Perspective | Open Access

Intelligence & Robotics

Fu *et al. Intell. Robot.* **2026**, *6*(4), 1030-40 **DOI:**10.20517/ir.2026.44

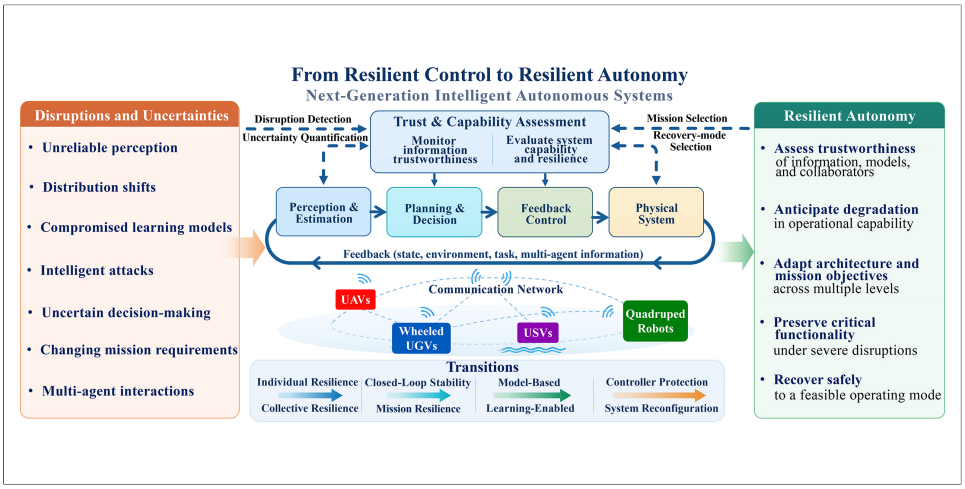
Resilient control for the next generation of intelligent autonomous systems

Chengcheng Fu, Lu Liu, Gang Feng

Citation: Fu, C.; Liu, L.; Feng, G. Resilient control for the next generation of intelligent autonomous systems. *Intell. Robot.* **2026**, *6*(4), 1030-40. https://dx.doi.org/10.20517/ir.2026.44

Received: 10 Oct 2026 **First Decision:** 28 Oct 2026 **Revised:** 11 Nov 2026 **Accepted:** 13 Nov 2026 **Published:** 23 Nov 2026

Academic Editor: Hongtian Chen **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang



1. INTRODUCTION

Intelligent autonomous systems are increasingly deployed in transportation, robotics, manufacturing, aerospace, marine systems, energy, and critical infrastructure. Autonomous vehicles, unmanned aerial vehicles (UAVs), unmanned marine vehicles (UMVs), mobile robots, and robotic networks integrate physical dynamics with sensing, communication, computation, learning, planning, and feedback control. Intelligence is therefore becoming an integral part of closed-loop operation.

This integration improves autonomous capability while introducing new sources of uncertainty and vulnerability. Perception errors influence estimation and planning; inappropriate decisions produce undesirable control actions; and continual learning may incorporate information acquired under abnormal conditions.

Robust and fault-tolerant control provide the classical foundations for maintaining closed-loop performance under uncertainty and component failures[1-3]. Networked feedback introduced additional cyber-physical vulnerabilities, motivating secure control[4], attack detection and identification[5], secure state estimation[6], adversarial-control analysis[7], and control under denial-of-service (DoS) attacks[8]. More recent

Check for updates

Department of Mechanical Engineering, City University of Hong Kong, Hong Kong 999077, China.

Correspondence to: Prof. Lu Liu, Department of Mechanical Engineering, City University of Hong Kong, Hong Kong 999077, China. E-mail: luliu45@cityu.edu.hk

work has extended these foundations toward broader notions of cyber-resilience[9].

For intelligent autonomous systems, resilience may be defined as the capability to anticipate, withstand, adapt to, and recover from adverse events while preserving safety and essential mission functionality. In this article, secure control emphasizes protection against malicious interference, while resilient control addresses maintaining and recovering acceptable closed-loop behavior under disruptions. Mission resilience focuses on preserving essential task functionality or safely adapting mission objectives. Resilient autonomy integrates these concerns across perception, learning, decision-making, cooperation, and control, including reconfiguration and safe recovery.

Resilient autonomy is therefore a key capability of next-generation intelligent autonomous systems. Four complementary transitions characterize this development: from individual to collective resilience; from closed-loop stability to mission resilience; from model-based to learning-enabled resilience; and from controller protection to autonomous system reconfiguration. Existing surveys synthesize advances in cyber-resilience[9], resilient consensus[10], and safe learning in robotics[11]. This article contributes an integrated framework that organizes the four transitions around information trustworthiness, residual operational capability, and mission feasibility. The framework highlights the interdependencies among the four transitions and connects their established foundations to a common question: which safety and mission requirements remain achievable when disruptions change both available resources and the assumptions supporting autonomous decisions? Guided by this question, we review the state of the art, outline future research directions, and discuss open challenges for resilient autonomy.

