Federal Register Notice: 89 FR 78915, Federal Register :: Networking and Information Technology Research and Development Request for Information on Cyber-Physical Systems Resilience Research, September 26, 2024.

Request for Information on the National Cyber-Physical Systems Resilience Plan

Hongkai Chen

Shan Lin

Nicola Paoletti

Scott A. Smolka

Specification, Verification, and Control for Resilient Cyber-Physical Systems:

Response to NSF RFI on CPS Resilience

Hongkai Chen[1], Shan Lin[2], Nicola Paoletti[3], and Scott A. Smolka[4]

[1]Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong
[2]Department of Electrical and Computer Engineering, Stony Brook University
[3]Department of Informatics, King's College London, UK
[4]Department of Computer Science, Stony Brook University

October 25, 2024

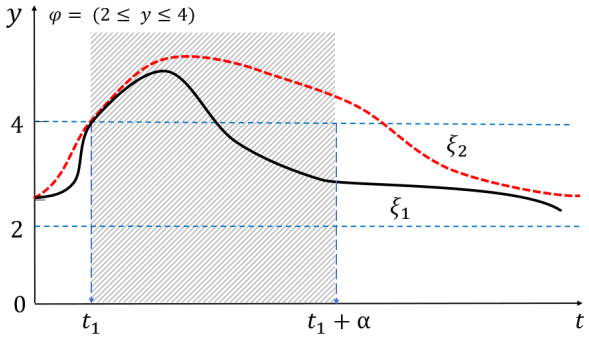
1 Introduction

Resiliency (*syn.* resilience) is defined as the ability to recover from or adjust easily to adversity or change [7]. Resiliency is of fundamental importance in Cyber-Physical Systems (CPS), which are expected to exhibit safety- or mission-critical behavior even in the presence of internal faults or external disturbances. Consider for example the *lane keeping* problem for autonomous vehicles (AVs), which requires a vehicle to stay within the marked boundaries of the lane it is driving in. The standard temporal-logic-based notion of safety is not ideally suited for specifying the AV's behavior when it comes to lane keeping. This is because AV technology is not perfect and driving conditions (e.g., being crowded by a neighboring vehicle) and other external disturbances may require occasional or even intermittent violations of lane keeping. Rather, the AV should behave resiliently in the presence of a lane violation, recovering from the violation in a timely fashion (*recoverability*), and avoiding future lane departures for as long as possible (*durability*). Unfortunately, there is no widely agreed upon notion of resiliency within the CPS community, despite several efforts to settle the issue,

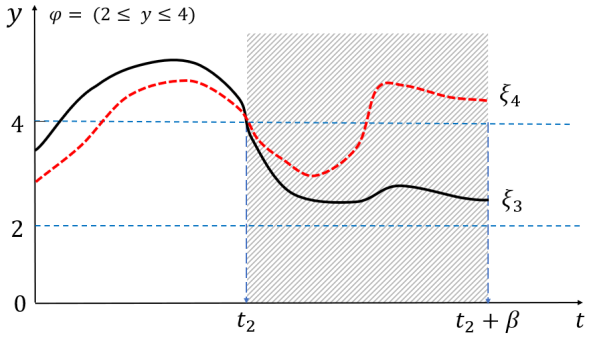
Moreover, reasoning about resiliency in CPS and creating resilient systems are considered to be open problems in the recently published *Control for Societal-scale Challenges: Road Map 2030* [2]. To address this state of affairs—and to simultaneously point out how well-aligned our research on CPS Resilience is with the *NSF RFI*, we have recently made two foundational contributions to the theory and practice of CPS resiliency:

Contribution 1: In [4], we developed an STL [6]-based framework for specifying and reasoning about resiliency in CPS in which resiliency has a syntactic characterization in the form of an *STL-based Resiliency Specification* (SRS). Given an STL formula φ, time bounds α and β, the SRS of φ, denoted $R_{\alpha,\beta}(\varphi)$, requires that recovery from a violation of φ occur within time α (*recoverability*), and subsequently that φ be maintained for duration β (*durability*). These R-expressions, which are atoms in our SRS logic, can be combined using STL operators, allowing one to express composite resiliency specifications, e.g., multiple SRSs must hold simultaneously, or the system must eventually be resilient. Further details about this work are provided in the Prior Work section of this proposal (Section 2.1).

