

INTISARI

Schema matching merupakan solusi penting pada tahap awal integrasi informasi dari sumber basis data yang heterogen. Secara umum, proses *schema matching* adalah menerima masukan berupa masukan dua *schema* dan/atau informasi tambahan, mencocokkan kesamaan pasangan atribut, menghasilkan luaran berupa pemetaan kesamaan pasangan atribut, dan verifikasi oleh pengguna untuk menentukan apakah hasil pemetaan kesamaan pasangan atribut yang diperoleh sudah valid atau masih perlu direvisi. Permasalahan *schema matching* adalah kesulitan dalam menentukan pemetaan kesamaan pasangan atribut secara manual sehingga diperlukan model semi otomatis dan lebih efektif. Solusi untuk permasalahan tersebut adalah mengembangkan model berbasis komputer dengan cara menggabungkan metode-metode yang telah ada.

Dalam penelitian ini dikembangkan model dan prototipe *schema matching* gabungan *constraint-based method* dan *instance-based method* yang dilaksanakan secara *hybrid*. Model yang diperoleh kemudian dimodifikasi untuk meningkatkan efektivitas hasil dan efisiensi proses. Model tersusun atas modul-modul, meliputi 1) pembacaan masukan pasangan basis data sumber (DBSource) dan dan basis data target (DBTarget), 2) ekstraksi basis data, *constraint*, dan *instance*, 3) pencocokan, perhitungan nilai kesamaan pasangan atribut (SIM), dan penentuan pasangan atribut yang cocok, 4) verifikasi oleh pengguna, serta 5) pengujian efektivitas model. Efektivitas model diukur menggunakan parameter *precision* (P), *recall* (R), dan *f-measure* (F). Pengujian model dilakukan menggunakan basis data relasional.

Berdasarkan hasil penelitian diketahui bahwa model *schema matching* gabungan *constraint-based method* dan *instance-based method* lebih efektif dibandingkan dengan *constraint-based method* atau *instance-based method* secara individual. Model yang dikembangkan dapat diterapkan untuk proses *schema matching* pada basis data relasional dengan DBMS dan/atau domain aplikasi yang berbeda. Model yang dikembangkan juga dapat diterapkan baik secara gabungan maupun secara individual pada masing-masing metode penyusunnya.

Kata kunci: *constraint-based method, instance-based method, schema matching.*

ABSTRACT

Schema matching is an important solution on the preliminary stage of information integration from the heterogeneous database. Basically, schema matching process is receiving a pair of a database as an input, matching a pair of an attribute, and verifying by the users to determine whether the result of pairing attribute mapping is already valid or need revision. The problem of schema matching is the difficulty on the mapping of pairing attribute manually, therefore, it is needed the semi-automatic and more effective model. The solution is by improving the computer-based model and combining the provided methods.

The hybrid model of schema matching combined the constraint-based method and instance-based method is developed in this research. The gained model is modified to increase the effectiveness of the result and the efficiency of the process. The model is arranged on the modules including, 1) input reading pair of the database as a DBSource and a DBTarget, 2) database, constraint, and instance extraction, 3) matching, calculation the similarity the pair of attribute (SIM), and the decision of the matched pair of the attribute, 4) user verification, and 5) model effectiveness test. The effectiveness of the model is measured using the parameter of precision (P), recall (R), dan f-measure (F). Model testing is conducted by applying relational database.

Based on our research it is found that a hybrid schema matching model combined the constraint-based method and instance-based method is more effective compared with the constraint-based method or instance-based method individually. The developed model can be applied to the process of schema matching on a relational database with DBMS and or with different application domain. The developed model can also be applied both in combination form and individually on their own method of the component.

Keywords: *constraint-based method, instance-based method, schema matching.*