B155 The Utilization of Receiver Operator Curves (ROCs) for the Evaluation of Fire Debris: The Influence of Population Distributions on Classifier Performance

Alysaa Allen, National Center for Forensic Science, 12354 Research Parkway, #225, Orlando, FL 32826; Michael E. Sigman, PhD, University of Central Florida, National Center for Forensic Science, PO Box 162367, Orlando, FL 32816; Mary R. Williams, MS, University of Central Florida, National Center for Forensic Science, PO Box 162367, Orlando, FL 32816*

After attending this presentation, attendees will understand how diverse populations can alter the covariance matrices that affect the performance of classification modeling in fire debris analysis. An introduction to the use of subpopulation distributions and their influence on Likelihood Ratio (LR) calculations will also be explored.

This presentation will impact the forensic science community by discussing the development of "relevant" populations and their impact on the evaluation of fire debris evidence.

An important aspect in fire debris analysis is determining whether a sample contains an ignitable liquid residue. In this work, fire debris samples are assigned to two classes, those containing Ignitable Liquid (IL) residue and those containing only Substrate (SUB). One consideration in the development of classification models for fire debris is defining a population. A population is comprised of SUB and a subpopulation of IL classes drawn from the eight American Society for Testing and Materials (ASTM) E1618-14 IL classes in a defined distribution.[1] The challenge of defining a relevant population is attributed to the lack of known ground-truth ASTM E1618-14 IL class distributions in casework samples. This work examines the effect of diverse population distributions on LRs calculated from a one-level multivariate normal classification model.[2] The one-level classification model only takes into account between-sample variances. Diverse population distributions were generated computationally, as previously reported.[3]

Receiver Operating Characteristic (ROC) curves for model performance were generated based upon the results of LR calculations. Models were built based on each distribution and were tested on data from all other distributions. An Area Under the Curve (AUC) was calculated to evaluate classifier performance from the ROC curves. The change in the population distribution altered the mean and covariance of SUB and IL classes. Changing the covariance matrices modifies the LRs used to generate the ROC curves and assigns a different ordering of ground-truth labels. The population models that demonstrated the best performance consisted of significant SUB contribution and subpopulation comprised of representatives of each ASTM E1618-14 IL class. Population distribution models that contained a minimum SUB contribution and lacked representations from an ASTM E1618-14 IL class resulted in poorer performance.

This work was supported by an award from the National Institute of Justice, Office of Justice Programs. The content of this publication does not necessarily reflect the position or the policy of the government and no official endorsement should be inferred.

Reference(s):

1. International, A. ASTM E1618-14, Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas-Chromatography- Mass Spectrometry. ASTM International. West Conchohocken, PA, 2014.
2. Zadora G., Martyna A., Ramos D., Aitken C., *Statistical Analysis in Forensic Science: Evidential Value of Multivariate Physicochemical Data*. John Wiley & Sons, Ltd.: The Atrium, Southern Gate, Chichester, West Sussesex, United Kingdom, 2014.
3. Sigman, M.E.; Williams, M.R., Assessing evidentiary value in fire debris analysis by chemometric and likelihood ratio approaches. *Forensic Science International.* 2016, *264*, 113-121.

Fire Debris, Likelihood Ratio, Population Distribution

*Presenting Author