Training Dataset

Raw Log Events

Feature Generation

Semantic Similarity

Feature Vector

Normalization

Preprocessing

Filtering Stop-Words

Filtering Symbols

Fig. 1 Flowchart of Data Preprocessing and Feature Generation