

B0001: Index out of range

Bug ID: B0001

Title: Index out of range

Vender: ZHU Renjie	Product: News Clustering System	Version: 1.0
Severity: Critical	Updated: 2014-10-14	OS: Windows 7
Status: Closed	Assigned to: Fang Zhou	

Description:

When creating hash table based on input files, the error message saying that index of range will prompt out.

Steps to Reproduce::

- 1.) Prepare the articles under raw_data directory
- 2.) Start the program

Expected result:

A word count file will be built as a dictionary for each category in unsorted order.

Actual result:

The error message saying that index of range will prompt out.

Fang Zhou 2014-10-15 19:12 EDT

The cause of this problem is that when using getWords in other classes, we loop all the contents in words array from 0 to words.length. However, this would case IndexOutOfRangeException Exception since the last several elements may be null. So, I add a null check before reading words into array.

Before

```
private final int MAX_WORDS_NUMBER = 20000;
private String[] words;
private int index;

private void copyArray(String[] s) {
    for(int i=0;i<s.length;i++)
    {
        if(!isInteger(s[i])&&s[i].length()!=0)
        {
            words[index]=s[i];
            index++;
        }
    }
}

public String [] getWords() {
    return this.words;
}
```

After

```
private ArrayList<String> words;

private void arrayCopy(String [] arr){
    for(String token: arr){
        if(!token.trim().isEmpty() && token.matches("[a-zA-Z- _]+"))
            this.words.add(token.toLowerCase()); //all words in lower case
    }
}

public String [] getWords() {
    return this.words.toArray(new String[words.size()]);
}
```

Zhu Renjie 2014-10-16 20:40 EDT

Test succeeds. No IndexOutOfRangeException appears.