2. EVOLUTION AND STATE OF THE ART

2.1. Secure estimation and resilient feedback control

Reliable information is fundamental to feedback control. Research on replay attacks, attack-resilient state estimation, and secure state reconstruction established that corrupted information can compromise closed-loop operation even when the nominal controller satisfies standard stability requirements[12-15].

Adversarial uncertainty differs from conventional uncertainty because an attacker may exploit knowledge of plant dynamics, estimation algorithms, controllers, and communication architectures. Resilient-control research has also considered false-data-injection, DoS, and deception attacks[16-18]. However, these studies rely on specific assumptions, such as sufficient information redundancy or restrictions on attack capabilities. Detecting inconsistent measurements does not by itself establish accurate state recovery.

2.2. From individual to collective resilience

Research on resilient distributed systems has established algorithms tolerant of malicious agents, resilient consensus, and resilient operation under DoS attacks[10,19-23]. System architecture is becoming part of the resilience mechanism. Adaptive and activatable communication layers can provide alternative resources after network disruption[24].

Consensus ensures agreement under specified conditions but does not necessarily guarantee the accuracy of shared observations or successful task execution. Excluding a suspicious agent also requires checking whether the remaining sensing and communication resources are sufficient.

2.3. From closed-loop stability to mission resilience

Closed-loop stability alone is insufficient to characterize the resilience of autonomous systems. A study of vehicle platoons[25] further illustrates this distinction. In the investigated scenarios, external disturbances and sensor faults produced steady-state spacing errors without destroying string stability.

Resilience assessment should therefore account for performance degradation, residual capability, and recovery in addition to stability. Recoverability and durability have been introduced as resilience-related properties[26]. Temporal logic provides a language for specifying acceptable behavior, while robustness measures quantify the degree to which a trajectory satisfies a specification[27].

2.4. Time-critical and data-driven resilient cooperative control

Recent resilient cooperative-control research exhibits three main trends: explicit treatment of convergence time, reduced model dependence, and integration of cyber resilience with fault tolerance and learning. Data-driven approaches reduce reliance on accurate plant models[28,29]. Resilient formation control addresses mission-dependent coordination under DoS attacks[30] and Byzantine attacks[31]. Byzantine-resilient output optimization has also been studied[32]. Finite-time control of UMVs[33], finite-time cooperative regulation[34] and fixed-time approaches[35] provide increasingly explicit convergence-time guarantees.

Although finite- and fixed-time control methods provide convergence-time guarantees under their respective assumptions, these guarantees alone do not establish collision avoidance, input feasibility, or safe mission recovery.

2.5. Learning-enabled resilience of autonomous systems

Adaptive dynamic programming and reinforcement learning provide foundations for learning feedback policies[36]. Their integration with resilient control includes model-free control under DoS attacks, multi-agent reinforcement learning, and learning-based fault-tolerant control[37-41].

The principal unresolved issue concerns assurance. Learned perception may deteriorate under distribution shift, learned models may incorporate corrupted data, and learned policies may produce unsafe actions outside their training domains. Resilient control must therefore account for uncertainty associated with both physical dynamics and autonomous intelligence.

2.6. From controller protection to autonomous system reconfiguration

Resilient-control methods have been investigated for a range of autonomous platforms[42-48]. Digital twins offer complementary tools for monitoring and virtual testing[49], while intelligent robotics increasingly integrates perception, planning, and control[50].

A controller may tolerate an attack while perception becomes unreliable, or a vehicle may remain dynamically stable while its planner generates a plan that violates feasibility or safety constraints. Resilience mechanisms must therefore span the complete autonomous-system architecture and support its reconfiguration when necessary.

Table 1 links the aforementioned transitions to their established foundations, remaining gaps, required capabilities, and future research directions.

3. DIRECTIONS FOR NEXT-GENERATION RESILIENT CONTROL

3.1. End-to-end resilient autonomy

Future theoretical frameworks should treat the perception–decision–control loop as an integrated resilient system. A useful concept is a resilience envelope, which describes the degradation conditions under which specified safety and mission requirements remain achievable, subject to explicit information, threat, and resource assumptions. Nominal, degraded, and stopping modes have separate mission specifications under common mandatory safety constraints. For each predefined mission mode m, prescribed initial information, resource limits, and a finite horizon T, the resilience envelope can be expressed conceptually as

Table 1. Four complementary transitions toward resilient autonomy: established foundations, remaining gaps, required capabilities, and future research directions

