ISSN: 2450-6869

eISSN:2719-6763

No. 23, 2026

DEFENCE SCIENCE REVIEW

http://www.journalssystem.com/pno/

DOI: 10.37055/pno/242205

Adversary emulation based on MITRE ATT&CK as a method of assessing defence readiness: a review of frameworks, evidence and limitations

Original article

Received: 2026-07-27

Final review: 2026-09-22

Accepted: 2026-10-27

Double blind

Keywords:

Adversary emulation, MITRE ATT&CK, defence readiness, threat-led penetration testing, cyber defence exercises.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License

Artur Chlebicki[1, A-B, D, F]

ORCID 0009-0009-7304-0138

A - Research concept and design, B - Collection and/or assembly of data,
C - Data analysis and interpretation, D - Writing the article, E - Critical revision of the article, F - Final approval of the article

[1]Independent Researcher, Poland

Abstract

Objectives: This review asks how adversary emulation grounded in the MITRE ATT&CK knowledge base differs from older forms of offensive testing, what frameworks turn it into a repeatable procedure, and what the published evidence says about its value and limits as a method for assessing defence readiness.

Methods: A narrative review was conducted of peer-reviewed literature publish between 2016 and 2026, together with framework and normative documents fr MITRE, the European Central Bank, ENISA, NIST and CIS No experiments, exercises or measurements of the author's own were carr out, and every quantitative statement is attributed to its source. The synthesis matrix the article is an expert assessment calibrated against the cited studi not a measurement.

Results: Adversary emulation moves the question of readiness from whether an intruder can find a way in to whether a defender can see and stop a named adversary's documented behaviours. ATT&CK supplies the shared vocabulary that makes this repeatable, and mature frameworks now operationalise it, from open emulation plans to threat-led penetration testing mandated under DORA and to military cyber-defence exercises. Published assessments show that capable products still miss steps of realistic attack chains, confirming its diagnostic value; the same literature shows that technique coverage is not the same as security and that the approach inherits the blind spots of what is known.

Conclusions: Adversary emulation is the best-structured method available for continuous readiness assessment, provided its results are read as a lower bound on capability against known behaviours rather than a certificate of resilience.

Corresponding author Artur Chlebicki Independent Researcher, Starowiejska 62a/190, 80-534, Gdańsk, Poland; email: arturichlebicki@gmail.com

1. Introduction

Defence readiness has traditionally been judged against documents. An organisation is asked whether it holds the required policies, whether its controls match a checklist, and whether an audit can be passed. An adversary does none of these things. An adversary performs actions on a network, and those actions either succeed or are seen and stopped. The gap between a passed audit and a defended network is the problem this review addresses, and adversary emulation is an approach developed to address this gap.

The idea is straightforward to state. Rather than probing a system for any exploitable weakness, the tester reproduces the specific, documented behaviours of a real adversary and observes whether the defence detects and responds to them. The behaviours are drawn from cyber threat intelligence and expressed in a shared vocabulary, so that a test is no longer a one-off demonstration of a single intruder's ingenuity but a repeatable measurement of coverage against a named threat. That shared vocabulary is, in almost all current work, the MITRE ATT&CK knowledge base, which catalogues the tactics and techniques that adversaries have been observed to use (Strom et al., 2020).

For a defence organisation the appeal is direct. Military and critical-infrastructure operators cannot certify readiness by compliance alone, because the threats they face are precisely those that compliance regimes lag behind. They need a way to ask, repeatedly and against a moving target, whether their people, sensors and procedures actually catch what a determined opponent would do. Adversary emulation promises exactly that, and the appearance of regulated threat-led penetration testing in the European financial sector shows that the promise is now being written into law (European Union, 2022).

