Documentation

1 Variables

- **bins_count:** for each bin size it counts how many bins have been created ;
- **chi_tests_count:** for each bin size, it counts how many χ^2 tests were negative and how many χ^2 tests have not been performed due to insufficient data;
- **flags:** for each bin size, it indicates the number of the message at the end of a flagged bin, and the number of the bin;
- **gap_vector:** it sets the bin sizes;
- **messages_count:** it counts how many syslog messages have been analyzed;
- **M_T:** message number-template vector;
- **new_message:** new syslog message to be analyzed;

2 Functions

- **analyze_bins_and_write:** given a bin it reports why it was marked as anomalous indicating unique or rare templates;
- **anomalous_bins_frequencies:** it finds the template distributions of given bins of messages;
- **determine_anomalous_bins_parents:** for the messages that belong to anomalous bins for each bin size (child bins) it finds the distribution of such anomalous bins (parent bins);
- **determine_child_bins:** once the anomalous parent bins have been determined, it finds messages that belong to bins with anomalous parent distribution;
- **find_overlapping_indeces_parents:** finds messages that belong to anomalous bins for each bin size (child bins);
- **find_template:** it matches the new message with a previously observed template or it creates a new template;
- **sliding_window_flags_eff_new:** for each bin it compares the distribution of the templates with previous bins and it keeps track of the bin distributions found so far;
- **reindexing_bins_cluster:** it gives different indexes for each type of distribution so they are easily distinguished;
- **set_log_files:** it sets the log files to be analyzed;