

Medical Surveillance, Frequent Sets, and Closure Spaces

Alan P. Sprague
University of Alabama at Birmingham

A task that hospitals must perform is to monitor the bacteria residing there, to detect increasing resistance to antibiotics. To aid in this monitoring, we developed a computer program. When processing data describing a clone of multiply resistant bacteria, the program experienced a combinatorial explosion in time and space.

Computation of frequent sets, a concept which is often useful in Data Mining applications, is an important step in processing. The input file is a set of records, each describing a bacterial isolate. Every record in this file is a list of items pertaining to that isolate. A frequent set is a set of items which co-occur in at least some preset threshold of records. It was while generating frequent sets that the combinatorial explosion occurred.

This talk will be centered on a modified algorithm to generate frequent sets, whose proof of correctness involves the notions of closure, independent sets, and circuits of closure spaces.