This paper asks three review questions. First, how does adversary emulation differ from penetration testing and red teaming, and what does ATT&CK contribute that makes it repeatable (RQ1)? Second, which frameworks operationalise the method, from open emulation plans to regulated testing and military exercises (RQ2)? Third, what does the published evidence say about the method's value and its limits for assessing readiness (RQ3)? To answer them, a narrative review was conducted of peer-reviewed publications from 2016 to 2026, identified through Crossref, IEEE Xplore, the ACM Digital Library, SpringerLink and MDPI, together with framework and normative documents from MITRE, the European Central Bank, ENISA, NIST and CISA. Sources were included when they addressed adversary emulation, the ATT&CK knowledge base, or the testing and exercise practices that surround them. This is a secondary study: it presents no exercises or measurements of the author's own, every quantitative claim is attributed to its source, and any claim that could not be verified against an accessible source was excluded.

The remainder of the paper distinguishes emulation from the older testing methods (Section 2), examines ATT&CK as the knowledge base that structures it (Section 3), surveys the frameworks that operationalise it (Section 4), weighs the published evidence and the limits of the method (Section 5), draws out the implications for defence readiness assessment (Section 6), and concludes (Section 7).

2. From penetration testing to adversary emulation

Offensive security testing is not one activity but a family of them, and the differences matter for what each can tell an assessor about readiness. The oldest member of the family is the penetration test. Its goal is to find a way in: the tester searches for exploitable weaknesses and demonstrates impact by chaining them into a successful intrusion. The established methodological guidance treats the exercise as a structured search for vulnerabilities and their exploitation (Scarfone et al., 2008), and the research literature has mapped its workflow and open problems in detail (Dalalana Bertoglio & Zorzo, 2017). A penetration test answers a real question, but a narrow one: does at least one path exist? It says little about whether the defence would notice an intruder who took a different path.

Red teaming widens the question. A red team pursues a defined objective, such as reaching a particular system or dataset, using whatever combination of technical, physical and social means an adversary might, and in doing so it tests the defence as a whole rather than a single control. The realism is greater, but the exercise is a manual effort whose course and findings depend on the skill and choices of the team performing it, and two red teams pursuing the same objective may exercise entirely different parts of the defence (Applebaum et al., 2016).

Adversary emulation makes a further, decisive move. Instead of pursuing an objective by any means, the tester reproduces the documented behaviours of a specific, named adversary, in the sequence and with the techniques that threat intelligence attributes to it. The point is no longer to prove that an intrusion is possible but to measure whether the defence covers the particular tactics and techniques that this adversary is known to use. The approach was crystallised in early work on automated red team emulation, which showed that an adversary's behaviour could be encoded as a model and executed automatically against a target environment (Applebaum et al., 2016). Because the behaviours come from intelligence rather than from the tester's imagination, two emulations of the same adversary are comparable, and a result becomes a coverage statement rather than an anecdote.

A related practice, purple teaming, keeps the emulation but removes the secrecy: the offensive and defensive sides work together, so that each emulated technique is immediately checked against the defence's ability to see it, and gaps are closed as they are found. This collaborative variant has since been written into regulated testing as a mandatory phase (European Central Bank, 2025). At the far end of automation, breach and attack simulation turns emulation into a continuous, tool-driven process that replays attack chains on a schedule rather than as a single engagement, extending coverage to multi-step chains and exploit sequences (Chen et al., 2023). Table 1 sets these methods against each other on the dimensions that matter for readiness assessment.

Table 1. Forms of offensive security testing compared (compiled from the cited sources).

Method	Guiding goal	Adversary realism	Role of TTP knowledge	Result	Typical cadence
Penetration testing	Find any exploitable path in	Tester-defined	Low; ad hoc	Proof that a path exists (Dalalana Bertoglio & Zorzo, 2017; Scarfone et al., 2008)	One-off
Red teaming	Reach a defined objective by any means	Objective-driven	Moderate	Whole-of-defence test against one goal (Applebaum et al., 2016)	Occasional
Adversary emulation	Reproduce a named adversary's documented behaviours	High; intelligence-led	Central; ATT&CK-structured	Coverage against specific TTPs (Applebaum et al., 2016)	Repeatable
Purple teaming	Emulate while co-operating with defenders	High	Central	Gaps found and closed in session (European Central Bank, 2025)	Iterative
Breach and attack simulation	Automate and repeat emulation	Tool-encoded	Central	Continuous coverage of attack chains (Chen et al., 2023)	Continuous

Source: own elaboration based on the cited sources.