Contribution 2: In [5], we present *ResilienC*, a framework for resilient control of Cyber-Physical Systems subject to STL-based requirements. We define the *Resilient STL Control Problem* as one of *multi-objective optimization*, where the recoverability and durability of the desired STL specification are maximized. When neither objective is prioritized over the other, the solution to the problem is a set of *Pareto-optimal* system trajectories. We present a precise solution method to the problem using a mixed-integer linear programming encoding and an *a posteriori* ϵ-constraint approach for efficiently retrieving the complete set of optimally resilient solutions. In ResilienC, at each time-step, the optimal control action selected from the set of Pareto-optimal



(a) Signal ξ_1 satisfies recoverability.



(b) Signal ξ_3 satisfies durability.

Figure 1: Resilience w.r.t. STL formula $\varphi = (2 \leq y \leq 4)$

solutions by a *Decision Maker* strategy realizes a form of *Model Predictive Control*. Further details about ResilienC are available in Section 2.2.

2 Our Prior Work on CPS Resilience

In this section, we provide the reader with further details regarding our two contributions to CPS Resilience outlined in Section 1.

2.1 An STL-based Formulation of Resiliency

In a recent paper [4],[1] we formulate CPS resiliency using *STL-based Resiliency Specifications* (SRSs). Given an arbitrary STL formula φ (STL stands for Signal Temporal Logic [6]), time bounds α and β, the SRS of φ, $R_{\alpha,\beta}(\varphi)$, is the STL formula $\neg\varphi \mathbf{U}_{[0,\alpha]} \mathbf{G}_{[0,\beta)}\varphi$, specifying that recovery from a violation of φ occur within time α, and subsequently φ be maintained for duration at least β. The SRS of φ captures the requirement that a system quickly recovers from a violation of φ (*recoverability*) and then satisfy φ for an extended period of time (*durability*). $R_{\alpha,\beta}(\varphi)$ expressions, which are atoms in our SRS logic, can be inductively combined using STL operators, allowing one to express composite resiliency specifications; e.g., multiple SRSs must hold simultaneously ($R_{\alpha_1,\beta_1}(\varphi_1) \wedge R_{\alpha_2,\beta_2}(\varphi_2)$), and the system must eventually be resilient ($\mathbf{F}_I R_{\alpha,\beta}(\varphi)$).

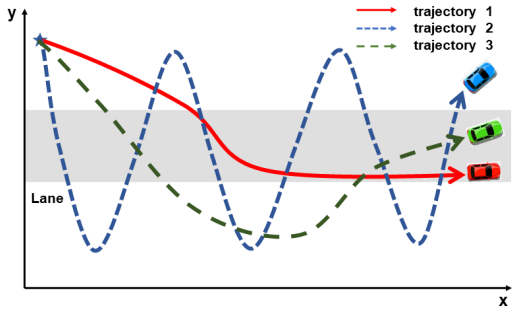
Consider for example the STL specification $\varphi = (2 \leq y \leq 4)$, where y is a signal. In Figure 1(a), signals ξ_1 and ξ_2 violate φ at time t_1. Given recovery deadline α, we see that only ξ_1 satisfies recoverability of φ w.r.t. α because φ becomes true before $t_1 + \alpha$. In the case of ξ_2, φ becomes true only after $t_1 + \alpha$. In Figure 1(b), signals ξ_3 and ξ_4 recover to satisfy φ at time t_2. Given durability bound β, we observe that only ξ_3 is durable w.r.t. β.

We can quantify the level of resiliency exhibited by a system trajectory by defining a so-called "quantitative semantics" for SRSs in the form of a *Resiliency Satisfaction Value* (ReSV) function r. Our semantics for $R_{\alpha,\beta}(\varphi)$ atoms is a singleton set of the form $\{(rec, dur)\}$, where *rec* quantifies how early before bound α recovery occurs, and *dur* indicates for how long after bound β property φ is maintained. The r-value for a composite SRS formula results in a set of non-dominated recoverability-durability pairs, given that the ReSVs of subformulas might not be directly comparable (e.g., one subformula has superior durability but worse recoverability than another). To the best of our knowledge, this is the first *multi-dimensional* quantitative semantics for STL.

