeterrence and Cyberwar, xvi. See also Cooper, New Approaches to Cyber-Deterrence, 2–4 and passim.


10. Ibid., 51.


13. Ibid., 19.

14. Ibid.

15. An alternative minimum deterrent proposal based on infrastructure targeting and intended as a way station to nuclear zero is outlined in Hans M. Kristensen, Robert S. Norris, and Ivan Oelrich, From Counterforce to Minimal Deterrence: A New Nuclear Policy on the Path Toward Elimi-


17. I gratefully acknowledge Michael Noonan, Foreign Policy Research Institute, for this insight. He is not responsible for its application here.


20. Grateful acknowledgment is made to Dr. James Tritten for use of a model he originally developed. Dr. Tritten is not responsible for any analysis or arguments here.


25. However, it deserves emphasis that even “small” nuclear attacks against infrastructure would leave historically unprecedented damage with large recovery times. For illustrations, see Office of Technology Assessment, The Effects of Nuclear War (Washington: Government Printing Office, 1979), 63–80.

26. Interestingly the attacks might not have been conducted directly by the Russian government but by sympathetic shadowy groups such as the Russian Business Network, notable for cyber crimes including the kind of DDOS (distributed denial of service) attack visited upon Georgian websites. See Ariel Cohen and Robert E. Hamilton, The Russian Military and the Georgian War: Lessons and Implications (Carlisle, PA: Strategic Studies Institute, June 2011), esp. 44–49; and Timothy Thomas, Recasting The Red Star: Russia Forges Tradition and Technology through Toughness (Fort Leavenworth, KS: Foreign Military Studies Office, 2011), esp. 247–51.

27. For this distinction, following Russian nomenclature, see Thomas, Cyber Silhouettes, 74.