Physics-Informed Neural Networks for Secure Connected and Autonomous Traffic Modeling

Eunhan Ka

The rapid evolution of transportation systems toward the integration of Connected and Autonomous Vehicles (CAVs) is reshaping the landscape of urban mobility. CAVs promise significant enhancements in traffic efficiency, safety, and environmental sustainability. By enabling vehicles to communicate with each other and with infrastructure, CAV technology has the potential to optimize traffic flow, reduce congestion, and minimize accidents caused by human error. However, the transition from traditional human-driven vehicles to CAVs introduces complex challenges in modeling traffic dynamics and ensuring the cybersecurity of transportation networks.

Accurate traffic modeling is critical for the efficient management of road networks and the development of strategies to mitigate congestion and other traffic-related issues. Traditional traffic models often rely on simplified assumptions that may not capture the intricate dynamics of modern traffic systems, especially with the advent of CAVs. These models may fall short in accounting for the interactions between vehicles, variability in traffic patterns, and the impact of external factors such as cyber threats.

The increasing reliance on digital technologies and communication systems in CAVs raises concerns about potential cybersecurity vulnerabilities. Cyber threats, such as Route Guidance Attacks (RGAs), pose significant risks to the stability and efficiency of traffic networks. Attackers can manipulate route guidance systems, leading to increased congestion, longer travel times, and gridlock situations. There is a pressing need for advanced modeling techniques that can accurately represent traffic dynamics in the context of CAVs and for strategies that can detect, analyze, and mitigate the effects of cyber threats on transportation systems.

This dissertation develops a physics-informed deep learning framework for network-level traffic state estimation and for analyzing, attacking, and hardening urban traffic systems under cyber threats. It progresses from model construction to adversarial manipulation, network-level impact assessment with behavioral heterogeneity, and defense design. The main contributions are:

1. **Physics-Informed Machine Learning of the Generalized Bathtub Model (PIML-GBM):** A hybrid learning framework that calibrates and evolves the generalized bathtub model on large urban networks with mobile location–based data. The approach estimates and forecasts trip volumes and network states with physics-guided regularization, learns network parameters directly from data, and represents traffic dynamics over continuous time–distance fields for city-scale inference and prediction.

2. **Adaptive Spatial–Temporal Domain Decomposition in Physics-Informed Neural Networks (Ada-STDPINN):** A physics-informed neural architecture that partitions the space–time domain to capture localized, rapidly varying traffic phenomena while preserving global conservation laws. It introduces adaptive error control across subdomains and interfaces, and demonstrates superior accuracy and generalization to classical PINNs and data-driven baselines on real networks.

3. **Bounded-Rational Route Guidance Attacks on Urban Navigation Systems:** A threat model and attack generator (e.g., *PhantNav*) that manipulates route suggestions while explicitly modeling travelers' bounded rationality and tolerance for suboptimality. The framework optimizes attack objectives subject to behavioral and platform constraints, and shows that accounting for heterogeneous compliance produces stronger and more realistic disruptions than perfect-compliance assumptions.

4. **Driver Behavior–Aware Resilience of Traffic Networks under Route Guidance Attacks:** An integrated behavior–traffic modeling and experimental analysis that quantifies how user heterogeneity shapes network-level outcomes under attacks. The study identifies regimes where heterogeneity buffers congestion growth and delays collapse, thresholds where attack intensity overwhelms these benefits, and trade-offs between peak severity and recovery dynamics at the city scale.

5. **Physics-Informed Neural Networks via Adversarial Training under Sensor Attacks:** A defense framework that hardens PINN-based estimators against corrupted sensing. It combines adversarial training with worst-case sensor perturbations, a confidence-weighted robust loss that down-weights outliers, and a physics-guided anomaly filter that suppresses inconsistent observations. The result is materially improved robustness and stable reconstruction of traffic states under malicious noise.

The overarching goal is to enhance CAV traffic network resilience against cyber threats through advanced modeling and mitigation techniques. This cohesive approach contributes to safer, more efficient transportation systems and paves the way for integrating autonomous vehicle technology into daily life.

Chapter 2: Physics-Informed Machine Learning of Generalized Bathtub Model

Chapter 2 introduces the PIML-GBM, an innovative hybrid model that effectively combines physical traffic flow models with machine learning techniques. By leveraging the strengths of both approaches, the model enhances the accuracy and interpretability of traffic state estimations (traditional model's RMSE: 0.4748; PIML-GBM's RMSE: 0.0533). Validation with comprehensive road network data from Indianapolis demonstrates its applicability and effectiveness. This foundational work is crucial for transitioning to CAV traffic modeling, providing the baseline upon which subsequent chapters build.