Transition	Established foundations and remaining gap	Required capabilities	Future research directions
From individual to collective resilience	Resilient consensus, activatable network layers, and Byzantine-resilient formation tracking support cooperation under specified threats[10,20-22,24,31]. Agreement alone does not establish trustworthy collective information or mission success	Assess shared information; exploit effective diversity; coordinate sensing, task allocation, and control; preserve essential collective functionality	How can heterogeneous agents combine local trust assessment and resource coordination to maintain collective functionality despite compromised agents and common-mode failures?
From closed-loop stability to mission resilience	Temporal-logic formulations and finite- and fixed-time output regulation extend stability-based analysis[26,34,35]. The remaining issue is to connect these guarantees with residual resources and feasible degraded missions	Quantify residual mission capability; preserve safety-critical functionality; select feasible degraded objectives; recover safely within feasible time bounds	Which combinations of degradation, uncertainty, and residual resources permit safe mission continuation, timely recovery, or controlled termination?
From model-based to learning-enabled resilience	Data-driven regulation and learning-based fault-tolerant control reduce reliance on exact models[28,39-41]. Their guarantees remain conditional; corrupted data and distribution shifts require additional assurance	Assess information and policy reliability; anticipate degradation; regulate learned-policy authority; adapt without invalidating safety guarantees	How can learning, uncertainty assessment, safety supervision, and fallback control maintain compatible guarantees under corrupted data, distribution shifts, and continual adaptation?
From controller protection to autonomous system reconfiguration	Resilient control for a range of autonomous platforms[42-46]; digital twins support monitoring and virtual testing[49]. Coordinated reconfiguration across autonomy layers still requires system-level assurance	Reconfigure trusted resources and operating modes; adjust mission objectives; preserve critical functions and support safe recovery and control handover	When should architecture or mission reconfiguration be triggered, and how can cross-layer changes be verified under constraints on sensing, communication, computation, and actuation?

$$\mathcal{E}_m = \left\{ \delta \;\middle|\; \exists\, \pi \in \Pi_{\mathrm{adm}},\ \xi_{[0,T]}^{\pi,\omega} \in \mathcal{T}_{\mathrm{S},m}, \quad \forall\, \omega \in \Omega(\delta) \right\}, \tag{1}$$

where δ is a vector of degradation bounds, and $\omega \in \Omega(\delta)$ specifies compatible initial uncertainty, disturbances, attacks, and faults. The policy $\pi \in \Pi_{\mathrm{adm}}$ selects control inputs and reconfiguration actions using only currently available and past information, subject to the prescribed constraints. The same causal policy must satisfy the requirements for every admissible scenario. Here, $\xi_{[0,T]}^{\pi,\omega}$ denotes the closed-loop trajectory generated by policy π under scenario ω, and $\mathcal{T}_{\mathrm{S},m}$ is the set of trajectories satisfying the specified safety and mission requirements of mode m.

Consider a robot approaching a stationary obstacle at speed $v_0 > 0$, with measured clearance $\hat{d}_0$ and absolute measurement error bounded by $\epsilon \geq 0$. Let $\tau \geq 0$ bound the braking-response delay under communication attacks, and $\rho \in [0, 1)$ bound the fractional loss of nominal braking deceleration $a_0 > 0$ due to actuator faults. Assume straight-line motion, no acceleration during the delay, braking deceleration of at least $(1 - \rho)a_0$, and reliable holding at rest. A brake-and-hold policy preserves minimum clearance $d_{\min} > 0$ and stops by the end time $T > 0$ if

$$\epsilon + v_0\tau + \frac{v_0^2}{2(1-\rho)a_0} \leq \hat{d}_0 - d_{\min}, \qquad \tau + \frac{v_0}{(1-\rho)a_0} \leq T. \tag{2}$$

For $\delta = (\epsilon, \tau, \rho)$, these conditions certify a policy-dependent inner approximation of the stopping-mode resilience envelope. Characterization of the resilience envelope relates system degradation and available resources to achievable closed-loop behavior.

3.2. Self-aware and anticipatory resilience

Beyond responding to detected disruptions, next-generation autonomous systems should anticipate degradation using component-health estimates, communication integrity, perception confidence, environmental conditions, and closed-loop performance.

This capability requires quantitative assessment of uncertainty, trustworthiness, degradation, and residual capability. Such information can support online decisions concerning mission continuation, controller reconfiguration, and fallback operation. Real-time evaluation of cyberattack-resilient control in large networked systems also illustrates the value of connecting theoretical resilience guarantees with system-level implementation evidence[51].

The corresponding control problem is therefore not only to estimate the physical state, but also to assess whether the assumptions underlying current autonomous decisions remain justified.

3.3. Learning for intelligence and control for assurance

Learning-based and model-based methods provide complementary capabilities. Learning supports perception, prediction, adaptation, and high-level decision-making, whereas model-based control provides mechanisms for enforcing physical constraints and safety requirements[11].

Control barrier functions (CBFs)[52,53], predictive safety filters[54], backup model predictive control[55], and runtime assurance[56] provide complementary mechanisms for enforcing safety requirements on advanced autonomous behavior.

