A12 Classification of Smokeless Powders by Cluster Analysis

Dana-Marie K. Dennis, BS, Michael E. Sigman, PhD, and Mary R. Williams, BS, Univ of Central Florida, NCFS, PO Box 162367, Orlando, FL 32816*

The goal of this presentation is to identify statistically valid sub-classifications of smokeless powders.

This presentation will impact the forensic community by providing a statistical basis for identification of smokeless powders beyond single- and double-base designations.

The purpose of this research is to determine if sub-classes (clusters) can be identified within two broad classes of smokeless powders, namely those that are single-based or double-based. Similarities between the total ion spectrum (TIS) were calculated and dendrograms of the similarity data were generated. Cophenetic correlation and inconsistency coefficients were calculated in order to determine the number of clusters which were present among the samples.

In 2009, the National Center for Forensic Science (NCFS) in collaboration with the Technical Working Group for Fire and Explosions (TWGFEX) developed a smokeless powders database, which consists of a compilation of analytical data for commercially available smokeless reloading powders. Methods utilized for analysis of these powders include stereomicroscopy, Fourier-Transform Infrared (FTIR) spectroscopy, and Gas Chromatography-Mass Spectrometry (GC/MS).

For this study, the GC/MS data for 87 smokeless powders analyzed in-house was exported as a condensed data file (CDF) format, and were used to generate the TIS. The TIS was generated by averaging the mass spectra across the entire chromatographic profile for each of the samples. The intensity values for m/z 43-400 were normalized to sum to one, and this data was used to calculate the similarities between the samples. Similarities that approach one indicate that the samples are more similar, while those that depart from one indicate that the samples are less similar.

Agglomerative hierarchical clustering was used to analyze the similarity data, and a study was done in order to determine the most appropriate distance (i.e., Euclidean, Manhattan, etc.) and linkage (i.e., Centroid, Complete, etc.) clustering method for further data analysis. The cophenetic correlation coefficient was calculated for each distance and linkage combination. Cophenetic correlation coefficients closer to one indicate that the clusters in the data, as defined by the dendrograms, closely reflect distances between objects in the original data. Agglomerative hierarchical clustering allows for the determination of natural divisions that may be present within data. This method initially assigns each object to its own cluster and continues by merging objects and/or clusters together until all samples are merged into one cluster. Natural divisions (i.e., optimal clusters) are determined based on the inconsistency coefficient. The inconsistency coefficient compares the link height in a given cluster hierarchy with the average of the link heights directly below it. A link that is approximately the same height as the links below it is said to be consistent, indicating that there are no natural divisions within the data. On the other hand, a link that has a significantly different height from that of the links below it is said to be inconsistent, which may indicate natural divisions within the data.

Preliminary results indicate that average linkage and correlation distance, as well as centroid linkage and correlation distance, methods are the most suitable for analysis of the smokeless powders data. The cophenetic correlation coefficient in both cases was 0.96, and the data clustered into two primary groups, namely single-based and double-based. In addition, there are a number of sub-clusters apparent within the two main groups. Some sub-clusters correspond primarily to a single manufacturer, while others are predominantly populated by powder kernels of a given shape. Additional investigations are underway to better understand the chemical nature of these sub-clusters.

Clusters, Similarity, Dendrograms

 ** Presenting Author*