Cybersecurity risk assessment in connected intelligent systems for designing resilient systems

Data description

The main objective of this study is to focus on analyzing the vulnerability of cooperative driving relying on infrastructure-based communication from a real field experimental data collected at the Aberdeen center in Maryland. The research would emulate multiple sensor anomalies and cyberattack to assess the impacts of cyberattacks.

We will obtain, process and analyze multiple data sources in this project: 1) Real-world driving data for cooperative driving from a field experiment;; and 2) Cyberattack data generated from emulated cyberattacks, 3) Crash data managed by VDOT. The data will be stored and managed in distributed servers across CMU. The data engine offers organization, visualization and analytics of a wide array of mobility data.

Data format and metadata standards

While integrating data extracted from various entities, such as in-cabin sensors and kinematics and AV sensor data from cooperative driving, traffic flow data, traffic counts, traffic speed data, crash data etc., the raw data will first be converted to .py database file for cleansing, fusion and processing. The PI and research assistants will code algorithms using within python platform. The output data will mainly be system performance metrics. The aggregated data, without any personally identifiable information, can be provided in the standard .cvs format to any interested party.

Policies for access and sharing

The PI will communicate any significant findings with the scientific community in accordance with USDOT policy through journal publications, national and international conference presentations, and seminars throughout the project duration. The reported results will be made available to the research community upon request, where possible and permitted.

The PI commits to protect privacy, confidentiality, and security while sharing the data The PI will work with our respective Technological Transfer Offices to protect potential proprietary data if any research and project discoveries can be secured with intellectual property. Additionally, no raw data will be posted to any publicly available site, that violates the data usage agreement with private sector.

Policies for re-use, redistribution, derivatives

Data derived from this project shall be retained for at least one year. The data processing method developed in this project will be open source and shared along with research results to research community. No private or confidential information will be contained in the project data.

Plans for archiving and preservation

Any used data will reside on PCs and workstations belonging to the PI's university. External hard drives or centralized cloud will be used to back up all data periodically. This process would enable data recovery in the event of equipment failure.