An appropriate architectural principle is to combine learning for performance and adaptation with control and supervisory mechanisms whose safety guarantees explicitly account for information uncertainty and physical constraints. At the control-input level, a representative safety filter selects the control input according to

$$u^{\star}(t) \in \underset{u \in \mathcal{U}_{\mathrm{safe}}(\mathcal{I}_t)}{\arg\min} \frac{1}{2} \left\| u - u_{\mathrm{L}}(t) \right\|^2, \tag{3}$$

where $u^{\star}(t)$ is the filtered control input, $u_{\mathrm{L}}(t)$ is the input proposed by a learned policy, $\mathcal{I}_t$ denotes the available information history, and $\mathcal{U}_{\mathrm{safe}}(\mathcal{I}_t)$ contains inputs certified under the adopted uncertainty description and physical constraints[52,56]. The information history $\mathcal{I}_t$ should support a state set $\mathcal{X}_t$ whose coverage of the true state is justified under the declared sensing and attack assumptions. The key challenge is to construct and update $\mathcal{U}_{\mathrm{safe}}(\mathcal{I}_t)$ while preserving the validity of the safety certificate and maintaining closed-loop feasibility. Online updates should distinguish a performance policy from the models and uncertainty bounds supporting its safety supervisor. After a policy update, the safety filter adjusts the proposed control input to ensure that the applied input still satisfies the safety constraints. Changes to the dynamics model or uncertainty bounds require revalidation of the safety certificate and continued feasibility before deployment. Updates that cannot be certified as safe should be rejected. The previous controller may be retained only while its safety assumptions and feasibility conditions remain satisfied. This formulation motivates confidence-aware resilient autonomy, in which authority assigned to learned policies depends on their estimated reliability.

3.4. Effective diversity and collective intelligence

The resilience benefit of redundancy depends on the extent to which alternative resources share failure mechanisms. Common-mode failures can substantially undermine the protection provided by nominal redundancy. Multiple learning models trained on similar data may exhibit correlated vulnerabilities, and nominally independent communication channels may share common infrastructure.

The relevant system property is therefore effective diversity, encompassing sensing, learning algorithms, computation, communication, and actuation. Effective diversity can be assessed by identifying shared dependencies and testing whether multiple resources fail under the same disruptions. Dependencies may

include common power supplies, communication infrastructure, software, or training data. Simulation or controlled fault-injection tests should record joint failures against predefined performance requirements and evaluate the remaining sensing, control, communication, and task capabilities. This assesses whether nominally diverse resources provide usable alternatives under the tested conditions.

For multi-agent autonomous systems, resilience must also extend beyond consensus to collective sensing and planning[50]. Shared information should be checked against physical predictions and independent observations using an attack-resilient fusion rule. Information authentication alone does not establish physical correctness. For example, an inspection team may isolate a robot reporting inconsistent observations. The remaining agents may still reach consensus, yet lose the sensing coverage or resources needed to complete the inspection. Task reallocation or an explicitly permitted degraded mission therefore requires reassessing residual sensing, communication, and actuation capabilities.

3.5. Recovery, learning, and graceful degradation

Resilient autonomy should integrate anticipation, detection, isolation, disruption tolerance, reconfiguration, recovery, and learning. Digital twins may support monitoring and virtual evaluation[49], but their usefulness during and after a disruption depends on the validity of the models and data used to update them.

Resilience should also support graceful mission degradation. When nominal objectives become infeasible, the system should assess whether a degraded mission or safe fallback remains feasible and select an appropriate operating mode. Operating modes may progress from nominal operation to degraded operation, safe fallback, and controlled shutdown. A fallback must be initiated while a safe transition remains feasible, and an empty certified input set does not itself imply that safe recovery is still possible. This formulation requires control and planning methods capable of relating residual resources and uncertainty directly to achievable mission objectives. Classical autonomous planning provides a useful basis for such integration[57].

3.6. Verification and resilience by design

A central issue is compositional assurance, namely deriving system-level resilience claims from heterogeneous evidence concerning perception, estimation, communication, learning, planning, and control. Resilience must also be considered during architectural design. Sensor diversity, network topology, computational resources, actuator capability, energy reserves, fallback modes, learning algorithms, and supervisory mechanisms determine which recovery strategies are feasible. This life-cycle perspective is consistent with approaches that combine design-time formal verification with runtime monitoring to assess whether the assumptions underlying verification remain valid during operation[58].

4. OPEN CHALLENGES

To advance resilient autonomy in the next generation of intelligent autonomous systems, we summarize the following open challenges.

4.1. Quantitative characterization of resilience

A central challenge is to characterize safety preservation, performance degradation, residual capability, and recovery within a common, mission-relevant assessment framework. Future resilience measures should also support online assessment so that resilience can influence real-time mission and control decisions. A practical route is to compute inner approximations of the resilience envelope using robust reachability or model predictive control, with information, resource, and mission constraints included jointly.

