

Implementation of Boyer Moore Algorithm to Search Criminal Law Based on Criminal Code

A. Hanifa Setianingrum
Syarif Hidayatullah State Islamic University
Jakarta, Indonesia
anif.hanifa@uinjkt.ac.id

Tabah Rosydi
Syarif Hidayatullah State Islamic University
Jakarta, Indonesia
TabahRosydi@gmail.com

Reza Hermawan
Syarif Hidayatullah State Islamic University
Jakarta, Indonesia
rezherma@gmail.com

Abstract—Balai Pemasyarakatan (BAPAS) or Penitentiary is one of the institutions to implement guidances. The Penitentiary (“Badan Pemasyarakatan in Indonesia. BAPAS”) has a role to make ex Prisoner are not committing a crime or a similar offence. KUHP is to provide knowledge of the articles of the Criminal Code. KUHP (Kitab Undang-Undang Hukum Pidana) consist of many kinds of violations of law such as crimes and abuse, etc. The aim of those guidelines is to provide counselling a deterrent effect to the Ex-Prisoner or “Klien Pemasyarakatan”. Information retrieval is a material document search form or another that is not structured to meet the needs of information in a collection of large text that is stored in the computer. Boyer Moore algorithm is an algorithm that does the matching characters from the rightmost position to the left. The results of Boyer Moore algorithm that is implemented into the search clauses of a felony offence under the Criminal Code and provides quick results. Speed in terms of doing a keyword search of article Clients criminal offences committed in total 422.194 characters and 1.459 chapters with proven character in the search for the worst and best quickly in few minutes running time.

Keywords: *Boyer Moore algorithm, information retrieval, Criminal Code*

I. INTRODUCTION

The rise of a felony and violations committed at this time caused not least the people who tangled case law. Various crimes and offences contained in the Criminal Code dealt with through the stages of law in Indonesia. BAPAS is one of the institutions that interact directly with offenders serving a sentence in West Jakarta.

Approaches to Client is one of the tasks of BAPAS. The approach in question is the provision of guidance or education about crime or offence against the law. With article describes a similar or related to a crime or offense committed client, is expected to provide a deterrent effect against him.

Criminal Code is a book of guidelines for the implementation of criminal law in a country, all issues regarding criminal offences contained therein (Ali 2015). In the Criminal Code there are many articles and some of them there are changes over time.

Although the client at the hearing made when found guilty by a court already knows that he violated article, BAPAS felt the need to be educated about the additional articles of the Criminal Code.

Information retrieval is finding material such as documents that are unstructured text that satisfy the information needs of the large collections which is stored in the computer. (D., Raghavan, & Schütze, 2009).

Every day people are always dealing with data in the form of various kinds, but the data in the form of text is the most common data used for the exchange of information and other benefits. Often operating in the text search process involves the emergence of a string, and location of the text in question. (Sulistyo, Pradipto, & Perwira, 2006).

From some search algorithms are the advantages and disadvantages of each. In the longer Brute Force algorithm pattern search a fixed search time and the algorithm Knuth Morris Pratt sought the longer pattern search time tends to increase, while the Boyer Moore algorithm the longer pattern search a shorter search time (Utomo, Harjo, & Handoko, 2008). From some research related to Boyer Moore above, the authors apply the Boyer Moore algorithm in finding the article of criminal act to support the system in the search string to the article.

Therefore the researchers apply a string search algorithm with the Boyer Moore algorithm for the search process a crime and violation of article according Penal Code. Boyer Moore used because it can do a search string efficiently (Ali, 2015). Because of the efficiencies expected to support web-based system. Web-based because according to the needs of BAPAS Barat to publish matters related to the general need for the clients or the general public regarding the BAPAS.

II. LITERATURE REVIEW

A. Boyer Moore Algorithm

In 1977, Robert Stephen Boyer (Bob Boyer), and J. Strother Moore made an algorithm that does the matching characters from the rightmost position to the

left (rightmost characters in the pattern is the first character to be matched with the characters in the text). Such as statements of Boyer and Moore (1997), is “*The basic idea behind the algorithm is that more information is gained by matching the pattern from the right than from the left*”. (Faradita, 2014)

Boyer Moore uses two heuristics to decide how far to jump: Bad Character Heuristic Heuristic often referred Occurrence, and heuristic Good Character called Heuristic Match. Heuristic information for each maintained in an array that was built at the beginning of the operation from which. (Orwant, Heitaniemi, & Macdonald, 1999).

