

## INTISARI

### **PENGEMBANGAN KOMPONEN PERANTARA SISTEM VISUALISASI GRAF HASIL TRANSFORMASI DATA BRT MENGGUNAKAN ALGORITMA MULTIPLEX VISIBILITY GRAPH PADA GRAPHX**

Afif Izzul Falakh  
13/347531/PA/15289

Visualisasi data dalam bentuk graf dapat digunakan untuk menganalisis pola data dan mengeksplorasi data untuk bisa mendapatkan pengetahuan baru. Visualisasi graf tersebut dapat diterapkan pada analisis halte *Bus Rapid Transit* (BRT) untuk menemukan pola penggunaan halte dari waktu ke waktu maupun untuk menemukan informasi dengan mengeksplorasi relasi antarhalte. Namun karena ukuran data BRT yang besar, diperlukan mesin pengolahan graf skala besar seperti GraphX yang merupakan bagian dari mesin pengolahan big data Spark. Dengan menggunakan Spark, data BRT yang merupakan data deret waktu ditransformasikan menjadi graf lalu dimuat dalam GraphX. Kemudian data graf dalam GraphX tersebut diekspor menjadi berkas untuk dapat divisualisasikan menggunakan program visualisasi di luar lingkungan pengolahan data Spark.

Berdasarkan latar belakang tersebut, dalam penelitian ini diusulkan penggunaan komponen perantara berupa HTTP *server* yang tertanam pada kode program untuk bisa menyediakan data graf secara langsung kepada antarmuka visualisasi yang berada di luar sistem Spark. Dengan demikian, tidak diperlukan proses ekspor data graf menjadi berkas. Sehingga menyingkat proses yang diperlukan untuk memvisualisasikan graf dan memungkinkan interaksi secara dinamis antara antarmuka visualisasi dengan GraphX.

Dari hasil pengujian didapatkan bahwa algoritma MVG dapat digunakan untuk mentransformasikan data BRT menjadi graf dengan beberapa langkah prapemrosesan untuk menyiapkan data BRT. Pertambahan waktu transformasi menggunakan algoritma MVG juga relatif konstan terhadap penambahan jumlah data olah dengan rata-rata pertambahan waktu transformasi sebesar 18,21 detik tiap penambahan 5.500.000 baris data olah. Kemudian didapatkan juga bahwa penyediaan data siap visualisasi menggunakan komponen perantara sedikit lebih cepat namun lebih stabil dibandingkan tanpa menggunakan komponen perantara, dengan rata-rata pertambahan waktu penyediaan data dengan komponen perantara sebesar 1 detik dan standar deviasi sebesar 1,89 dibandingkan tanpa menggunakan komponen perantara sebesar 1,27 detik dan standar deviasi sebesar 6,49.

Kata kunci : transformasi data, data deret waktu, graf, komponen perantara, HTTP *server*, GraphX, Spark.

## **ABSTRACT**

### **MIDDLEWARE DEVELOPMENT FOR VISUALIZATION SYSTEM OF GRAPH RESULTED FROM TRANSFORMING BRT DATA USING MULTIPLEX VISIBILITY GRAPH ALGORITHM ON GRAPHX**

Afif Izzul Falakh  
13/347531/PA/15289

Graph visualization can be used to analyze data pattern and to explore the data to gain new insight. This graph visualization can be implemented on Bus Rapid Transit (BRT) shelter analysis to find the dynamic usage pattern of the shelter from time to time and to gain new information by exploring the relation of each shelter. But, because the sheer size of the BRT data, a large scale graph processing engine such as GraphX which is a part of big data processing engine Spark is needed. Using Spark, BRT data which is a time series data transformed into graph and then loaded into GraphX. Then, the graph data inside GraphX is exported into files which can be used to build the visualization using external visualization program outside Spark environment.

Based on that background, this research proposed the use of middleware in the form of HTTP server, which is embedded inside the program code, to serve the graph data directly to the visualization interface. Thus, no need to export the graph data into files. so that the process to visualize the graph can be shortened and enabling a dynamic interaction between visualization and GraphX.

From the evaluation result, it is found that MVG algorithm can be used to transform BRT data into graph by performing some preprocessing steps to prepare the BRT data. The time needed to transform the data using MVG algorithm also shows a relatively constant increase along the increasing amount of processed data with an average increase of 18.21 seconds for each 5,500,000 additional data rows. It is also proved that the processing time needed to serve visualization ready data using middleware is a bit faster but much stable than without the use of middleware with the average increase of processing time using middleware comes to 1 second with standard deviation of 1.89 compared to the average increase of processing time without the use of middleware which comes to 1.27 seconds with standard deviation of 6.49.

Keywords : data transformation, time series data, graph, middleware, HTTP server, GraphX, Spark