4.2. Unknown and intelligent threats

Existing methods generally assume restrictions on attack magnitude, sparsity, frequency, or duration. Such assumptions do not by themselves exclude adaptive adversaries. The open question is whether the attack classes defined by these restrictions adequately capture relevant attacker capabilities and cross-layer effects. Unknown attacks refer to mechanisms or patterns that are not assumed to be classified in advance. Adaptive adversaries are considered within application-specific limits on compromised information channels, communication availability, and loss of actuation capability. The independent observations and supervisory mechanisms on which a safety certificate relies are assumed to remain available and satisfy their prescribed error bounds. Future resilient-control methods should exploit physical invariants, heterogeneous sensing, and input–output consistency to assess the reliability of feedback information without requiring prior classification of the attack mechanism.

4.3. Resilience of perception and decision-making

Perception and high-level decision-making determine both the information used for feedback and the objectives supplied to the controller. Sensor degradation and adversarial inputs may compromise state estimates and environmental representations, while reasoning errors may alter task interpretation or planning decisions. Autonomous reasoning also introduces attack surfaces beyond conventional sensing and control[59]. The challenge is to represent these different failure mechanisms without assuming they can all be reduced to additive disturbances. Set-valued estimation can propagate physical uncertainty, while independent task-specification checks can reject objectives incompatible with authorized missions or mandatory constraints.

4.4. Guarantees for learning and adaptation

Building on Section 3.3, the remaining challenge is to derive update-admission conditions that preserve safety and recursive feasibility when policy learning, state estimation, and supervisory switching interact. These conditions must account for experience collected during disruptions and changes to the models underlying the safety certificate. A tractable starting point is to preserve recursive feasibility using robust invariant sets with explicitly bounded model uncertainty[60].

4.5. Scalability, heterogeneity, and common-mode failures

For next-generation autonomous systems operating over large-scale heterogeneous networks, a key open challenge is to assess shared failure dependencies and certify residual task capability using local information without incurring excessive communication or computational costs. Dependency graphs can represent shared infrastructure or software and couple admissible joint failures to observability, connectivity, and task-allocation constraints.

4.6. Safe recovery and mission reconfiguration

Recovery should be formulated as a constrained transition from the post-disruption condition to a feasible operating mode, rather than assumed to be a return to nominal operation. The central problem is to characterize which post-event states, information conditions, and remaining resources admit a recovery strategy that preserves safety throughout the transition. A practical route is to compute reachable sets or robust predictive trajectories under residual actuation, sensing, and energy constraints. A candidate recovery plan should be accepted only when its entire transition satisfies safety constraints and reaches a feasible nominal, degraded, or stopping mode. Evaluation should report minimum safety margin, recovery time, and retained task coverage.

4.7. Compositional certification and human-autonomy interaction

Component-level guarantees do not automatically imply system-level or mission-level resilience. A practical route is to specify component assumptions and guarantees using assume-guarantee contracts[61], then verify

their compatibility after reconfiguration. Resilience architectures should define conditions for continued autonomous operation, reduced autonomy, transfer of authority, and safe fallback. Cognitive-physical shared control offers one possible approach to incorporating operator state and adaptive authority allocation directly into control. Certification methods should ultimately provide evidence not only that individual components satisfy specified requirements, but also that the integrated autonomous system can maintain an acceptable level of safety and mission capability under defined classes of degradation.

5. CONCLUSIONS

Resilient control is evolving from protecting individual feedback loops toward sustaining safe autonomous operation across perception, learning, decision-making, cooperation, and physical execution. The four complementary transitions discussed in this article motivate resilient autonomy, in which systems assess information trustworthiness and residual capability, adapt their architecture and mission objectives, preserve essential functionality, and recover safely. The central challenge is to connect these capabilities to verifiable closed-loop and mission-level guarantees under explicit threat and resource assumptions. Stability, robustness, observability, invariance, and feedback remain fundamental, but must be integrated with information assessment, planning, and recovery. Progress should be judged not only by whether a system tolerates a disruption but also by whether it retains the capacity to continue its mission, adopt a feasible degraded objective, or terminate the mission safely.

DECLARATIONS

Authors' contributions

Preparation of the original draft and figures: Fu, C.
Supervision and guidance: Liu, L.; Feng, G.

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

During the preparation of this manuscript, the AI tool ChatGPT (version 6.1, released 2026-09-29) was used for language editing and to provide suggestions on the color scheme of the graphical abstract. The tool did not influence the study design, data collection, analysis, interpretation, or the scientific content of the work. All authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Financial support and sponsorship

This work was supported in part by the National Natural Science Foundation of China (Grant 62373314), and in part by the Research Grants Council of the Hong Kong Special Administrative Region of China (Projects CityU/11206625 and CityU/11212225).