The progression down the table is a progression in what an assessor can claim. A penetration test yields a binary existence proof; emulation yields a structured statement about which adversary behaviours are and are not covered; and that statement is only repeatable and comparable because the behaviours are named in a common language. That language is the subject of the next section.

3. ATT&CK as the structuring knowledge base

MITRE ATT&CK is a curated knowledge base of adversary behaviour, organised as a matrix of tactics, the adversary's short-term goals, and the techniques and sub-techniques used to achieve them, cross-referenced to the threat groups, campaigns and software in which each behaviour has been observed (Strom et al., 2020). It is not a maturity model or a control catalogue; it is a description of what adversaries do, derived from public reporting and revised as reporting accumulates. The knowledge base is versioned so that assessments can be pinned to a fixed reference: the enterprise matrix in the current release contains fourteen tactics and several hundred techniques and sub-techniques, and successive major versions add, merge

and deprecate entries as the evidence changes (MITRE Corporation, 2026). This versioning is what allows an emulation performed today to be compared with one performed a year ago.

The knowledge base earns its central place by being useful for more than emulation. A substantial research literature has grown around extracting ATT&CK techniques from unstructured threat reports, so that the raw material of emulation can be produced at scale; TTPDrill demonstrated the automatic extraction of threat actions from prose intelligence and their mapping to a structured representation (Husari et al., 2017). Others have studied the statistical structure of the matrix itself, learning which techniques tend to co-occur so that partial observations can be extended into likely chains (Al-Shaer et al., 2020). ATT&CK has been used as the backbone of formal threat modelling, attaching attack behaviours to system architectures to reason about exposure (Xiong et al., 2022), and as a bridge between offensive knowledge and defensive frameworks, mapping techniques onto control catalogues such as the NIST cybersecurity framework (Kwon et al., 2020). Empirical studies have characterised how the knowledge base is populated and where its coverage is uneven, whether by analysing techniques and trends in real malware (Oosthoek & Doerr, 2019) or by examining the database as a whole to surface the threats it does and does not represent well (Al-Sada et al., 2024).

Placing ATT&CK beside the models it succeeded clarifies what it added. The Lockheed Martin cyber kill chain describes intrusion as a linear sequence of phases, and the Diamond Model relates adversary, capability, infrastructure and victim; both are analytical lenses rather than behavioural catalogues. A direct comparison shows ATT&CK to be the most granular and operational of the three, precisely because it enumerates concrete techniques rather than abstract stages, which is what an emulation needs to reproduce (Naik et al., 2022). That granularity is also a hazard. Mapping observed activity to techniques is a matter of judgement, and inconsistent or over-confident mapping degrades everything built on top of it, which is why national guidance now sets out best practices and common mistakes for ATT&CK mapping (Cybersecurity and Infrastructure Security Agency [CISA], 2023). The knowledge base makes emulation repeatable, but only to the extent that the mapping into it is disciplined.

4. Frameworks that operationalise emulation

A knowledge base is not a procedure. Turning ATT&CK into a readiness assessment requires an emulation plan that says which techniques to execute, in what order, and how to judge the result, and several distinct frameworks now supply exactly this at different levels of formality.

The most direct are published emulation plans. The Center for Threat-Informed Defense maintains an open library of adversary emulation plans that translate the documented behaviour of specific threat actors into executable sequences, offering both full plans that emulate an adversary from initial access to exfiltration and micro plans that focus on a single compound behaviour seen across several actors (Center for Threat-Informed Defense, 2020). Because every step is annotated back to ATT&CK and to the public intelligence it rests on, such a plan is simultaneously a test script and an audit trail, which is what makes its results defensible. Automation frameworks then execute these plans with less human effort: the early work on automated emulation showed the principle (Applebaum et al., 2016), and later systems have

pushed adversary emulation towards proactive, repeatable threat hunting (Ajmal et al., 2021), where the techniques executed during an emulation double as the hypotheses a hunt sets out to confirm, an equivalence made explicit in work that carries defenders from an adversary's tactics and techniques through to the indicators they leave behind (Berady et al., 2021). Breach and attack simulation tools drive the same idea to a continuous cadence (Chen et al., 2023).