Systematically, the steps undertaken Boyer Moore algorithm when matching string is: (Sarno, Anistyasari, & Fitri, 2012).

- 1) Boyer Moore algorithm started matching pattern at the beginning of the text.
- 2) From right to left, this algorithm will match the character-by-character pattern with the character in the corresponding text, until one of the following conditions are satisfied:
 - The characters in the pattern and in the *text* being compared do not match (mismatch).
 - All the characters in the pattern matched then the algorithm will inform the invention in this position.
- 3) The algorithm then shift pattern to maximize the value of panning Good Suffix shift Bad Character Heuristics and Heuristics, and then repeat steps 2 through pattern at the end of the text.

B. The Procedure on Boyer Moore Algorithm

By using this algorithm average string search process will be faster than the other algorithms. Here's how the Boyer Moore:

D	A	T	A		K	L	I	E	N
K	L	I	E	N					

In the example above, with comparisons of the position of the most end pattern can be seen that the character "N" on the string "KLIEN" does not fit with the character "space/empty" because it is the distance between the string "DATA" and "KLIEN", and "space/empty" was not found in the string "KLIEN" who sought to pattern "KLIEN" can be shifted past the string "DATA" and "space/empty" so that its position be:

D	A	T	A		K	L	I	E	N
					K	L	I	E	N

In this example it appears that the Boyer Moore algorithm has great character stepping up the search pattern so as to examine a little character can be known that the search string is not found and can be shifted to the next position.

Boyer Moore algorithm has been proven as one of the most efficient algorithms in applications string search using natural language. This algorithm has been frequently implemented for the function "Search" and "Substitute" in the text editor.

III. METHOD

Collecting data on observations carried out with a direct view of procurement processes and activities that run on Penitentiary in Jakarta. Observation was conducted to determine the process or workflow applied.

IV. RESULTS AND DISCUSSION

A. Design Algorithm Model

As described in the previous chapter design modelling using the Boyer Moore algorithm to be implemented on a chapter search process in the application. Next will be explained the workings of the algorithm is implemented in this study:

- 1) The first thing to do is application will read the data on the search form Article/Case. The application will perform a search of the array that contains data from Article/Case entered. Sample text: “ketentuan pidana”

T = ketentuan pidana

N (total) = 16

- 2) Then, the application captures the character or string as a keyword entered by the user, in this discussion is referred to as pattern. Example *pattern*: pidana

P = pidana

M (total) = 6

- 3) Create a pattern shift value is sought by the Match Heuristic approach (MH) and Occurrence Heuristic (OH) to determine the number of shifts to be done if it gets the characters do not fit in the matching process with the string in the text. Here is a picture in the table:

<i>Pattern</i>	p	i	d	a	n	a
<i>Occurence Heuristic</i>	5	4	3	0	1	0
<i>Match Heuristic</i>	5	4	3	1	1	1

OH obtained by breaking words into in word and used the array by the system. In the source code, step is conducted in make CharTable function. Where a

criminal stored last word with a value of 0. Then n stored with a value of 1. And further that should have a value of 2, but returned with a value of 0 because it was read at the beginning as a value of 0, is due in accordance uses Boyer Moore algorithm that searches of said the far right. Then d is stored with a value of 3, i saved with a value of 4, and p is stored with a value of 5.

MH which determines the position ic in step displacement. MH in the system is determined by calculation $\max(\$ta$ so in get:

$$p = \max(5,1) = 5$$

$$i = \max(4,1) = 4$$

$$d = \max(3,1) = 3$$

$$a = \max(0,1) = 1$$

$$n = \max(1,1) = 1$$

$$a = \max(0,1) = 1$$

$\$stable[\$text[i]]$ is scoring OH. While the number 1 is used as a shift of at least done one time if the rightmost character is found.

4) Then the pattern (P) aligned with the leftmost text, so that the first character of the pattern aligned vertically with the first characters of text (T). I and J are used to transfer.

5) In the first step $i=5$ in getting the character 'a' in parallel to the pattern character 't' in the text. 'a' \neq 't';

Characters doesn't match, $\$i += \$patlen; 5 + 6 = 11$

$\$patlen$ is a long pattern, pattern of criminal has a length of 6 characters.

6) Next step $i = 11$ in getting the character 'a' in parallel to the pattern character 'i' in the text

'a' \neq 'i';

But the characters do not match the character 'i' contained in the pattern then performed formula $\$i += \max(4,1); 11 + 4 = 15$

Value 4 is taken as the MH value of the characters 'i' is 4.

