

ABSTRACT

The purpose of the present invention is to generate a dictionary for monitoring text information such that it is possible to achieve high-precision detection compared to conventional art. A feature degree calculation unit 3 compares the statistics of a positive example group and a negative example group, and calculates the degree by which a phase of interest appears in the positive example group as the feature degree. A usefulness degree calculation unit 21 calculates a usefulness degree by using the length of the phrase, the frequency at which the phrase appears within the positive example group, and an index pertaining to an inclusion relationship between phrases for each phrase extracted by means of a phrase extraction unit 1. A detection condition determination unit 22 uses the usefulness degree calculated by means of the usefulness degree calculation unit 21 and the feature degree calculated by means of the feature degree calculation unit 3 to evaluate the appropriateness of each phrase as a detection condition by means of the product of the usefulness degree and the feature degree, and determines that the phrase is appropriate for a detection condition when the value of the product is greater than a threshold value.