At the most formal end sits regulated threat-led penetration testing. The European framework for threat intelligence-based ethical red-teaming, TIBER-EU, was updated in 2025 to align with the regulatory technical standards of the Digital Operational Resilience Act, and in the process made purple teaming a mandatory phase and standardised the roles and terminology of a controlled test (European Central Bank, 2025).

The underlying regulation requires significant financial entities to undergo advanced resilience testing by means of threat-led penetration testing on a defined cycle (Regulation (EU) 2022/2554, 2022). This matters beyond finance because it is among the first instances of intelligence-led emulation being made a binding legal obligation across a whole sector, with a prescribed methodology, and it provides a template that defence and critical-infrastructure operators can adopt: a controlled, intelligence-driven emulation conducted against production systems under agreed rules.

Between the open plan and the regulated test lies the exercise, which is the form most familiar to defence organisations. Cyber-defence exercises give structure to planning, execution and evaluation of a simulated engagement (Seker & Ozbenli, 2018), and they are increasingly hosted on cyber ranges, instrumented environments whose scenarios, functions and architecture have been surveyed in detail (Yamin et al., 2020). Complex hands-on defence exercises in a range yield practical lessons about what such events can and cannot teach (Vykopal et al., 2017), and dedicated frameworks now exist to develop and assess the competences of the people taking part, so that an exercise measures human readiness and not only tool coverage (Brilingaitė et al., 2020). When the scenario driving one of these exercises is an ATT&CK-based emulation plan, the exercise becomes a readiness assessment in the sense this review is concerned with. Figure 1 places these frameworks on a single pipeline, from intelligence to a readiness decision.

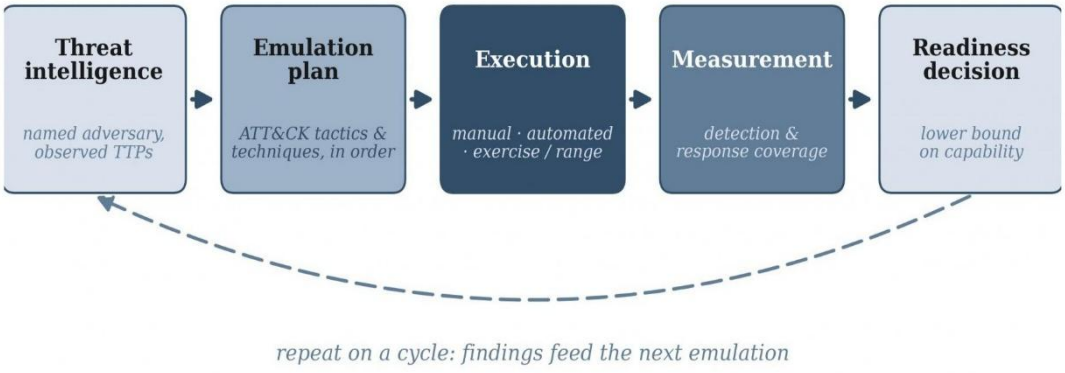

Fig. 1. Adversary emulation as a readiness-assessment pipeline.

Source: own elaboration based on Applebaum et al. (2016), Strom et al. (2020), European Central Bank (2025) and Seker and Ozbenli (2018).

5. Evidence and limitations

The case for emulation rests ultimately on what it reveals, and the published record shows that it reveals a great deal. Controlled assessments of endpoint detection and response products against the attack vectors of advanced persistent threats have found that capable, widely deployed tools miss or fail to alert on steps of realistic chains, so that a defence assumed to be adequate is shown not to be (Karantzas & Patsakis, 2021). Results of this kind are the strongest argument for the method: they translate an abstract worry into a specific, reproducible gap. The public ATT&CK Evaluations extend the same logic to a whole market, emulating a named adversary against many products under identical conditions; the seventh enterprise round, released at the end of 2025, emulated behaviours inspired by the Scattered Spider and Mustang Panda actors across eleven participating vendors and, for the first time, exercised a cloud-based scenario (MITRE Corporation, 2025). The evaluations are deliberately published without an overall ranking, and this restraint is itself instructive: the data describe coverage against one emulated adversary, not a general verdict of security. Independent research has shown how the underlying evaluation data can nonetheless be put to work, using it to reason quantitatively about defender policy and the allocation of defensive resources (Outkin et al., 2023).