Chapter 3: Adaptive Spatio-Temporal Decomposition-Based Physics-Informed Neural Network of Traffic Flow Models

Building upon the foundational model in Chapter 2, Chapter 3 introduces the Adaptive Spatio-Temporal Decomposition Physics-Informed Neural Network (AdaSTDPINN). This framework tackles the challenge of capturing localized traffic variations and abrupt changes by dividing the spatial-temporal domain into smaller, overlapping subdomains. The method ensures smooth transitions between subdomains through weighted neural network combinations, significantly enhancing the accuracy of traffic dynamics analysis. Validation studies using real-world traffic data demonstrate substantial improvements in both accuracy and computational efficiency over global domain models. This advancement is critical for analyzing large, complex urban networks and sets the stage for modeling scenarios involving disruptions such as cyber attacks.

Chapter 4: Route Guidance Attack Modeling with Bounded Rationality

Chapter 4 leverages the advanced modeling techniques from Chapters 2 and 3 to focus on the development of RGAs. It introduces a mathematical framework that models how travelers perceive and respond to manipulated route suggestions, taking into account their cognitive limitations and tolerance for suboptimal routes through an indifference band. An algorithm strategically adjusts suggested routes in real-time to increase travel time while

keeping changes subtle enough to remain within the traveler's indifference band. Simulations conducted in detailed urban traffic networks demonstrate how incorporating bounded rationality leads to more effective attack strategies, resulting in higher compliance rates and greater disruption potential. This development marks a significant advancement in understanding the mechanisms of RGAs and their real-world implications for transportation cybersecurity.

Chapter 5: Driver Behavior–Aware Resilience of Traffic Networks under Route Guidance Attacks

Chapter 5 advances the dissertation by modeling how route guidance attacks reshape city-scale traffic when drivers respond in heterogeneous ways. It fuses a generalized bathtub traffic model with three route-choice formulations—perfect rationality, logit-based stochastic choice, and bounded rationality—to propagate manipulated guidance into network accumulation, delay, and recovery. The threat model specifies attack intensity and distortion of perceived travel times, which alters realized trip lengths and feeds the macroscopic flow dynamics. Controlled experiments quantify peak congestion, time to recovery, total delay, and a resilience index across behavior regimes. Results show that behavioral diversity mitigates early cascades, lowers peak accumulation, and accelerates recovery under moderate attacks, while extreme attacks erode these gains and drive convergence of outcomes. The analysis also reveals an informational Braess effect: aggressive compliance with guidance can steer the network toward a worse equilibrium. The chapter contributes a behavior-aware simulation framework and threshold characterizations that link attack intensity and user heterogeneity to network vulnerability, offering actionable levers for resilient operations.

Chapter 6: Physics-Informed Neural Networks via Adversarial Training under Sensor Attacks

Chapter 6 develops a detection and mitigation approach for falsified traffic measurements by extending adaptive spatio-temporal domain-decomposed physics-informed neural networks. The method augments training with bounded adversarial perturbations and scripted attack scenarios, applies a robust Huber data loss with confidence weighting at each sensor, and performs residual-based anomaly detection and masking so that corrupted inputs do not dominate learning. The framework preserves conservation laws and interface compatibility by shifting trust from compromised data to governing physics during attacks. Evaluations on high-resolution highway data compare against conventional data-driven and standard PINN baselines using root-mean-square, demonstrating improved state estimation under targeted falsification while maintaining low false positives on genuine congestion. The chapter delivers a practical blueprint for securing traffic state estimation pipelines: it detects sensor attacks, contains their impact during training and inference, and maintains physically consistent reconstructions needed for resilient network management.

Each chapter builds upon the previous ones, ensuring a seamless transition from foundational concepts to advanced applications:

- **Chapter 2 to Chapter 3:** Chapter 2 establishes the PIML-GBM for traffic state estimation. Chapter 3 introduces AdaSTDPINN-TFM to address its limitations by capturing localized traffic variations, enhancing modeling capabilities for complexities in later chapters.
- **Chapter 3 to Chapter 4:** The advanced modeling from Chapter 3 enables precise traffic dynamics modeling essential for developing effective RGAs in Chapter 4, facilitating understanding and exploitation of traveler behavior.
- **Chapter 4 to Chapter 5:** Chapter 4's attack strategies provide context for Chapter 5's analysis of network failures and recovery, assessing the impact of cyber threats using prior models.

- **Chapter 5 to Chapter 6:** Insights into vulnerabilities from Chapter 5 inform the mitigation strategies in Chapter 6, crucial for designing effective detection and response mechanisms.

This dissertation presents a cohesive and comprehensive approach to enhancing the resilience of network traffic dynamics in CAVs against cyber threats. By integrating theoretical models with practical applications, the research contributes significantly to the development of safer, more efficient transportation systems. The progression from foundational modeling techniques to advanced attack strategies and mitigation measures exemplifies how each chapter synergizes with and supplements the others. This holistic approach ensures that the findings are robust, applicable, and valuable for the future integration of autonomous vehicle technology into daily life.