IJEMD-SS, 5 (2) (2026) https://doi.org/10.54938/ijemdss.2026.05.2.676

International Journal of Emerging Multidisciplinaries: Social Science

Research Paper
Journal Homepage: www.ojs.ijemd.com
ISSN (Online): 2958-0277

Machine-Speed Deterrence: Institutional Capacity as the Moderating Variable in AI-Driven Cyber Crises

Noureen Akhtar & Dr. Nazish Mahmood

PhD scholar at School of Politics and International Relations, Quaid-i-Azam University, Islamabad

Noureenakhtar_dss@live.com (Corresponding Author)

Assistant Professor at School of Politics and International Relations, Quaid-i-Azam University, Islamabad

nazish@qau.edu.pk

Abstract

In this paper, Erosion of Deterrence in Rivalrous Dyads under Conditions of Cyber Evolution via Artificial Intelligence (AI) is advanced as moderated by absorptive institutional capacity of the receiving state, rather than by the offensive volume of cyber capability enabled by AI. Drawing on the literature on the techno-realist synthesis of technology and power (Sukma, 2024), and the cyber-deterrence literature which has been long preoccupied with the attribution issue problem (Lindsay, 2015; Nye, 2017), the paper captures three erosion processes: operational acceleration, attribution uncertainty, and signalling ambiguity, and models institutional capacity, defined in four dimensions and six observable indicators, as the variable that determines the absorption or amplification of these erosion

processes. The India-Pakistan cyber competition is used as an illustrative case to illustrate the model, which is not tested. The paper's contribution is a framework that is portable and independent of the adversary, called machine-speed deterrence, to assess the stability of deterrence in an age of AI-compressed conflict, as well as policy recommendations that are geared toward institutional investment as opposed to capability investment.

Keywords: deterrence stability, artificial intelligence, institutional capacity, cyber conflict, techno-realism

1- Introduction

The prevailing thinking in modern security studies is that as a state gains more sophisticated AI powers for its cyber offensive, its threat to strategic stability increases.The prevailing notion in contemporary security studies is that, the more advanced a state's AI capabilities in the offensive domain, the more worrisome it is for strategic stability. This paper proposes that this intuition is only half correct and that, in an important sense, it is incorrect. Two states could be exposed to similar cyber pressure from AI, and yet react very differently, depending on whether they have the absorptive capacity of the institutions of the state under pressure. When institutions can work across technical and strategic levels, coordinate incident response and escalation decision-making, operate within compressed decision cycles, and maintain operations in times of disruption, AI-driven cyber pressure is significantly mitigated. When these abilities are divided, that pressure will be increased to strategic instability.

This claim is elaborated upon in this article as a moderated-variable model of deterrence stability. In the techno-realist view that technology is a medium of power whose strategic impact is shaped by the political and organisational setting in which it is employed (Sukma, 2024), the paper identifies three pathways by which AI-based cyber evolution puts a strain on deterrence relationships: operational acceleration, attribution uncertainty, and signalling ambiguity. It then uses the construct of institutional capacity, defined across four dimensions and six observable indicators, to model the moderating role between these pressures and their absorption and/or amplification. This high tempo, compressed environment, where the human and institutional

tempo of strategic deliberation rapidly falls behind the speed of AI-supported operations, is known as machine-speed deterrence.

The framework is not statistically tested, but rather illustrated, using the conflict between India and Pakistan as an example of a dyad with a constant technology and/or adversarial relationship between the two actors, and a variable institutional configuration. The May 2025 crisis, which began with the attack on Pahalgam and followed India's Operation Sindoor, provides a first case study in the rivalry in which cyberspace served as a real, and AI-enhanced, arena of conflict, and is cited here as a case example, but not a definitive test of the model.

The paper is divided into five parts. Section two summarizes the literature on cyber-deterrence, the speed of AI decision-making and institutional capacity, while placing techno-realism as an integrating frame. Section three builds up the moderated-variable model and its three mechanisms. The paper's process-tracing and comparative-institutional approach and its evidentiary boundaries are outlined in Section four. Section five is illustrative of the application of the framework in the case of India and Pakistan, including the institutions they have declared and the May 2025 crisis. The paper concludes by considering theoretical and policy implications in Section six.

The stakes of the reframing are not limited to a dyadic relationship. The ability to compare capabilities risks is a research and policy agenda that focuses on keeping scarce resources on the offense, rather than developing the capacity to withstand an adversary's offense, as AI spreads across the cyber operations of more and more state and non-state actors. On the model's own terms, the case made here has a direct policy implication: For states with fewer resources than their rivals, where keeping up with them technologically is simply not a practical proposition, an institutional investment in coordination, integration, responsiveness, and resilience might be a more attainable and, on the model's own terms, more consequential path to deterrence stability than capability competition.

