

ABSTRACT

In current Business Scenario, industries and organizations are generating a huge volume of data in the process of online transaction. Data Mining and Business Intelligence (BI) are employed to analyze and extract the knowledge or insight from this data. These processes are implemented using component or services. Service oriented implementation leverages the advantage of both service computing and BI technologies. Existing systems use multi-layered approach for Service Oriented Business Intelligence (SOBI). These systems do not adapt to dynamically changing user requirement as they use static intermediate data store for the business intelligence process. In this research, a Three Tier SOBI Framework is proposed, implemented and tested for service discovery and composition of BI applications.

The three tier SOBI framework contains user or application tier, service tier and data tier, which creates a XML based dynamic Integrated Data Store (IDS) satisfying the user requirement. The BI services access the contents of IDS for analysis and knowledge extraction. The services in service tier are accessed using programmable interfaces of application tier. The data tier stores IDS and XML replica of the data sources along with original data

sources. Data federation is performed over the proposed architecture for experimentation. The experiments show the integrated view is created 12.1% faster than the existing data federation methodology that uses native database with data controllers. Since data federation is performed over the XML replica, no data are hidden for analysis or decision support which overcomes the drawback of using static reporting data store.

The service providers advertise business intelligence services and stores in a service repository using web service description language description. The Service Discovery (SD) process aids to extract the required service from the repository. The complexity of scanning of service in existing SD processes is linear. The proposed method reduces to logarithmic complexity by using Document Object Model and Virtual Token Descriptor Model for parsing service description over extended Three Tier SOBI framework. The proposed BI service discovery process bifurcate the service attributes as functional attributes and non functional or QoS attributes. The functional attributes are stored in a tree data structure similar to document object model and the non functional attributes are stored in an indexed sequential array (modified Virtual Token Descriptor). Two service discovery algorithms are formulated for efficient and swift service discovery. This SD methodology discovers the QoS aware service 31% (average) faster than the conventional in-memory parse tree based method.

The organization realizes complex BI processes using the composition of simple service. The existing dynamic composition methods consume more space for storing the composite service description. This research proposes an efficient Quality of Service aware service composition methodology using Frequent Service Sequence Patterns. This method converts the composition description to composition tree, containing activity and service nodes. The composition tree is subsequently converted into service sequences using symbolic notations. These service sequences are stored in a repository. The frequent service sequence patterns are extracted using Frequent Pattern Growth algorithm. They are stored in repository along with user request. For any new similar dynamic composition request, the frequently used services are allocated to respective service nodes thereby reducing the service discovery and allocation time. The storage requirement of the symbolic notation representation of composition tree consumes 80% of that in XML representation.

Three tier SOBI architecture is designed and implemented to satisfy dynamic changing user requirement. The three tier SOBI architecture reduces service discovery delay. QoS aware service composition using frequent service sequence is designed and implemented for swift service allocation and effective storage utilization.