THREAT AGENT DEFEAT MODELING AND TESTING

Harnessing The Power Of Machine Learning To Identify Novel Chemical Threats In Complex Forensic Environments

Carolyn Fisher LLNL **Colin Ponce** LLNL **Steven Magana-Zook** LLNL **Mark Dreyer** LLNL **Roald Leif** LLNL **Alex Vu** LLNL **Brian Mayer** LLNL **Audrey Williams** LLNL

Threat detection for chemical warfare agents (CWA) is necessary to identify use of weapons of mass destruction. The clandestine synthesis of CWA and use by state and non-state actors often aims to skirt regulations on chemicals or existing detection databases through continual development of novel analogs. The result is a never-ending arms race between nefarious synthetic chemists who make these chemical analogs and the chemists who must detect them. Samples are typically analyzed by liquid or gas chromatography (LC or GC) coupled to a mass spectrometer (MS). The resulting mass spectral data is then compared to databases of chemicals including CWA. However, these databases will always be incomplete for emergent CWA, and therefore will always be insufficient for reliable detection of novel CWA analogs. Manual identification of emerging CWA and illicit pharmaceuticals (e.g., fentanyl analogs) through MS is possible but not at the time scale required to agilely respond to national security needs. Our work towards developing Machine Learning (ML) methods to accelerate and enhance MS analysis to support national security will be presented.

Details on our work applying ML to perform binary classification to differentiate MS data of fentanyls (n = 250 types of fentanyls) and non-fentanyl chemicals (n = 440 structurally diverse pesticides) will be discussed. ML models have been built using both nominal mass GC-MS and high-resolution LC-MS instruments, with each dataset being independently used to develop ML models capable of differentiating fentanyls from non-fentanyls with >97% accuracy. A compelling finding was both nominal mass and high-resolution MS data performed equally well, suggesting that the ML methods were able to wholistically consider the “overall shape” of a spectrum to perform identification rather than focusing on a small number of high-precision measurements. The overall workflow, which includes the data processing, ML tools, and subsequent method development, will be described and represents the first demonstration of a broadly applicable approach for screening and identifying other classes of threat compounds including emergent CWAs, novel biotoxins, and explosives, all of which are relevant to the DTRA mission and the warfighter. This novel capability, initially developed around classification of compounds in the fentanyl family, advances threat agent identification capabilities and provides a means for rapid identification of novel CWA.