Conflicts of interest

Feng, G. is a Section Member of the Robot Control and Navigation Section of the journal *Intelligent & Robotics*, but was not involved in any steps of editorial processing, notably including reviewer selection, manuscript handling, and decision-making, while the other authors have declared that they have no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

REFERENCES

1. Isermann, R. *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*, 1st ed.; Springer Berlin, Heidelberg, 2006. DOI
2. Zhang, Y.; Jiang, J. Bibliographical review on reconfigurable fault-tolerant control systems. *Annu. Rev. Control.* **2008**, *32*, 229-52. DOI
3. Blanke, M.; Kinnaert, M.; Lunze, J.; Staroswiecki, M. *Diagnosis and fault-tolerant control*, 3rd ed.; Springer Berlin, Heidelberg, 2016. DOI
4. Cardenas, A. A.; Amin, S.; Sastry, S. Secure control: towards survivable cyber-physical systems. In *2008 The 28th International Conference on Distributed Computing Systems Workshops*, Beijing, China. Jun 17-20, 2008. IEEE; 2008. pp. 495-500. DOI
5. Pasqualetti, F.; Dorfler, F.; Bullo, F. Attack detection and identification in cyber-physical systems. *IEEE. Trans. Autom. Control.* **2013**, *58*, 2715-29. DOI
6. Fawzi, H.; Tabuada, P.; Diggavi, S. Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE. Trans. Autom. Control.* **2014**, *59*, 1454-67. DOI
7. Teixeira, A.; Shames, I.; Sandberg, H.; Johansson, K. H. A secure control framework for resource-limited adversaries. *Automatica* **2015**, *51*, 135-48. DOI
8. De Persis, C.; Tesi, P. Input-to-state stabilizing control under denial-of-service. *IEEE. Trans. Autom. Control.* **2015**, *60*, 2930-44. DOI
9. Segovia-Ferreira, M.; Rubio-Hernan, J.; Cavalli, A.; Garcia-Alfaro, J. A survey on cyber-resilience approaches for cyber-physical systems. *ACM. Comput. Surv.* **2024**, *56*, 1-37. DOI
10. Ishii, H.; Wang, Y.; Feng, S. An overview on multi-agent consensus under adversarial attacks. *Annu. Rev. Control.* **2022**, *53*, 252-72. DOI
11. Brunke, L.; Greeff, M.; Hall, A. W.; et al. Safe learning in robotics: from learning-based control to safe reinforcement learning. *Annu. Rev. Control. Robot. Auton. Syst.* **2022**, *5*, 411-44. DOI
12. Mo, Y.; Sinopoli, B. Secure control against replay attacks. In *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, USA. Sep 30 - Oct 02, 2009. IEEE; 2009. pp. 911-8. DOI
13. Teixeira, A.; Sandberg, H.; Johansson, K. H. Networked control systems under cyber attacks with applications to power networks. In *Proceedings of the 2010 American Control Conference*, Baltimore, USA. Jun 30 - Jul 02, 2010. IEEE; 2010. pp. 3690-6. DOI
14. Pajic, M.; Lee, I.; Pappas, G. J. Attack-resilient state estimation for noisy dynamical systems. *IEEE. Trans. Control. Netw. Syst.* **2017**, *4*, 82-92. DOI
15. Shoukry, Y.; Nuzzo, P.; Bezzo, N.; Sangiovanni-Vincentelli, A. L.; Seshia, S. A.; Tabuada, P. Secure state reconstruction in differentially flat systems under sensor attacks using satisfiability modulo theory solving. In *2015 54th IEEE Conference on Decision and Control (CDC)*, Osaka, Japan. Dec 15-18, 2015. IEEE; 2015. pp. 3804-9. DOI
16. Chen, B.; Tan, Y.; Sun, Z.; Yu, L. Attack-resilient control against FDI attacks in cyber-physical systems. *IEEE/CAA. J. Autom. Sinica.* **2022**, *9*, 1099-102. DOI
17. Liu, W.; Sun, J.; Wang, G.; Bullo, F.; Chen, J. Resilient control under quantization and denial-of-service: codesigning a deadbeat controller and transmission protocol. *IEEE. Trans. Autom. Control.* **2022**, *67*, 3879-91. DOI
18. Yu, M.; Wang, W.; Yan, J. Distributed adaptive resilient consensus control for uncertain nonlinear multiagent systems against deception attacks. *IEEE. Control. Syst. Lett.* **2025**, *9*, 1411-6. DOI
19. Sundaram, S.; Hadjicostis, C. N. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE. Trans. Autom. Control.* **2011**, *56*, 1495-508. DOI
20. Leblanc, H. J.; Zhang, H.; Koutsoukos, X.; Sundaram, S. Resilient asymptotic consensus in robust networks. *IEEE. J. Sel. Areas. Commun.* **2013**, *31*, 766-81. DOI
21. Zhou, J.; Lv, Y.; Wen, G.; Yu, X. Resilient consensus of multiagent systems under malicious attacks: appointed-time observer-based approach. *IEEE. Trans. Cybern.* **2022**, *52*, 10187-99. DOI PubMed
22. Huo, S.; Zhang, Y.; Lewis, F. L.; Sun, C. Observer-based resilient consensus control for heterogeneous multiagent systems against cyberattacks. *IEEE. Trans. Control. Netw. Syst.* **2023**, *10*, 647-58. DOI
23. Fang, F.; Li, J.; Liu, Y.; Park, J. H. Resilient control for multiagent systems with a sampled-data model against DoS attacks. *IEEE. Trans. Ind. Inform.* **2023**, *19*, 780-9. DOI
24. Venkateswaran, D. B.; Qu, Z. Resilient multi-agent systems against denial of service attacks via adaptive and activatable network layers. *IEEE. Open. J. Control. Syst.* **2024**, *3*, 472-82. DOI
25. Viadero-Monasterio, F.; Meléndez-Useros, M.; Jiménez-Salas, M.; López Boada, B.; Jesús López Boada, M. Key influencing factors in vehicle platoons: a systematic study and review. *Evol. Syst.* **2025**, *16*, 116. DOI
26. Chen, H.; Smolka, S. A.; Paoletti, N.; Lin, S. An STL-based approach to resilient control for cyber-physical systems. In *Proceedings of the 26th ACM International Conference on Hybrid Systems: Computation and Control*. Association for Computing Machinery; 2023. pp. 1-12. DOI