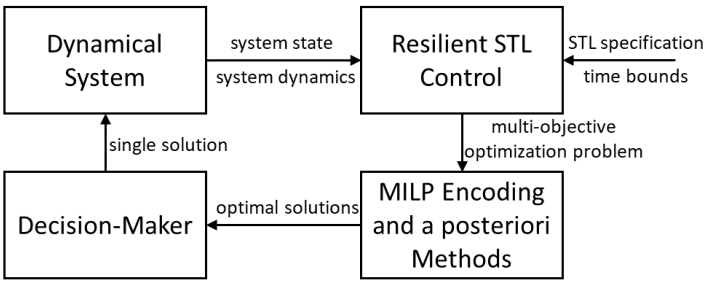
2.2 Resilient Control

The formulation of resilience presented in Section 2.1 naturally suggests the problem of *resilient STL control*: find an optimal control strategy that maximizes the system's resilience in terms of recoverability and durability. These two objectives are often at odds with each other. For example, in the *lane-keeping problem* (see Figure 2(a)), the initial location of the vehicle, marked by a star, is outside the lane, thereby violating a lane-keeping specification φ. Trajectory 1 is the slowest to recover from its violation of φ, resulting in the worst recoverability among the three trajectories shown in the figure. However, the vehicle subsequently maintains φ until the end of the trajectory, resulting in the best durability. In trajectory 2, the vehicle aggressively steers back into the lane whenever φ is violated; so it exhibits the best recoverability. It cannot, however, maintain φ post-satisfaction due

[1] Received the Best Paper Award at the 2022 FORMATS conference.



(a) Illustration of resilient optimal solutions.



(b) The ResilienC architecture.

Figure 2: Illustration of resilient trajectories and Overview of ResilienC control framework.

to overshooting, and thus has the worst durability. The behavior of trajectory 3 lies in-between trajectories 1 and 2. The three trajectories are *mutually non-dominated* and, hence, equally resilient.

To address the resilient STL control problem, we have developed the *ResilienC* control framework, which formulates this problem as one of multi-objective optimization, designed to maximize both the recoverability and durability of the CPS under consideration [5]. Unlike existing techniques for STL-based control [8, 9] which focus on optimizing a single objective (e.g., spatial robustness in [8] and time robustness in [9]) and thus produce a single solution, our method results in a set of *non-dominated, aka Pareto-optimal, solutions*. Such a method is also called *a posteriori* as it avoids making any *a priori* assumptions about the relative importance of the two objectives. We realize an MPC scheme with our approach by deploying a *Decision Maker* strategy that at each time-step, selects the next optimal control action from among the set of Pareto-optimal solutions for execution by the plant. See Figure 2(b) for an overview of the ResilienC framework.

3 Proposed Work

The NSF RFI on CPS Resilience is very thought provoking and leads us to consider the following research activities as "Proposed Work". In particular, we put forth the *ReSCVe CPS* framework for Multi-Dimensional Resiliency Specification, Control, and Verification of CPS. The *ReSCVe CPS* effort (pronounced "Rescue CPS") will be centered around the following research tasks (please also see Figure 3).

Task 1: Specifying and reasoning about both *spatial* and *temporal* resiliency in CPS. A logical framework for reasoning about CPS resilience in the spatial-temporal domain will be developed. Resilient CPS subjected to external disturbances are expected to recover from such events and subsequently maintain desired behavior that may be subject to spatial and temporal requirements. Case studies will be used to demonstrate the expressiveness and utility of spatial-temporal resiliency specifications.

Task 2: Accurate and efficient predictive monitoring of CPS resiliency properties. We will devise methods to reliably predict the recoverability and durability of CPS at runtime, deriving prediction intervals with probabilistic coverage guarantees.

Task 3: Resilient control under uncertainty using MPC and Reinforcement learning. We will investigate resilient control techniques for CPS subject to two kinds of uncertainty: stochasticity (e.g., random disturbances) and non-determinism (when uncertainty cannot be quantified).

Task 4: Investigate the *Inherently resilient vs. Engineered-to-be-resilient systems* dichotomy. We believe that much is to be gained from understanding the mechanisms deployed in inherently resilient CPS (especially those occurring in nature) and applying these techniques to engineered CPS. For example, the human pancreas is an inherently resilient system for regulating human blood glucose levels, while the artificial pancreas is an engineered-to-be-resilient version of this digestive system organ.

Task 5: Perform an extensive performance evaluation of the ReSCVe CPS framework centered around three significant Case Studies. Case studies will involve demonstrating the resilience of the following CPS: (1) the Artificial Pancreas [3] (AP) subject to hypo- and hyper-glycemia; (2) Networked Microgrids [1] (NM) in the presence of varying loads, disconnection of distributed energy resources, and islanding from the main power grid; and (3) collective autonomous agents in *flock formation* subject to various external disturbances, including (artificial) wind, obstacles, and highly agile predator agents.