These strengths come with limits that the same literature makes plain, and honesty about them is essential to using the method well. The first and most fundamental limitation is the reliance on previously documented adversary behaviours.. Emulation reproduces documented behaviour, and ATT&CK is a catalogue of what has been observed and reported; a genuinely novel technique is by definition absent from the plan, and studies of how the knowledge base is populated confirm that its coverage is uneven across platforms and adversary types (Al-Sada et al., 2024; Oosthoek & Doerr, 2019). A defence can therefore score well against every emulated technique and still fall to something the catalogue has not yet named. The second limit follows from the first: technique coverage is not the same as security. A matrix of green cells measures how many known behaviours are detected, not how resilient the organisation is, and treating coverage as a target invites the familiar failure of optimising the metric rather than the outcome. The third is that emulation captures technology far better than it captures people and process; the human and organisational dimension of readiness, from security culture to the competence of responders, needs its own instruments, whether a cyber-security culture framework applied to ATT&CK risk (Georgiadou et al., 2021) or the competence-assessment frameworks built for exercises (Brilingaitė et al., 2020). The fourth is practical: a high-fidelity emulation demands current intelligence, a skilled team and careful control of a live environment, and its result is valid only for the configuration and the moment in which it was run. The broader threat picture, in which intrusion vectors and dominant actors shift from year to year (European Union Agency for Cybersecurity [ENISA], 2025), guarantees that a readiness assessment loses validity over time and must be updated. Figure 2 summarises how well the main variants of emulation answer the distinct questions a readiness assessment must ask; its cells are an expert assessment calibrated against the studies cited in this section, not the result of measurement, and should be read as such.

