

In this section, we evaluate our algorithms on two datasets, a set of synthetic graphs and a set of graphs modeling active ingredients in antiviral drugs for the treatment of AIDS. We call the former the synthetic graph dataset and the later the real graph dataset. We perform two sets of tests; In the first set, we evaluate our algorithms on the retrieval of induced subgraph to query and In the second set, we evaluate retrieval of non induced sugraph to query. To test the efficacy of our algorithm with regard to induced subgraph query, we evaluate the proposed algorithm in comparison with Messmer’s Network Algorithm and a sequential SCAN using the state-of-the-art subgraph isomorphism detection algorithm VF2cordella2001\_vf2. *We also evaluate our algorithm with regard to non-induced subgraph query, where we compare proposed algorithm with programming language and run on a Intel Core i-3 CPU 3.09GHZ, 16Gbyte memory, Personal Computer running Windows 7.*

**Graph Datasets** We evaluate our algorithms by processing the retrieval of descriptors from the compounds dataset. We use AIDS Antiviral Screen dataset to provide a real graph dataset. This dataset contains around 43,000 chemical compounds and is available publicly from NCI National Cancer Institute <http://dtp.nci.nih.gov/> We denote this dataset as “AIDS” in our experiment. The graphs of AIDS have an average number of 25 vertices and 27 edges, a maximum of 438 vertices and 441 edges, 63 distinct vertex labels and 3 distinct edge labels.

In order to evaluate our algorithms over a larger database to test scalability, we use synthetic large dataset. The graph generator is configured to emit only connected graphs that have an edge probability of 50 percent. Except where explicitly stated, 10 distinct vertex labels and 10 distinct.

**Chemical Descriptor Search** In chemistry, substructures of compounds that imply a chemical, physical property of the compound are called as descriptor. Fast retrievals of descriptors from compounds aids research of compounds. We evaluate our algorithm by processing the retrievals of descriptors. We build a model graph database  $D$  by extracting graph database  $W$  composed of 10,000 compounds whose size are less than 40 from AIDS and applying frequent graph mining to  $W$ . We set minimum support as 5 percents.  $D$  is composed of 18930 distinctive frequent subgraphs.

We vary average size of query graphs and evaluate query processing time. For each plot, 100 query graphs are extracted from  $W$ . Query processing time is time to process the all query graphs and the time to construct the DAG is excluded.

Fig.fig:fig3 shows the processing time for induced subgraph isomorphism query. Fig.fig:fig3 indicates that proposed algorithm has no improvements from Messmer et al.’s algorithm. It is because that there are originally few decompositions that creates disconnected graphs and also few redundant subgraph isomorphism detections in DAG. So in this experiment, our improvements did not work.

Fig.fig:fig4 shows the processing time for subgraph isomorphism query. Fig.fig:fig4 indicates the subgraph isomorphism query is faster than SCAN same as Messmer et al’s algorithm.

**Synthetic Data Search** In chemical descriptor search, we can not show improvements of proposed algorithm in term of processing time. In order to demonstrate improvements of proposed algorithm, we process subgraph isomorphism query on synthetic data.

We vary average size of query graphs and evaluate query processing time. Model graph dataset is composed of 20,000 graphs whose average number of size are 10. 100 query graphs are generated for each plot. *fig-5 shows the plot of Processing Time for Induced Subgraph Isomorphism Query for Synthetic Data fig : fig5*