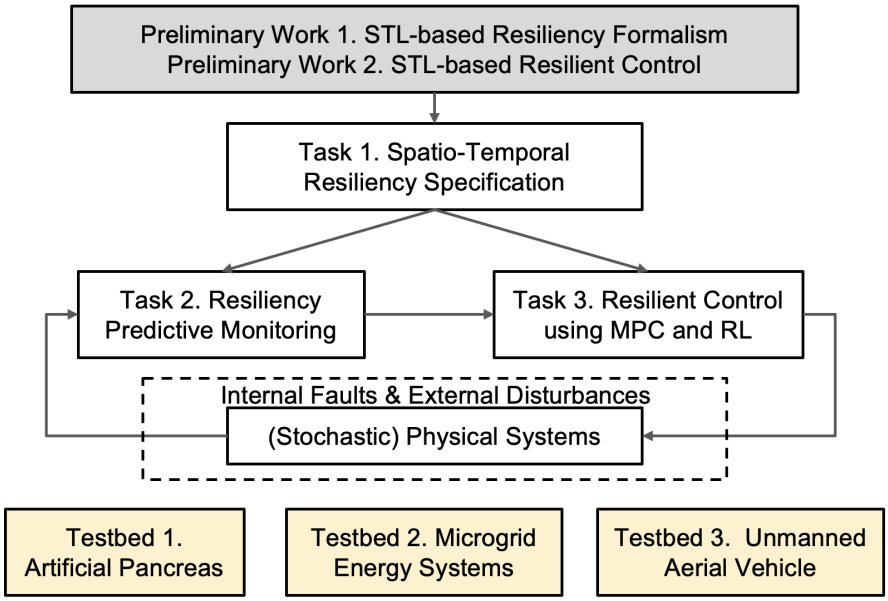


Figure 3: Architectural overview of proposed *ReSCVe CPS* framework.

4 Conclusion

We have presented/highlighted our recent work on CPS Resilience, especially when it come to formal specification, verification, and control, and outlined proposed work in the area focused around the ReSCVe CPS framework. Our approach to CPS resilience enjoys the following benefits.

- Our logic-based framework for CPS resilience is foundational because it can be applied to *runtime monitoring*, *online/MPC-based control* (resilient-by-intervention), and *synthesis of fixed policies/controllers/system parameters* (which can be seen as resilient-by-design). Each of these approaches can be traditional/mechanistic or data-driven.
- To make our approach resilient to unseen threats/adversaries, we can solve all of the above problems in a robust (minmax) or uncertainty-aware (or risk-aware) manner.
- Our current resilient control approach can be made more efficient by considering "offline" policies or solving the single-objective problem (rather than bi-objective, which may be much more expensive) with a single recoverability-durability tradeoff value (which could change/adapt at runtime).

We would welcome the opportunity to pursue our CPS Resilience research goals within the confines of the NSF CPS program, and plan to submit a proposal to the same.

References

[1] Alam, M. N., Chakrabarti, S., and Ghosh, A. Networked microgrids: State-of-the-art and future perspectives. *IEEE Transactions on Industrial Informatics 15*, 3 (Mar. 2019).

[2] Alleyne, A., Allgöwer, F., Ames, A., et al. *Control for Societal-scale Challenges: Road Map 2030.* IEEE Control Systems Society, May 2023.

[3] Boughton, C. K., and Hovorka, R. Advances in artificial pancreas systems. *Science translational medicine 11*, 484 (2019), eaaw4949.

[4] Chen, H., Lin, S., Smolka, S. A., and Paoletti, N. An STL-based formulation of resilience in cyber-physical systems. In *Proceedings of FORMATS 2022, 20th International Conference on Formal Modeling and Analysis of Timed Systems, Best Paper Award* (2022).

[5] Chen, H., Smolka, S. A., Paoletti, N., and Lin, S. An STL-based approach to resilient control for cyber-physical systems. In *HSCC '23: Proceedings of the 26th ACM International Conference on Hybrid Systems: Computation and Control* (May 2023).

[6] Maler, O., and Nickovic, D. Monitoring temporal properties of continuous signals. In *Proceedings of Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*. Springer, Sept. 2004, pp. 152–166.

[7] Merriam-Webster Dictionary. https://www.merriam-webster.com/dictionary/resiliency.

[8] Raman, V., Donzé, A., Maasoumy, M., Murray, R. M., Sangiovanni-Vincentelli, A., and Seshia, S. A. Model predictive control with signal temporal logic specifications. In *Proceedings of the IEEE Conference on Decision and Control* (Dec. 2014), IEEE, pp. 81–87.

[9] Rodionova, A., Lindemann, L., Morari, M., and Pappas, G. J. Time-robust control for STL specifications. In *Proceedings of IEEE Conference on Decision and Control* (Dec. 2021), IEEE.