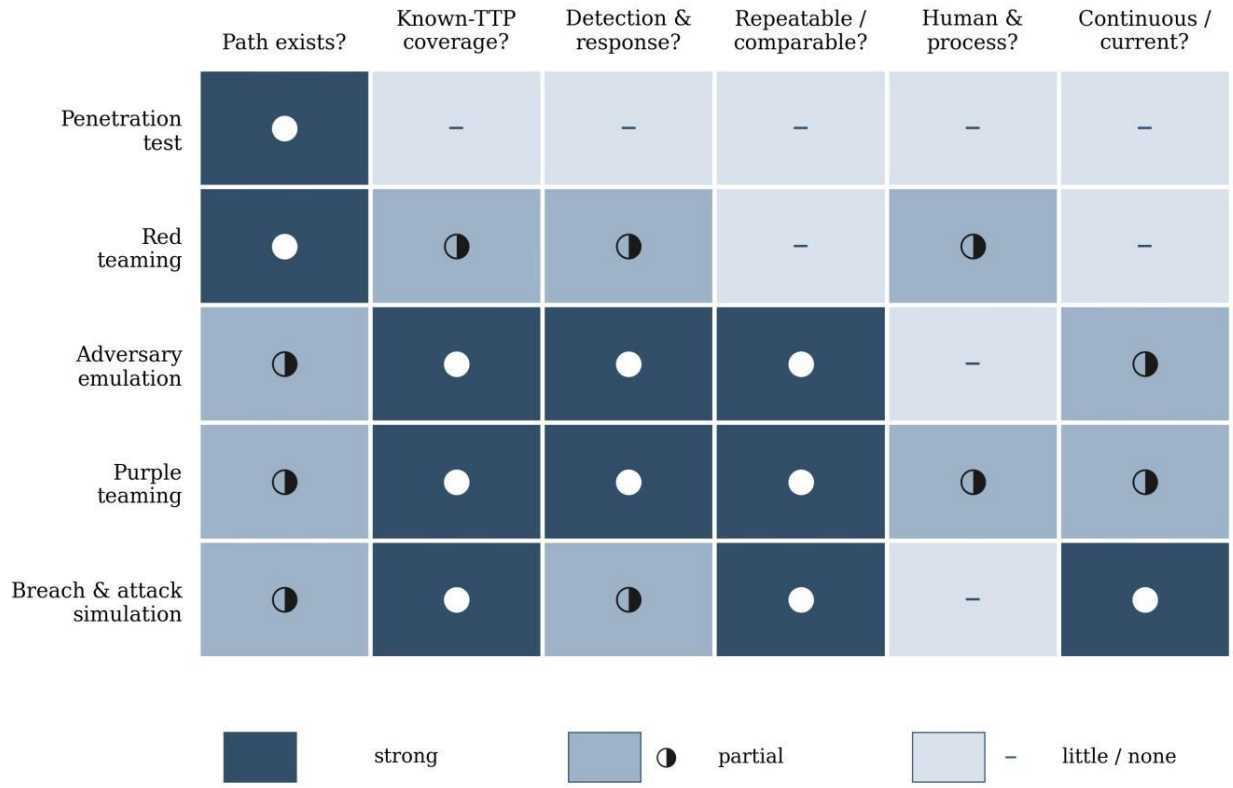

Fig. 2. Coverage of readiness questions by forms of offensive testing.
Values are an expert assessment calibrated against the cited studies, not a measurement.

Source: own elaboration (expert assessment) based on Scarfone et al. (2008), Dalalana Bertoglio and Zorzo (2017), Applebaum et al. (2016), Chen et al. (2023), European Central Bank (2025) and Karantzas and Patsakis (2021).

6. Implications for defence readiness assessment

Read together, the frameworks and the evidence point to a change in what readiness assessment can be. Compliance asks whether the required controls are present; emulation asks whether they work against a named opponent, and it asks the question in a form that can be repeated and compared over time. For a defence organisation this is the difference between certifying a state and measuring a capability. The unit that passes an inspection has demonstrated conformance; the unit whose sensors and responders catch an emulated adversary has demonstrated that, against that adversary on that day, it can effectively detect and respond to the emulated adversary.

The practical route to this capability is already visible in the frameworks surveyed above. An ATT&CK-based emulation plan can be embedded as the scenario in a cyber-defence exercise, so that a recurring exercise doubles as a recurring readiness measurement rather than a training event alone (Seker & Ozbenli, 2018; Vykopal et al., 2017). The regulated model from the financial sector supplies a template for higher-stakes systems: an intelligence-led, controlled

emulation conducted against production infrastructure on a fixed cycle, with a mandatory collaborative phase so that findings are remediated and not merely recorded (European Central Bank, 2025; Regulation (EU) 2022/2554, 2022). Adopting either form has preconditions that the literature is candid about. It requires threat intelligence current enough to make the emulated adversary representative, a team competent to execute and interpret the plan, and enough control over the target environment to run the test safely, together with instruments for the human dimension that the technical emulation does not reach (Brilingaitė et al., 2020; Georgiadou et al., 2021). Where those conditions can be met, emulation gives a defence organisation something compliance cannot: a repeatable, adversary-referenced measure of whether it is ready. The open research questions concern coverage of the parts the method reaches least well, the novel and the human, and the design of metrics that are resistant to manipulation.

7. Conclusions

This review set out to distinguish adversary emulation from older testing, to survey the frameworks that operationalise it, and to weigh its value and limits for assessing defence readiness. In answer to the first question, emulation differs from penetration testing and red teaming by reproducing the documented behaviours of a specific adversary rather than searching for any path or pursuing any objective, and it is ATT&CK that makes those behaviours nameable, repeatable and comparable. In answer to the second, the method is now operationalised across a spectrum that runs from open emulation plans and automation tools, through military cyber-defence exercises on instrumented ranges, to legally mandated threat-led penetration testing with a prescribed methodology. In answer to the third, the published evidence shows that emulation reliably exposes real gaps in otherwise capable defences, which is its central value, while the same literature warns that it measures only the known, that coverage is not security, and that its results are perishable.

