ABSTRAK

Saat ini web usage mining adalah salah satu topik dimana analisis data merupakan hal paling penting untuk menemukan pola-pola perilaku pengguna akses internet. Hal ini akan membantu desainer untuk meningkatkan situs web agar menarik pengunjung, membuat situs web adaptif, mendukung bisnis dan layanan, mempersonalisasi / memprofiling, menganalisis trafik jaringan, mendesain sistem cache untuk meningkatkan kualitas pengalaman pegguna internet, dan lain-lain. Fase terpenting dari web usage mining adalah rekonstruksi sesi pengguna dengan menemukan pola yang berguna dengan menggunakan teknik penemuan pola seperti Association rule mining, misalnya algoritma Apriori, algoritma Frequent Pattern, dan lain-lain.

Pada web usage mining, aturan asosiasi yang digunakan untuk menemukan halaman yang dikunjungi bersama-sama cukup sering. Aturan asosiasi digunakan untuk menemukan hubungan antara atribut dari item set. Item set adalah kumpulan dari halaman situs. Aturan yang diterapkan untuk membedakan halaman yang sering muncul untuk mengungkap hubungan antara kelompok pengguna dengan kepentingan tertentu dan untuk menentukan situs yang paling disukai. Informasi tersebut dapat digunakan oleh perusahaan untuk memperluas sistem cache dan memberikan kualitas yang lebih baik untuk pelanggan. Kelemahan dari aturan asosiasi adalah terlalu banyak aturan yang dihasilkan dan tidak ada jaminan bahwa aturan tersebut relevan. Banyak algoritma yang ada untuk menghasilkan pola perilaku pengguna kurang efisien dalam waktu eksekusi dan kebutuhan memori.

Penelitian ini terdiri dari beberapa tahap antara lain pengumpulan data, pembersihan data, identifikasi pengguna, identifikasi sesi, teknik asosiasi (membangun Frequent Pattern-Tree (FP-tree), menghitung information gain, menemukan item yang sering muncul dengan algoritma Frequent Pattern-Growth (FP-Growth)). Dengan improvisasi Association Rule Mining dengan menambahkan information gain, sistem ini membutuhkan waktu eksekusi lebih sedikit 1.34% dan memperoleh jumlah rule lebih sedikit 40-50% dibandingkan dengan metode sebelumnya.

Kata kunci: Web Usage Mining, Association Rule Mining, FP-tree, FP-Growth, Information Gain