Feature Generation

Semantic Similarity

Feature Vector

Normalization

Training Dataset

Preprocessing

Filtering Stop-Words

Filtering Symbols

Raw Log Events

Fig. 1 Flowchart of Data Preprocessing and Feature Generation