The implications for operators of defence and critical systems should therefore be interpreted with caution. Adversary emulation is the best-structured method available for assessing readiness continuously and against a realistic opponent, and it deserves the place it is being given in exercises and regulation. But its output is a lower bound on capability against known behaviours, not a certificate of resilience, and it is most valuable when it is read that way: as a repeated, honest measure of what a defence can currently catch, and a standing reminder of how much it has not yet been tested against.

Acknowledgements

The author thanks the maintainers of the open ATT&CK knowledge base and emulation libraries, whose public availability made a source-verifiable review of this subject possible.

Funding

This research received no external funding.

Disclosure of AI use and conflicts of interest

The author confirms that no generative artificial intelligence (AI) tools were used in the preparation, drafting, editing, analysis, or review of this work. All content was created

and verified exclusivelyy by the author, who takes full responsibility for its accuracy and integrity. The author declares no conflict of interest.

References

Ajmal, A. B., Shah, M. A., Maple, C., Asghar, M. N., & Islam, S. U. (2021). Offensive security: Towards proactive threat hunting via adversary emulation. *IEEE Access, 9*, 126023–126033. https://doi.org/10.1109/ACCESS.2021.3104260

Al-Sada, B., Sadighian, A., & Oligeri, G. (2024). Analysis and characterization of cyber threats leveraging the MITRE ATT&CK database. *IEEE Access, 12*, 1217–1234. https://doi.org/10.1109/ACCESS.2023.3344680

Al-Shaer, R., Spring, J. M., & Christou, E. (2020). Learning the associations of MITRE ATT&CK adversarial techniques. In *2020 IEEE Conference on Communications and Network Security (CNS)* (pp. 1–9). IEEE. https://doi.org/10.1109/CNS48642.2020.9162207

Applebaum, A., Miller, D., Strom, B., Korban, C., & Wolf, R. (2016). Intelligent, automated red team emulation. In *Proceedings of the 32nd Annual Conference on Computer Security Applications* (pp. 363–373). Association for Computing Machinery. https://doi.org/10.1145/2991079.2991111

Berady, A., Jaume, M., Viet Triem Tong, V., & Guette, G. (2021). From TTP to IoC: Advanced persistent graphs for threat hunting. *IEEE Transactions on Network and Service Management, 18*(2), 1321–1333. https://doi.org/10.1109/TNSM.2021.3056999

Brilingaitė, A., Bukauskas, L., & Juozapavičius, A. (2020). A framework for competence development and assessment in hybrid cybersecurity exercises. *Computers & Security, 88*, Article 101607. https://doi.org/10.1016/j.cose.2019.101607

Chen, J., Liang, R., Zhang, M., Zheng, C., Huang, X., Lu, H., Yu, X., & Tian, Z. (2023). Vulnerability correlation, multi-step attack and exploit chain in breach and attack simulation. In *2023 IEEE 12th International Conference on Cloud Networking (CloudNet)* (pp. 398–402). IEEE. https://doi.org/10.1109/CloudNet59005.2023.10490046

Dalalana Bertoglio, D., & Zorzo, A. F. (2017). Overview and open issues on penetration test. *Journal of the Brazilian Computer Society, 23*(1), Article 2. https://doi.org/10.1186/s13173-017-0051-1

Georgiadou, A., Mouzakitis, S., & Askounis, D. (2021). Assessing MITRE ATT&CK risk using a cyber-security culture framework. *Sensors, 21*(9), Article 3267. https://doi.org/10.3390/s21093267

Husari, G., Al-Shaer, E., Ahmed, M., Chu, B., & Niu, X. (2017). TTPDrill: Automatic and accurate extraction of threat actions from unstructured text of CTI sources. In *Proceedings of the 33rd Annual Computer Security Applications Conference* (pp. 103–115). Association for Computing Machinery. https://doi.org/10.1145/3134600.3134646

Karantzas, G., & Patsakis, C. (2021). An empirical assessment of endpoint detection and response systems against advanced persistent threats attack vectors. *Journal of Cybersecurity and Privacy, 1*(3), 387–421. https://doi.org/10.3390/jcp1030021