7) Next step $i = 15$ in getting the character 'a' in parallel to the pattern character 'a' in the text.

'a' = 'a';

Suitable character, the character that matches the run loop $i--; 15 - 1 = 14; dan j--; 5 - 1 = 4;$

Note, focus to I and J in the next step.

8) In the next step $i = 14$ dan $j = 4$ in getting the character 'n' in parallel to the pattern character 'n' in the text.

'n' = 'n';

Suitable character, the character that matches the run

loop $i--; 14 - 1 = 13; dan j--; 4 - 1 = 3;$

9) Next step $i = 13$ dan $j = 3$ in getting the character 'a' in pattern aligned with the character 'a' in the text.

'a' = 'a';

Suitable character, the character that matches the run loop $i--; 13 - 1 = 12; dan j--; 3 - 1 = 2;$

10) In the next step $i = 12$ and $j=2$ in get character 'd' on pattern aligned with the character 'd' in the text.

'd' = 'd';

Suitable character, the character that matches the run loop $i--; 12 - 1 = 11; dan j--; 2 - 1 = 1;$

11) In the next step $i = 11$ dan $j = 1$ in get character 'i' aligned with the character 'i' in the text.

'i' = 'i';

Suitable character, the character that matches the run loop $i--; 11 - 1 = 10; dan j--; 1 - 1 = 0;$

12) In the next step $i = 10$ dan $j = 0$ in get character 'p' the pattern aligned with character 'p' in the text.

'p' = 'p';

The characters fit and pattern have been met, then the search is complete characters fit and pattern have been meet, then the search is completed.

B. Algorithm Boyer Moore Pseudocode

There is an overview of the search pseudocode Article/Case

on application:

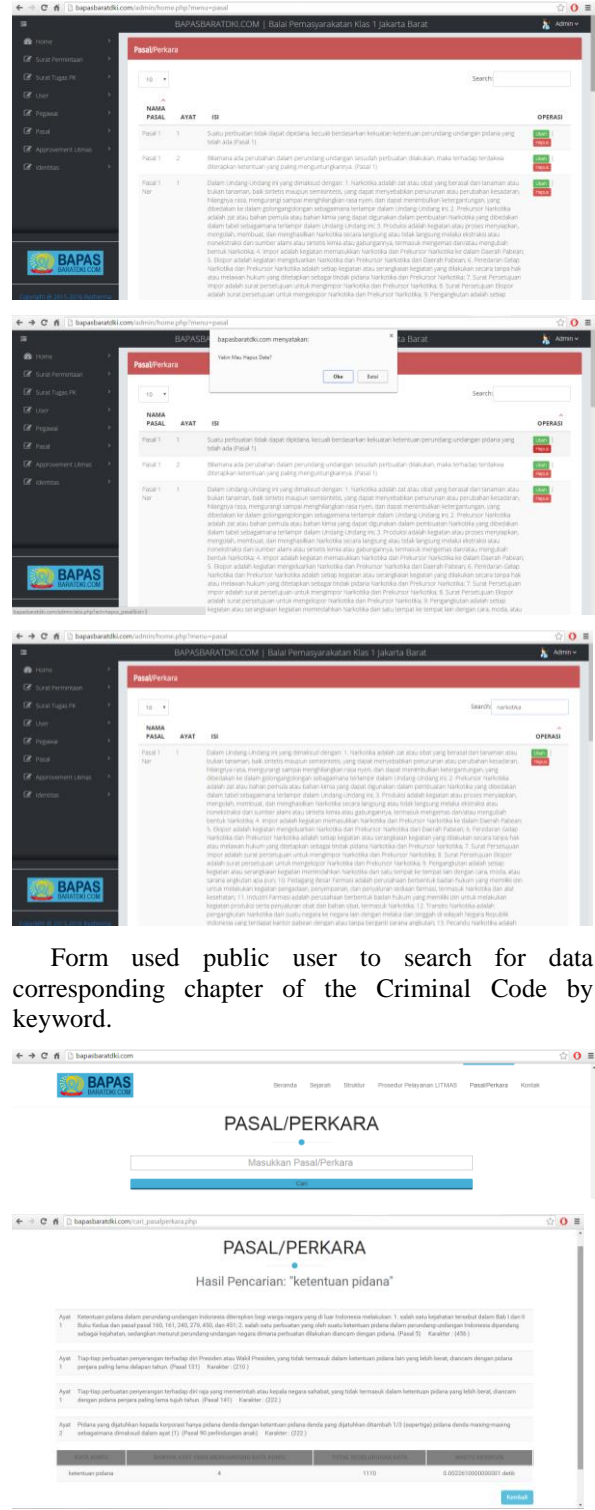
```

Function BoyerMoore(
input text      : array of char
input pattern   : array of char
)
Deklarasi:
table          : array of char
patlen, textlen, t, i      : integer
Algoritma:
patlen := STRLEN(pattern) //returns the length of
a string
textlen := STRLEN(text) //returns the length of a
string
table := makeCharTable(pattern) // makes the
jump table based on the mismatched character
information
for(i := patlen-1 to textlen - 1) LOOP1
    t := i;
    for(j:=patlen-1 to pattern[i] = text[i])
        LOOP2
            if(j = 0)
                return i;
            endif
        endFor
        i := t
        if(array_key_exist(text[i], table)) //checks if
the given key or index exist in the array
            i := i + max(table[text[i]], 1)
//returns the highest value in array
        endif
    else
        i += patlen
    endElse
endFor
return -1
endFunction

Function makeCharTable(
input string : array of char
)
Deklarasi:
len, k      : integer
table : array of integer
Algoritma:
len := STRLEN(string)
For(k = 0 to len - 1)
    table[string[k]] = len - k - 1
endFor
return table
endFunction
    
```