2- Literature Review

2.1 Cyber Deterrence and the Attribution Problem

Classical deterrence theory relies on a condition of credible threat of retaliation: if an adversary is expected to be deterred without committing to an action, the cost of that action (sent by a "legible signal") is greater than the expected gain from (or benefit of) that action (Schelling, 1966). Moving this reasoning into cyberspace has been challenging. The Cold War notion of "massive retaliation" is not a good analogy for cyber behaviour, Nye (2017) says, and proposes four mechanisms of cyber prevention , threats of punishment, threats of denial, entanglement, and normative taboos , of which the former, punishment, is significantly limited by the attribution problem. In other words, low-value, high-volume targets are harder to attribute to and hence harder to deter with punishment, whereas high-value targets create more investigative "footprint.As Lindsay (2015) formalises it, low-value, high-volume incursions are more difficult to attribute to and therefore more difficult to deter with punishment, while high-value targets create more investigative "footprint. This scale-dependence is important to the present argument because it is at the level of low value, high volume intrusions that AI-powered tooling will make deterrence by punishment less effective, thereby increasing the space in which it is less likely to work.

A second line of the literature places cyber conflict into the framework of the security dilemma. The same applies when it comes to cybersecurity, notes Jervis (1978) when intentions behind an actor's capability cannot be reliably discerned; and Buchanan (2017) explicitly brings it to cyberspace, where the very practice that states engage in to secure their own networks are often indistinguishable from preparations for attacks, creating a structurally embedded cybersecurity dilemma. Even if international order is not directly affected by the proliferation of offensive cyber capability, Kello (2017) contends that it disturbs the conceptual foundations of international order. In a nuclear dyad, this uncertainty adds to Snyder's (1965) long-standing paradox of the restraint posed by a nuclear overhang and the greater risk-taking that could occur at the sub-conventional level of strategic interactions, as both sides understand that a nuclear escalation is unlikely.

2.2 Artificial Intelligence and the Compression of Decision Time

The strategic implications of AI-driven speed-ups in military decision-making are explored in a different literature, generally using the heuristic of Boyd's Observe-Orient-Decide-Act (OODA) cycle. Johnson (2022) warns that devising speed is the wrong way to apply the OODA framework because its primary impact is not on how quickly a battle proceeds but on how well the

commanders orient themselves to the situation, and that automation of the decision-making process may have strategic implications that neither commanders nor AI systems can fully predict. The Joint Air Power Competence Centre (2022) notes that a RAND simulation carried out in 2020 revealed that a wide-spread deployment of AI and autonomous decision-support could result in unintended escalation and crisis instability, as decision cycles become shorter and quicker when AI and autonomous systems are deployed. This is related to an observation made in the nuclear-deterrence literature, which is that the psychological and organizational limits on the ability to assess accurately the extent and capability of one's own enemy are more important in determining the success or failure of deterrence than the actual balance of capabilities; see Jervis (1979). When AI makes the decision interval smaller than the one demanded to operate for that perception, the mechanism that Jervis is talking about is directly stressed.

2.3 Institutional Capacity and Cybersecurity Governance

A third and less widely explored literature considers institutional and organisational capacity instead of technological capacity to be the determinant of whether a new technology has strategic impact. Horowitz (2010) demonstrates that the spread of military innovations between states depends on financial and institutional capabilities: different technology outcomes yield different strategic effects, depending on the institutional context of the states in which they are adopted. In literature specifically to cyber security-governance, Nfuka and Rusu (2019) identify that the maturity of the information-technology governance structure is far more significant than the sophistication of the technology used to predict the performance of public-sector cyber security systems in a developing-country context. The literature rarely puts into direct dialogue with cyber-deterrence theory which has largely been used as an explanatory variable, rather than a background condition. The crux of the present paper is to make that variable explicit and weighty.

2.4 The AI Arms-Race Debate

A closely related question is whether there is an arms race in the diffusion of military AI (mil-AI) to the first place. The arms-race analogie is also descriptively wrong, as AI is a technology that is not like a weapon; it is a dual-use technology that is emerging out of commercial development; and it is an extremely diffuse and hard-to-control technology that is not readily

apparent as a weapon, but rather as a civilian tool whose genesis is commercial, and therefore not exclusively military, and which is so hard to control that the same technology that could be used for good purposes may also be used for bad ones. Other scholars argue that, even if a bit of a oversimplified metaphor, the competitive dynamics that the metaphor calls forth are real and meaningful for stability and that the spread of military AI is a real pressure toward instability, which the metaphor calls forth rather usefully, even if it is somewhat overstated. The present paper takes a particular lesson from this debate: the strategic impact of AI will be determined by organisational uptake and institutional absorption of the technology and not its raw sophistication, so a research agenda based on counting and comparing capability is measuring the wrong variable. With this shift in focus, the institutions that enable states to adopt, deploy, and govern AI-enabled cyber capability are analytically the most important and interesting objects of study, and this paper aims to shift the focus in this direction.

The emphasis on institutions has origins beyond the cybersecurity-governance literature. The institutions that allow for a state's developmental capacity is not one of resources, but a specific institutional configuration, namely, "embedded autonomy": bureaucratic coherence and dense structured connections with sectors the state aims to govern. (Evans 1995) In the current context, this means the coordination and integration aspects listed below are not only desirable administrative elements but a precondition for the absorption of pressure imposed by artificial intelligence technologies, what Evans describes as embedded autonomy is a precondition for the absorption of the available economic resources in the context of industrial development, by the state. The parallel is telling, as both are about institutional structure, not about resource levels, the latter is just not fully digested yet in the cyber-deterrence literature.

2.5 Techno-Realism as an Integrating Frame

There is one framework, techno-realism, in which these three literatures may be bound together. While classical realism posits that technology serves as a form of power and that technology's strategic impact should be found in its technical characteristics, Sukma (2024) argues that technology's strategic impact cannot be separated from the identity, institutions, and organisational context of the state that uses it or absorbs it. This is the license being given to the argument here: that cyber evolution, driven by AI, does not inevitably lead to erosion of deterrence,

but rather is filtered through the capacity of the state in a state of stress. The ontological commitments behind this move are scientific realism in that it treats causal mechanisms as real even if they are not directly observable, and a critical-realist epistemology in that it treats knowledge of causal mechanisms as mediated and fallible (Bhaskar, 1975) and that it allows causal claims about mechanism plausibility without overstating what any given illustrative case can show. As noted in the literature above though, the current literature on cyber deterrence is disproportionately focused on great power dyads, mainly the U.S.-China and the U.S.-Russia dynamics; the effects of institutional capacity on cyber pressure in regional nuclear rivalries under more constrained resources are comparatively less studied.

3- Theoretical Framework

3.1 Three Erosion Mechanisms

The model includes three ways in which AI-enabled cyber evolution imposes stress on deterrence stability.

Operational acceleration is the shrinking of the decision cycle that results from the use of AI to augment detection, analysis and response tooling. The compressed time between deliberation, verification, and signaling as indicated in the literature reviewed on the OODA-loop above reduces the time span within which decision makers can deliberate, verify, and signal (Johnson, 2022; Joint Air Power Competence Centre, 2022). Operational acceleration imposes direct pressure on the signalling function on which deterrence relies: Schelling (1966) talks of a model of "bargaining" based on the notion that both parties have time to read and react to signals, but operational acceleration means that neither party can be sure of having the time to read and react. The impact is magnified at the strategic, not just tactical, level: although automation might have benefits on the speed and quality of an individual detection process, the overall pace at which it can freeze a crisis for a state is not necessarily commensurate with the institutional cycle, inter-agency consultation, ministerial briefing and cross-checking against independent sources, that states have traditionally used to transform raw information into considered judgement. Instead of the acceleration per se, it is this difference in tempo between machine and institutional that is the destabilising force of the mechanism.

The classical attribution problem deepened with AI-enabled obfuscation is attributed to attribution uncertainty. AI-assisted tooling, automated intrusion generation, synthetic media, and machine-optimised evasion minimises the expense of producing ambiguity and high volume activity just in the space where the target value was already low and the investigative effort was also low, making the zone wide enough for deterrence by punishment to be ineffective. In a further complication, the same crisis window is also used by a myriad of non-state and quasi-state actors whose operations may overlap with state-sponsored ones, even if they can be distinguished in hindsight, after undertaking a forensic investigation. Hacktivist collectives, criminal groups, and advanced-persistent-threat units linked to state bodies operate in parallel and during the same crisis window, thereby adding to the acceleration mechanism above, increasing the burden of analysis and separation.

Signalling ambiguity is the loss of interpretive clarity that costs and observability of the signal typically creates. Signalling the signalling logic described by Schelling (1966) is already reversible and deniable in cyber operations, and AI exacerbates this by making operations perhaps even more difficult to understand as they happen, or even afterwards, furthering Jervis's (1978) description of security dilemma dynamics, which has been extended to cyber by Buchanan (2017). As such, observing state cannot tell if something is a capability or an intent and as Buchanan (2017) notes, AI is likely to make it harder to determine the intent of an adversary, as it makes it harder for an adversary to understand what is happening on their systems, because the same toolkit can be used defensively to harden networks and detect intrusions or offensively to seek to intrude into an adversary's network.

Each one of these mechanisms is conceptually different, but reinforcing each other: The smaller the time frame of the window available for assignment of the attribution and signalling task, the less reliable is the assessment of the attribution and the ambiguity of the signalling task the less reliable is the attribution assessment in that time frame.

3.2 Institutional Capacity as the Moderating Variable

The model's key finding is that the three mechanisms do not all necessarily lead to an erosion of deterrence. The impact that they have is mitigated by institutional capacity of the state

under pressure (ICsUP), which is defined as the capacity of the state to absorb, not amplify, AI-driven cyber pressure. It can be measured in four aspects. Coordination is the ability of the technical, intelligence, military and civilian components involved in cyber defence to work together. Integration is the ability to tie technical incident management into the strategic level, where the escalation is determined. Responsiveness is the ability to detect, interpret and decide in the short timeframes provided by AI-powered operations. Resilience is the ability of these functions to remain unaffected by the disruption caused by a serious operation.

Table 1 summarises these four dimensions together with the demand each is designed to meet.

Table 1

Four Dimensions of Institutional Capacity

Dimension	Definition	Demand It Meets
Coordination	Capacity of technical, intelligence, military, and civilian bodies to act in concert	Reduces fragmentation across the many actors involved in cyber defence
Integration	Capacity to connect technical incident management to strategic escalation judgement	Ensures decision-makers possess an accurate technical picture
Responsiveness	Capacity to complete detection, interpretation, and decision within compressed AI-driven intervals	Directly counters the operational-acceleration mechanism
Resilience	Capacity to sustain coordination, integration, and responsiveness through disruption	Preserves institutional function during a live crisis

Note. Indicators operationalising each dimension are developed in the following subsection and applied comparatively in Table 2.