Kwon, R., Ashley, T. D., Castleberry, J. E., McKenzie, P. L., & Gourisetti, S. N. G. (2020). Cyber threat dictionary using MITRE ATT&CK matrix and NIST cybersecurity framework mapping. In *2020 Resilience Week (RWS)* (pp. 106–112). IEEE. https://doi.org/10.1109/RWS50334.2020.9241271

Naik, N., Jenkins, P., Grace, P., & Song, J. (2022). Comparing attack models for IT systems: Lockheed Martin's cyber kill chain, MITRE ATT&CK framework and Diamond Model. In *2022 IEEE*

International Symposium on Systems Engineering (ISSE) (pp. 1–7). IEEE. https://doi.org/10.1109/ISSE54508.2022.10005490

Oosthoek, K., & Doerr, C. (2019). SoK: ATT&CK techniques and trends in Windows malware. In S. Chen, K.-K. R. Choo, X. Fu, W. Lou, & A. Mohaisen (Eds.), *Security and privacy in communication networks: 15th EAI International Conference, SecureComm 2019, proceedings, part I* (Vol. 304, pp. 406–425). Springer. https://doi.org/10.1007/978-3-030-37228-6_20

Outkin, A. V., Schulz, P. V., Schulz, T., Tarman, T. D., & Pinar, A. (2023). Defender policy evaluation and resource allocation with MITRE ATT&CK evaluations data. *IEEE Transactions on Dependable and Secure Computing, 20*(3), 1909–1926. https://doi.org/10.1109/TDSC.2022.3165624

Seker, E., & Ozbenli, H. H. (2018). The concept of cyber defence exercises (CDX): Planning, execution, evaluation. In *2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (pp. 1–9). IEEE. https://doi.org/10.1109/CyberSecPODS.2018.8560673

Vykopal, J., Vizváry, M., Ošlejšek, R., Čeleda, P., & Tovarňák, D. (2017). Lessons learned from complex hands-on defence exercises in a cyber range. In *2017 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). IEEE. https://doi.org/10.1109/FIE.2017.8190713

Xiong, W., Legrand, E., Åberg, O., & Lagerström, R. (2022). Cyber security threat modeling based on the MITRE Enterprise ATT&CK matrix. *Software and Systems Modeling, 21*(1), 157–177. https://doi.org/10.1007/s10270-021-00898-7

Yamin, M. M., Katt, B., & Gkioulos, V. (2020). Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. *Computers & Security, 88*, Article 101636. https://doi.org/10.1016/j.cose.2019.101636

Other sources

Center for Threat-Informed Defense. (2020, September 10). *Adversary emulation library*. Retrieved July 26, 2026, from https://ctid.mitre.org/resources/adversary-emulation-library/

Cybersecurity and Infrastructure Security Agency. (2023, January). *Best practices for MITRE ATT&CK mapping*. https://www.cisa.gov/sites/default/files/2023-01/Best%20Practices%20for%20MITRE%20ATTCK%20Mapping.pdf

European Central Bank. (2025, February 11). *TIBER-EU framework updated to align with DORA*. https://www.ecb.europa.eu/press/intro/news/html/ecb.mipnews250211.en.html

European Union Agency for Cybersecurity. (2025). *ENISA threat landscape 2025*. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025

MITRE Corporation. (2025, December 10). *MITRE ATT&CK evaluations advance cloud security and counter-espionage capabilities in latest round* [Press release]. https://www.mitre.org/news-insights/news-release/mitre-attck-evaluations-advance-cloud-security-and-counter-espionage

MITRE Corporation. (2026, May 12). *Enterprise ATT&CK* (Version 19.1) [Data set]. https://raw.githubusercontent.com/mitre-attack/attack-stix-data/master/enterprise-attack/enterprise-attack-19.1.json

Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011,

Official Journal of the European Union, L 333, 1–79 (2022). https://eur-lex.europa.eu/eli/reg/2022/2554/oj

Scarfone, K., Souppaya, M., Cody, A., & Orebaugh, A. (2008). *Technical guide to information security testing and assessment* (NIST Special Publication 800-115). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-115

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2020). *MITRE ATT&CK: Design and philosophy* (Report No. MP180360R1). The MITRE Corporation. https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf