Abstrak

Sistem pendeteksi plagiat sangat dibutuhkan di berbagai bidang. Salah satunya yang marak saat ini, sistem tersebut digunakan untuk mengindikasi dokumen plagiat dalam jumlah besar (*massive*). Pada Sistem tersebut, proses indikasi dokumen menggunakan metode *string matching* dengan mencocokan tiap dokumen input dengan dokumen *corpus*. Pada proses tersebut tentunya dokumen yang tidak relevan pun ada kemungkinan ikut dicocokkan, hal ini menyebabkan sistem berjalan kurang efektif. Dapat dilihat bahwa pencocokan dokumen dalam jumlah besar (*massive*) pada sistem pendeteksi plagiat tentunya akan membutuhkan waktu komputasi yang besar sebanding dengan besar jumlah dokumen. Hal ini tentunya perlu ada pendekatan metode untuk mereduksi dokumen yang tidak relevan dan perlu ada proses pencarian yang cepat di dalam sistem pendeteksi plagiat. Struktur data *indexing* menggunakan tabel *hash* multi-level dengan fungsi *hash* FNV-1a dapat diimplementasikan untuk proses pencarian cepat. Pada struktur data tersebut juga digunakan nilai *threshold* untuk mengurangi dokumen yang tidak relevan. Hasil penerapan *indexing* menggunakan tabel *hash* multi-level memberikan performansi yang stabil dalam pencarian similaritas dokumen. Implementasi proses *indexing* pada proses LCS dapat mempercepat sistem pendeteksi plagiat secara signifikan, yaitu sebesar 84.14% menggunakan nilai threshold 0.25 dengan metode pencarian similaritas simetrik pada dokumen.

Kata kunci: *Indexing*, FNV-1a, Tabel *Hash* Multi-Level, Sistem Pendeteksi Plagiat.