27. Fainekos, G. E.; Pappas, G. J. Robustness of temporal logic specifications for continuous-time signals. *Theor. Comput. Sci.* **2009**, *410*, 4262-91. DOI
28. Gao, W.; Jiang, Z. Data-driven cooperative output regulation of multi-agent systems under distributed denial of service attacks. *Sci. China. Inf. Sci.* **2023**, *66*, 190201. DOI
29. Deng, C.; Gao, W.; Wen, C.; Chen, Z.; Wang, W. Data-driven practical cooperative output regulation under actuator faults and DoS attacks. *IEEE. Trans. Cybern.* **2023**, *53*, 7417-28. DOI
30. Wen, H.; Li, Y.; Tong, S. Distributed adaptive resilient formation control for nonlinear multi-agent systems under DoS attacks. *Sci. China. Inf. Sci.* **2024**, *67*, 209201. DOI
31. Gong, X.; Li, X.; Shu, Z.; Feng, Z. Resilient output formation-tracking of heterogeneous multiagent systems against general Byzantine attacks: a twin-layer approach. *IEEE. Trans. Cybern.* **2024**, *54*, 2566-78. DOI
32. Yan, C.; Yan, L.; Lv, Y.; Dong, B.; Xia, Y. Security output Byzantine - optimization of heterogeneous multiagent systems via adaptive self-triggered detection approach. *IEEE. Internet. Things. J.* **2026**, *13*, 49326-39. DOI
33. Ye, Z.; Zhang, D.; Deng, C.; Yan, H.; Feng, G. Finite-time resilient sliding mode control of nonlinear UMV systems subject to DoS attacks. *Automatica* **2023**, *156*, 111170. DOI
34. Cao, W.; Zhang, D.; Feng, G. Resilient semi-global finite-time cooperative output regulation of heterogeneous linear multi-agent systems subject to denial-of-service attacks. *Automatica* **2025**, *173*, 112099. DOI
35. Cao, W.; Liu, L.; Ye, Z.; Zhang, D.; Feng, G. Resilient global practical fixed-time cooperative output regulation of uncertain nonlinear multiagent systems subject to denial-of-service attacks. *IEEE. Trans. Autom. Control.* **2026**, *71*, 1247-54. DOI
36. Vamvoudakis, K. G.; Lewis, F. L. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* **2010**, *46*, 878-88. DOI
37. Ma, Y. S.; Che, W. W.; Deng, C.; Wu, Z. G. Distributed model-free adaptive control for learning nonlinear MASs under DoS attacks. *IEEE. Trans. Neural. Netw. Learn. Syst.* **2023**, *34*, 1146-55. DOI
38. Chen, P.; Liu, S.; Chen, B.; Yu, L. Multi-agent reinforcement learning for decentralized resilient secondary control of energy storage systems against DoS attacks. *IEEE. Trans. Smart. Grid.* **2022**, *13*, 1739-50. DOI
39. Deng, C.; Yue, D.; Che, W. W.; Xie, X. Cooperative fault-tolerant control for a class of nonlinear MASs by resilient learning approach. *IEEE. Trans. Neural. Netw. Learn. Syst.* **2024**, *35*, 670-9. DOI
40. Chakraborty, S.; Gao, W.; Vamvoudakis, K. G.; Jiang, Z. Robust learning-based output regulation under denial-of-service attacks. *IEEE. Trans. Autom. Control.* **2026**, 1-8. DOI
41. Chakraborty, S.; Gao, W.; Jiang, Z. A review on resilient reinforcement learning for linear cyber-physical systems. *J. Syst. Sci. Complex.* **2026**, *39*, 729-58. DOI
42. Ye, Z.; Zhang, D.; Cheng, J.; Wu, Z. Event-triggering and quantized sliding mode control of UMV systems under DoS attack. *IEEE. Trans. Veh. Technol.* **2022**, *71*, 8199-211. DOI
43. Wu, C.; Zhu, Y.; Xu, L.; et al. Attack detection and security control for quadrotor under stealthy attacks. *Syst. Control. Lett.* **2025**, *197*, 106031. DOI
44. Hilmi, M.; Widyotriatmo, A.; Hasan, A. Resilient control based on secure state estimation for autonomous systems under cyberattacks. *IEEE. Trans. Cybern.* **2025**, *55*, 3560-72. DOI PubMed
45. Liu, S.; Guo, G. Distributed resilient control for autonomous surface vehicles via self-triggered communication under denial-of-service attacks. *IEEE. Trans. Autom. Sci. Eng.* **2026**, *23*, 3501-10. DOI
46. Han, Q.; Ma, J.; Zuo, Z.; Wang, X.; Yang, B.; Guan, X. Resilient event-triggered control of vehicle platoon under DoS attacks and parameter uncertainty. *IEEE. Trans. Intell. Veh.* **2025**, *10*, 169-78. DOI
47. Gong, J.; Ding, L.; Jia, C.; Liu, Y.; Cao, J. Toward resilient vehicle platooning: a two-layer secure control architecture against hybrid cyber-physical threats. *IEEE/CAA. J. Autom. Sinica.* **2026**, *13*, 1082-96. DOI
48. Khoshnevisan, L.; Liu, X. A secure adaptive resilient neural network-based control of heterogeneous connected automated vehicles subject to cyber attacks. *IEEE. Trans. Veh. Technol.* **2025**, *74*, 8734-44. DOI
49. Li, J.; Yang, S. X. Digital twins to embodied artificial intelligence: review and perspective. *Intell. Robot.* **2025**, *5*, 202-27. DOI
50. Schlotfeldt, B.; Tzoumas, V.; Pappas, G. J. Resilient active information acquisition with teams of robots. *IEEE. Trans. Robot.* **2022**, *38*, 244-61. DOI
51. Oshnoei, S.; Hassani, S.; Gheisarnejad, M.; Mahboubi-Moghaddam, E. Real-time evaluation of cyberattack-resilient control for secure large-scale power networks. *IEEE. Internet. Things. J.* **2026**, *13*, 4449-62. DOI
52. Ames, A. D.; Xu, X.; Grizzle, J. W.; Tabuada, P. Control barrier function based quadratic programs for safety critical systems. *IEEE. Trans. Autom. Control.* **2017**, *62*, 3861-76. DOI
53. Ames, A. D.; Coogan, S.; Egerstedt, M.; Notomista, G.; Sreenath, K.; Tabuada, P. Control barrier functions: theory and applications. In *2019 18th European Control Conference (ECC)*, Naples, Italy. Jun 25-28, 2019. IEEE; 2019. pp. 3420-31. DOI

54. Wabersich, K. P.; Zeilinger, M. N. A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica* **2021**, *129*, 109597. DOI
55. Tao, R.; Kim, H.; Yoon, H.; et al. Backup plan constrained model predictive control with guaranteed stability. *J. Guid. Control. Dyn.* **2024**, *47*, 233-46. DOI
56. Hobbs, K. L.; Mote, M. L.; Abate, M. C.; Coogan, S. D.; Feron, E. M. Runtime assurance for safety-critical systems: an introduction to safety filtering approaches for complex control systems. *IEEE. Control. Syst.* **2023**, *43*, 28-65. DOI
57. LaValle, S. M. *Planning algorithms*. Cambridge University Press, 2006. https://books.google.fr/books?id=-PwLBAAAQBAJ. (accessed 2026-11-18).
58. Luckcuck, M.; Farrell, M.; Dennis, L. A.; Dixon, C.; Fisher, M. Formal specification and verification of autonomous robotic systems: a survey. *ACM. Comput. Surv.* **2020**, *52*, 1-41. DOI
59. Kim, J.; Liu, X.; Wang, Z.; et al. The attack and defense landscape of agentic AI: a comprehensive survey. *Superintell. Robot. Saf. Align.* **2026**, *3*. DOI
60. Lorenzen, M.; Cannon, M.; Allgöwer, F. Robust MPC with recursive model update. *Automatica* **2019**, *103*, 461-71. DOI
61. Sharf, M.; Besselink, B.; Johansson, K. H. Contract composition for dynamical control systems: definition and verification using linear programming. *Automatica* **2024**, *164*, 111637. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.