These dimensions are operationalized in terms of six observable indicators, rather than by inference or third-party report, to a state's own official record. Doctrine and strategy are published statements to determine if a state has officially expressed a stance on cyber and AI threat issues and how specific. Coordination and integration – as manifested in dedicated agencies and mandates – reveals if the coordination and integration has been an institutionalized process or remains ad hoc across ministries. Civil-military command arrangements points to whether the integration dimension is applied to the interface that will be most likely to be tested in a real crisis. The level of public-private and inter-agency coordination mechanisms reflects the extent to which the state has developed lateral linkages needed for responsiveness under compressed tempo, since most states have a large proportion of critical infrastructure that is privately owned or operated. Coordination and integration is a function of legal and governance frameworks, as are binding authority and informal convention, which are important for resilience during the stress of an actual crisis. Human capital suggests the extent to which the specialised technical and analytical skills required to make practice of the other five indicators possible, not just possible, is available in adequate numbers. This indicator set needs to be limited to what the state itself formally commits to, consistent with the evidentiary caveat in the methodology section below: The state's commitments in the official documents are not operational performance commitments, but commitments of institutional design, which must not be confused.

3.3 The Moderated Model and Machine-Speed Deterrence

This model can be summarized as a moderated-variable relationship where the process of AI-driven cyber evolution is responsible for three cyber pressures (operational acceleration, attribution uncertainty, and signalling ambiguity) that affect deterrence stability, all of which are moderated by institutional capacity. This pressure is absorbed by high institutional capacity (coded across the four dimensions above) or boosted to strategic instability by low institutional capacity. The internal logic of the model needs to be clearly expressed: Capacity does not form a fourth pressure mechanism; it does not achieve this by itself. It is the moderating process that controls the transformation of the three mechanisms already identified to deterrence erosion, or, in some

cases, the failure to transform them. The collapse of this distinction would be a misstatement of the model because institutional mediation is not in itself a causal phenomenon but rather is a function of the three mechanisms that it conditions.

This gives rise to the paper's main thesis: the level of institutional capacity determines the stability of deterrence in an era of cyber-AI evolution, with higher levels of institutional capacity resulting in more stable deterrence and lower levels leading to less stable deterrence, much of which is independent of the offensive cyber capability of the two rivalries in question, as measured by the cyber-AI they can bring to the table. The strategic environment this model describes is one in which the tempo of AI-enhanced operations rapidly exceeds that of human and institutional operations that have long been relied upon for deliberate crisis management, and is known as machine-speed deterrence. The main thrust of the paper is the main reorientation, which is not whether an adversary's cyber-AI ability is accelerating, but whether state institutions can run, coordinate and make decisions at the speed of the adversary's ability.

4- Methodology

The framework is elaborated with theory-guided process tracing (Beach & Pedersen, 2013) and the comparative institutional analysis, applied in an illustrative way to one critical case. The India-Pakistan cyber rivalry is chosen because it is a good approximation of the theoretical condition of a critical case: both countries have a long history of hostility and both face a comparable nuclear overhang with many elements of the external threat environment being largely fixed, while the institutional structures, the key variable of theoretical interest, can be demonstrated to differ in specific and reliably documented ways. This design enables the analysis to separate out the plausibility of the institutional-capacity mechanism, while not asserting that it controlled for all the confounding factors that a large-N design would need to control for.

Van Evera's (1997) four inferential tests are evaluated: the straw-in-the-wind test is a weak test that is consistent with, but not conclusive evidence of, the hypothesis; the hoop test is a weak test that does not confirm but could refute a hypothesis; the smoking-gun test is a strong test that confirms a hypothesis, but its absence doesn't eliminate the hypothesis; the doubly-decisive test is a strong test that confirms the hypothesis and eliminates its competitors. This evidence is presented

at the most, to the specific claims, the "smoking gun" standard and, for the comparative institutional picture as a whole, the "straw in the wind" standard, given the scope of a single illustrative case.

The main evidentiary basis are the state-specific official documentation, which include the National AI Policy (Government of Pakistan, 2025); the National Cyber Security Policy 2021 and its accompanying CERT Rules (Government of Pakistan, 2021, 2023); the National Security Policy 2022-2026; and the Digital Pakistan Policy (Government of Pakistan, 2018); in the case of India, the National Cyber Security Policy (Government of India, 2013); the 2025 AI Governance Guidelines (Government of India, Ministry of Electronics and Information Technology, 2025); and the Joint Doctrine for Multi-Domain Operations (Government of India, Ministry of Defence, 2025). The illustrative case of the May 2025 crisis relies on some secondary reporting from policy and technical-intelligence sources like the Observer Research Foundation, the Royal United Services Institute and a technical threat-intelligence reporting from the CYFIRMA and Silobreaker.

This design has two inherent restrictions that should be clearly acknowledged:Two restrictions immediately follow from this design that must be forthrightly stated: First, official documents reflect a commitment to institutional design rather than institutional performance; where the two differ, this paper's document-based approach does not identify the gap; and hence this institutional claim should be understood as an institutional commitment rather than as an institutional performance. Second, much of the reporting on the May 2025 crisis, casualty figures, and contested claims of aircraft losses, are open source and contested; consistent with the paper's principles of attribution, this is attributed to its source rather than adjudicated. The paper then makes claims of mechanism plausibility based on a single case not causal proof based on a representative sample.

The six-indicator specification presented in the theory above is directly followed by the comparative institutional analysis aspect of the method. Instead of creating a composite numerical index, as is done in cross-national instruments like the International Telecommunication Union's Global Cybersecurity Index (GCI), which the present paper does not use because it hides the indicator-level differences that the theoretical model makes analytically salient, each indicator is

evaluated qualitatively based on the documentary record and reported separately. This decision is part of a general methodological approach of the paper, which is to differentiate between institutional design, as declared, and institutional performance, as demonstrated, so that a single aggregate score would risk giving an impression of getting the precision of the performance that is not warranted by a document-based approach. The qualitative, indicator-by-indicator approach, however, marks exactly where the evidentiary base is strong and where it is weak, which is analytically useful as an output of the analysis for an illustrative, rather than confirmatory, document.

5- Illustrative Application: The India-Pakistan Cyber Rivalry

5.1 Comparative Institutional Snapshot

Applying the six-indicator framework to each state's official record produces the comparative snapshot summarised in Table 2.

Table 2

Comparative Institutional Snapshot: Pakistan and India

Indicator	**Pakistan**	**India**
Published doctrine & strategy	National AI Policy (2025); National Cyber Security Policy (2021); National Security Policy 2022–2026	National Cyber Security Policy (2013); AI Governance Guidelines (2025); MDO Doctrine JP 2.06 (2025)
Dedicated agencies & mandates	AI Council; Policy Implementation Cell; National CERT	Proposed AI Governance Group; sectoral cyber regulators

Indicator	Pakistan	India
Civil–military command arrangements	Statutory CERT Council with mandatory MoD and NSD membership (CERT Rules 2023)	Cross-domain integration named as a foundational MDO characteristic
Public–private & inter-agency coordination	Referenced in NSP 2022–2026 hybrid-warfare commitment	Private-sector collaboration named as a defining MDO characteristic
Legal & governance frameworks	CERT Rules gazetted under PECA 2016 (statutory)	NCSP 2013 supplemented by 2025 guidelines (recommendatory, graded liability proposed)
Human capital	Capacity-building pillar with defined KPIs (National AI Policy)	Capacity-building pillar under AI Governance Guidelines (2025)

Note. Entries record declared institutional design in official documents, not demonstrated operational performance. Sources: Government of Pakistan (2018, 2021, 2022, 2023, 2025); Government of India (2013); Government of India, Ministry of Defence (2025); Government of India, Ministry of Electronics and Information Technology (2025).

Two observations follow. First, both states exhibit significant declared institutional design on the majority of indicators, meaning that any simple narrative of either state's lack of institutions needs to be complicated. Pakistan's CERT Rules, gazetted in 2023 under the Prevention of Electronic Crimes Act, mandate a statutory "Defence CERT", mandatory membership of the National CERT Council by the Ministry of Defence and National Security Division and a National CERT mandate that explicitly includes cyber terrorism and cyber warfare (Government of Pakistan, 2023) . The Joint Doctrine for Multi-Domain Operations (MDO) set out by the Government of India, Ministry of Defence, 2025, independently mentions concepts that are very similar to the mechanisms mentioned above: Cross-Domain Operations, which is correlated to multi-domain, and includes cooperation with the private sector in fields such as cyber defence, artificial intelligence, and space-based operations. Both the 2025 AI Governance Guidelines in India and the National AI Policy in Pakistan feature dedicated pillars for Institutions with clear

key-performance indicators to monitor the implementation of AI governance across the ministries and sectoral regulators (Government of India, Ministry of Electronics and Information Technology, 2025) and (Government of Pakistan, 2025).

Secondly, and more importantly for the model to be developed here, the indicators that actually test the responsiveness and integration aspects of the model are the ones that are the most difficult to assess from documents on their own. The National Security Policy 2022-2026 explicitly pledges the country to fight hybrid warfare by improving information and cyber security, facing disinformation and influence operations, and does not mention the timetable for the fulfillment of this commitment. The methodological note at the end of the paper is apt: the design of the institutions can be reasonably documented, but the performance of the institutions, especially under the few years compressed by the model as most critical, can only be inferred from doctrine.

The other two indicators, legal and human capital, present a more mixed picture and explain why the six-indicator approach will not lend itself to an overall score. In terms of institutional architecture, while both are based on MOUs, the latter are inter-ministerial in Pakistan, the CERT Rules have been gazetted under the Prevention of Electronic Crimes Act and thus are more explicit (Government of Pakistan, 2023). India's National Cyber Security Policy 2013, on the other hand, is not related to the current moment of AI-governance, having been issued more than 10 years ago and since then significantly supplemented, but not replaced, by other guidance, such as the recommendation for a graded liability system based on risk and due diligence contained in the AI Governance Guidelines 2025 (Government of India, Ministry of Electronics and Information Technology, 2025; Government of India, 2013) which, though a recommendation, is not yet in the stage of legal consolidation that Pakistan's gazetted CERT Rules are. The record is slightly more sparse for human capital for both states, and as in most cases, a commitment to building capacity is more commonly made as a policy objective than reported as a measurable stock of trained personnel; Pakistan's National AI policy does, however, assign capacity building/talent exchange as a separate implementation pillar with specific key performance indicators (Government of Pakistan, 2025).

The comparative snapshot alone does not lend itself to an easy conclusion about the relative institutional strengths of one state vs. another. It facilitates a more precise and, for the purposes of

this paper, more illuminating statement: that both states have made significant declared commitments in most of the six indicators, that certain indicators bear more directly on responsiveness in the context of compressed tempo and are more difficult to verify from either state's documents, and that the analytical value of the framework lies less in a single comparative score than in identifying precisely where the gap between design and performance is most consequential, and least visible from open sources, precisely the gap the May 2025 crisis can illuminate illustratively.

5.2 The May 2025 Crisis as an Illustrative Episode

The crisis caused by the attack on Pahalgam on the 22nd of April 2025 and the subsequent Operation Sindoor by the Indian army on 7th May, provides an illustrative rather than confirmatory test of the three mechanisms. According to sources from the Observer Research Foundation, this became the first India-Pakistan crisis where cyberspace became an active and coordinated theatre of confrontation, with the Pakistan campaign known as Khaima being responsible for defacing websites, Pakistan-sponsored propaganda activities, and even more advanced POTA (malware and APT) campaigns targeting India in the days leading up to and after the operation (Observer Research Foundation, 2025). Silobreaker's independent data on technical-intelligence (TI) reports shows over 1.5 million attacks against Indian infrastructure in the days following Operation Sindoor, with an intersection of advanced-persistent-threat groups and ideologically motivated hacktivist collectives operating side by side (Silobreaker, 2025). The technical report from CYFIRMA reveals the infrastructure and coordinated phishing attack using a fake cloud migration notice from the National Informatics Centre (NIC) of India targeting government employees (CYFIRMA, 2025).

The three mechanisms are illustrated in these developments. The attack triggered the military response within days of the incident, following the military operation and a surge of cyber attacks, which was in line with the operational-acceleration mechanism adopted in earlier India-Pakistan crises, that typically took weeks or months. State-supported APTs and loosely organized hacktivist groups operate concurrently, making attribution difficult, if not impossible, to separate the categories, and particularly to tie a particular intrusion to a state actor for any type of punishment-based deterrent action. In the kinetic dimension of the crisis, the ambiguity of

Pakistani claims of downed Indian aircraft, including French-built Rafales, early reporting of which was also contested and unverifiable, was observed, and is attributed to the sources of the claims in this paper but not adjudicated.

Meanwhile, the crisis provides at least a straw in the wind signpost to the complementary absorptive tendencies that are at work in institutions. In the aftermath of the operational engagement, RUSI's analysis suggests that India acted with good sense and calibrated its actions in the operation, as there seemed to be deliberate decisions to not exceed certain limits that were tested but not breached (Royal United Services Institute, 2025). This is similar to, but does not prove, the action of those coordination mechanisms outlined in the documents of each state, such as the civil-military coordination architecture mandated in Pakistan's CERT Rules, and the cross-domain coordination mandated in India's multi-domain doctrine. This paper is designed as a cautionary warning, this should not be interpreted as an assertion about the absence of any institutional design during the high speed crisis, but as an observation that it was not clearly absent. This would be the type of primary interview evidence that goes beyond this paper's documentation based approach.

6- Discussion

The paper's central theoretical shift is that the strategic impact of AI power in cyber pressure is not determined by the relative sophistication of the tech platforms involved, but instead mediated by the institutional ability to coordinate, integrate, respond and withstand the cyber pressure of each state. This is in line with Horowitz' (2010) general finding that the spread of military innovation is a function of organisational and financial capability, not that of the innovation itself, as well as with Nfuka and Rusu's (2019) finding that governance maturity rather than technical sophistication is the key determinant of cybersecurity outcomes in resource-constrained states. The present model speaks to the institutional insight in direct dialogue with cyber-deterrence theory, which has received comparatively limited attention so far; and to the mechanisms through which AI adds to the demand on institutional capacity.

The key theoretical take-away from the framework is that the focus on a comparative approach to strategic stability must shift from the offensive deployment of capabilities to

comparative analysis of institutions. The model is designed as a portable analytic tool to be applied, in principle, to other nuclear/near-nuclear rivalries in the region that are subject to similar pressure and have varying levels of institutional capability. Implication of the policy: If a state like Pakistan is competing with a bigger and better resourced opponent then the model implies that investment in closing the gaps in coordination and responsiveness identified above brings more stability to the system than investment on parity with an opponent's offensive capability using AI and cyber.

These conclusions are limited by three restrictions. The fact that the analysis is based on a single critical case, as well as the nature of the claims that are made in this analysis (mechanism plausibility and not statistical generalisation), is indicative of the fact that a larger comparative study across multiple rivalrous dyads would be needed to firmly establish the model's external validity. Second, a large portion of the evidence presented in the paper is open-source and contested, and the paper ignores actual performance to make claims of comparative institutional design, meaning that there may be a significant design-performance gap on either side. Third, the paper does not include primary interview data of the type that could explore directly the performance of institutions; this is another important area for further inquiry, along with the extension of the framework to other dyads and specific testing of the responsiveness indicator for which the analysis suggests it is the one most theoretically relevant and challenging to assess from documents only.

The results also have implications for the wider area of discussion raised above on the question of whether cyberspace challenges the conceptual framework of international order in general. Diffusion of offensive cyber capability doesn't change the balance of power between states in the aggregate, but the level of reliability of the institutional processes on which each state relies to transform raw capability into calibrated strategic behaviour (Kello 2017). On the contrary, while in the reading of Buchanan (2017) the dilemma is a structural feature of cyberspace, it is not in this reading, but rather a dynamic one that is exacerbated by the level and type of institutional capacity, a state with high coordination and integration capacity can in principle communicate more credibly a defensive posture than a state whose institutions fail to reliably distinguish, internally, between defensive hardening and offensive preparation. This indicates another path for future theoretical research: the institutional capacity can mean more than just a moderator of

deterrence stability, which could be a partial solution to the cybersecurity dilemma, a scenario that the present paper can only indicate but not develop.

6.1 Policy Recommendations

The analysis leads to four recommendations which can be extrapolated to the case beyond. First, the responsiveness indicator is particularly well suited for this, because it is the dimension that is most explicitly called for by operational acceleration, and it's the one that's least easily verifiable by doctrine. This is a reason for more frequent, realistic, crisis-tempo exercises to determine whether declared coordination arrangements can actually be used when the decision cycle is compressed, rather than relying on the presence of a coordinating body as an indicator of its functioning. Second, we need a statutory, not informal, basis for the coordination of civil and military affairs, as is the case with a cross-institutional, membership-determined council, because the informal approach is too susceptible to the kind of disruption a serious AI-enabled operation is intended to cause. Third, public-private mechanisms of coordination, in addition to official civil-military mechanisms, are important, because the resilience dimension is as much based on these as on the coordination between the state and its internal system; more than 80% of critical infrastructure in most states is privately owned and therefore requires coordination with the private sector. Fourth, and in keeping with the paper's central theoretical point, defence and technology spending focused on offensive or defensive AI-cyber capability will do less to buy deterrence stability per dollar spent than a more balanced approach to spending will do.

7- Conclusion

It has been suggested in this paper that, while offensive capability is a necessary precondition for strategic evolution of deterrence stability, it is not sufficient. Institutional capacity must also be taken into consideration. Drawing on the continued argument in the literature of techno-realism that the strategic effect of technology is moderated by the institutional and political context in which it is applied (Sukma, 2024), the paper identified three mechanisms through which AI usage for cyber pressure diminishes deterrence, namely: rapidisation of the use of technology, uncertainty of attribution, and ambiguity of signals, and modelled institutional capacity in four dimensions and six observable indicators as the moderating variable determining whether that

pressure is absorbed or amplified. What emerges is a new central question for policy makers and scholars: not whether the other side is building up its offensive capacity, but whether a state's system of institutions can coordinate, integrate, respond, and withstand at the speed of the other side's capacity. It presents a portable analytic framework based on the India-Pakistan rivalry and the crisis of May 2025, for a rapidly expanding set of dyads where the rhythms of institutional deterrence are being outpaced by the evolution of AI.

References

[1] Ashraf, M. (2024). Cyber deterrence: Challenges and strategic approaches. *IPRI Journal, 24*(1), 1-28. https://journal.ipripak.org/wp-content/uploads/2024/06/Article-4-IPRI-Journal-XXIV-I-Mubeen-Ashraf-1-28.pdf

[2] Beach, D., & Pedersen, R. B. (2013). *Process-tracing methods: Foundations and guidelines*. University of Michigan Press. https://www.researchgate.net/profile/Derek_Beach/publication/268072732_Process-Tracing_Methods_Foundations_and_Guidelines/links/5bf1a13492851c6b27c87aa4/Process-Tracing-Methods-Foundations-and-Guidelines.pdf

[3] Bhaskar, R. (1975). *A realist theory of science*. Leeds Books. https://uberty.org/wp-content/uploads/2015/09/Roy_Bhaskar_A_Realist_Theory_of_Science.pdf

[4] Buchanan, B. (2017). *The cybersecurity dilemma: Hacking, trust and fear between nations*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190665012.001.0001

[5] CYFIRMA. (2025). *Firewalls and frontlines: The India-Pakistan cyber battlefield crisis*. CYFIRMA Research. https://www.cyfirma.com/research/firewalls-and-frontlines-the-india-pakistan-cyber-battlefield-crisis/

[6] Evans, P. (1995). *Embedded autonomy: States and industrial transformation*. Princeton University Press.

[7] Government of India. (2013). *National cyber security policy 2013*. Ministry of Electronics and Information Technology. https://www.meity.gov.in/static/uploads/2024/02/National_cyber_security_policy-2013_0.pdf

[8] Government of India, Ministry of Defence. (2025). *Joint doctrine for multi-domain operations (JP 2.06)*. https://usiofindia.org/pdf/file_68c3c9dd2b33e.pdf

[9] Government of India, Ministry of Electronics and Information Technology. (2025). *India AI governance guidelines*. https://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc2025115685601.pdf

[10] Government of Pakistan. (2018). *Digital Pakistan policy*. Ministry of Information Technology and Telecommunication. https://moib.gov.pk/Downloads/Policy/DIGITAL_PAKISTAN_POLICY%2822-05-2018%29.pdf

[11] Government of Pakistan. (2021). *National cyber security policy 2021*. Ministry of Information Technology and Telecommunication. https://pkcert.gov.pk/uploads/2023/06/2021-National-Cyber-Security-Policy-Final.pdf

[12] Government of Pakistan. (2022). *National security policy of Pakistan 2022-2026*. https://nsd.gov.pk/SiteImage/Misc/files/NSP%20summary.pdf

[13] Government of Pakistan. (2023). *Prevention of Electronic Crimes Act (PECA) CERT Rules 2023* [Gazette notification]. https://pkcert.gov.pk/uploads/2023/10/GAZETTE-CERT-Rules-2023.pdf

[14] Government of Pakistan, Ministry of Information Technology and Telecommunication. (2025). *National AI policy*. https://moitt.gov.pk/SiteImage/Policies/National%20AI%20Policy.pdf

[15] Horowitz, M. C. (2010). *The diffusion of military power: Causes and consequences for international politics*. Princeton University Press.

[16] Jervis, R. (1978). Cooperation under the security dilemma. *World Politics, 30*(2), 167-214. https://doi.org/10.2307/2009958

[17] Jervis, R. (1979). Why nuclear superiority doesn't matter. *Political Science Quarterly, 94*(4), 617-633. https://doi.org/10.2307/2149629

[18] Johnson, J. (2022). Automating the OODA loop in the age of intelligent machines: Reaffirming the role of humans in command-and-control decision-making in the digital age. *Defence Studies, 22*(4), 1-21. https://doi.org/10.1080/14702436.2022.2102486

[19] Joint Air Power Competence Centre. (2022). *Speeding up the OODA loop with AI*. JAPCC Essays. https://www.japcc.org/essays/speeding-up-the-ooda-loop-with-ai/

[20] Kello, L. (2017). *The virtual weapon and international order*. Yale University Press. https://doi.org/10.2307/j.ctt1trkjd1

[21] Lindsay, J. R. (2015). Tipping the scales: The attribution problem and the feasibility of deterrence against cyberattack. *Journal of Cybersecurity, 1*(1), 53–67. https://doi.org/10.1093/cybsec/tyv003

[22] Nfuka, E. N., & Rusu, L. (2019). The effect of IT governance maturity on public sector performance in developing countries: Case of Tanzania. *Government Information Quarterly, 36*(1), 1–15. https://doi.org/10.1016/j.giq.2018.10.001

[23] Nye, J. S., Jr. (2017). Deterrence and dissuasion in cyberspace. *International Security, 41*(3), 44–71. https://doi.org/10.1162/ISEC_a_00266

[24] Observer Research Foundation. (2025). *Operation Sindoor and India's cyber threat landscape*. ORF Expert Speak. https://www.orfonline.org/expert-speak/operation-sindoor-and-india-s-cyber-threat-landscape

[25] Pant, H. V., & Patil, S. (Eds.). (2025). *In the aftermath of Operation Sindoor: Escalation, deterrence, and India-Pakistan strategic stability* (ORF Special Report No. 263). Observer Research Foundation. https://www.orfonline.org/research/in-the-aftermath-of-operation-sindoor-escalation-deterrence-and-india-pakistan-strategic-stability

[26] Royal United Services Institute. (2025). *Calibrated force: Operation Sindoor and the future of Indian deterrence*. RUSI Commentary. https://www.rusi.org/explore-our-research/publications/commentary/calibrated-force-operation-sindoor-and-future-indian-deterrence

[27] Schelling, T. C. (1966). *Arms and influence*. Yale University Press. https://www.scribd.com/document/810652850/Schelling-T-C-2020-Arms-and-Influence

[28] Scharre, P. (2021). Debunking the AI arms race theory. *Texas National Security Review, 4*(3), 121–132. https://doi.org/10.26153/tsw/13985

[29] Silobreaker. (2025). *Hacktivism in the India-Pakistan conflict: Analysis of recent campaigns*. Silobreaker Blog. https://www.silobreaker.com/blog/geopolitical/hacktivism-in-the-india-pakistan-conflict-analysis-of-recent-campaigns/

[30] Snyder, G. H. (1965). The balance of power and the balance of terror. In P. Seabury (Ed.), *The balance of power* (pp. 184–201). Chandler Publishing.

[31] Sukma, I. M. (2024). Techno-realism: Navigating new challenges in the contemporary role of technology in politics. *Security and Defence Quarterly, 46*(2), 24-46. https://securityanddefence.pl/pdf-188303-113014?filename=Techno-Realism--Navigatin.pdf

[32] Van Evera, S. (1997). *Guide to methods for students of political science*. Cornell University Press. https://is.muni.cz/el/fss/jaro2013/MVZ453/um/Van_Evera_-_Guide_to_Methods_for_Students_of_Political_Science.pdf

[33] Waltz, K. N. (1979). *Theory of international politics*. Addison-Wesley. http://slantchev.ucsd.edu/courses/ps240/03%20Anarchy,%20Hierarchy,%20and%20Sovereignty/Waltz%20-%20Theory%20of%20International%20Politics%20(Ch%205-8).pdf