C. Design Interfaces

Form used the admin user to log into the application. And then the form used to change, add, delete, and search for data on the access provisions to be user admin.



Form used public user to search for data corresponding chapter of the Criminal Code by keyword.

V. CONCLUSIONS

Boyer Moore algorithm can be implemented into the search clause acts felony and misdemeanor cases under the penal law properly. Can display the results of the search process quickly and precisely in the search for articles. Boyer Moore algorithm is able to provide the search speed for a few seconds in total 422.194 Characters and 1.459 clausal of the content of article. Where a keyword search more quickly if the condition gets long pattern or total character input.

Boyer Moore algorithm, a search of articles in the Criminal Code to be faster to search than search for it using the Criminal Codebook.

Applications that make the author of course still not perfect, there are still many things that can be developed in order to make the benefits of the application to be even better for the future. Therefore, the authors also given some suggestions for further development, is.

- 1) Applications can be added features such as presentation and data search with auto text, auto-complete, as well as bookmarks article.
- 2) Security applications need to be considered for future development because the application is based online, the necessary security.

REFERENCES

- [1] Jingbo Yuan, Jisen Zheng, Shunli Ding, "An Improved Pattern Matching Algorithm", 978-0-7695-4020-7/10 2010 IEEE DOI 10.1109/IITSI.2010.73.
- [2] Yuting Han, Guoai Xu, "Improved Algorithm of Pattern Matching based on BMHS", 978-1-4244-6943-7/10 2010 IEEE.
- [3] Lin quan Xie, Xiao ming Liu and Guangxue Yue, "Improved Pattern Matching Algorithm of BMHS", 978-0-7695-4360-4/10 2010 IEEE DOI 10.1109/ISISE.2010.154
- [4] Zhengda Xiong, "A Composite Boyer-Moore Algorithm for the String Matching Problem", 978-0-7695-4287-4/10 2010 IEEE, DOI 10.1109/PDCAT.2010.58
- [5] http://en.wikipedia.org/wiki/Boyer_Moore_string_search_algorithm.
- [6] Horspool R.N. "Practical Fast Searching in Strings". Software Practice and Experience, 1980, 10(6):501-506.
- [7] Zhen LIU, Su XU, Jue ZHANG, "Improved Algorithm of pattern matching for Instusion Detection". 2009 International Conference on Multimedia Infoemation Networkong and Security, Wuhan, CHINA, 2009, pp.446-449
- [8] Thierry Lecroq. Fast exact string matching algorithms [J]. Information Processing Letters, 2007, 6(102):229-235.
- [9] Chuanhan Liu, PYongcheng Wang, PDerong Liu, et al. Two Improved Single Pattern Matching Algorithms [C]//Proceedings of the 16th International Conference on Artificial Reality and Telexistence- Workshops, 2006:419-422.
- [10] Yihui SHAN, Yuming JIANG, Shiyuan TIAN, "Improved Pattern Matching Algorithm of BMHS for Intrusion Detection". Computer Engineering, vol.35, 2009, pp.170-173.
- [11] Charras, Christian, danLecroq, Thierry. 1997. Boyer Moore algorithm. [online]. (<http://www-igm.univ-mlv.fr/~lecroq/string/node14.html>)