



Chapter 7

Internet Delivery of Distributed Data Mining Services: Architectures, Issues and Prospects

Shonali Krishnaswamy
Monash University, Australia

Arkady Zaslavsky
Monash University, Australia

Seng Wai Loke
RMIT University, Australia

ABSTRACT

The recent trend of Application Service Providers (ASP) is indicative of electronic commerce diversifying and expanding to include e-services. The ASP paradigm is leading to the emergence of several Web-based data mining service providers. This chapter focuses on the architectural and technological issues in the construction of systems that deliver data mining services through the Internet. The chapter presents ongoing research and the operations of commercial data mining service providers. We evaluate different distributed data mining (DDM) architectural models in the context of their suitability to support Web-based delivery of data mining services. We present emerging technologies and standards in the e-services domain and discuss their impact on a “virtual marketplace of data mining e-services.”

INTRODUCTION

Application Services are a type of e-service/Web service characterised by the renting of software (Tiwana & Ramesh, 2001). Application Service Providers (ASPs) operate by hosting software packages/applications for clients to access through the Internet (or in

certain cases through dedicated communication channels) via a Web interface. Payments are made for the usage of the software rather than the software itself. The ASP paradigm is leading to the emergence of several Internet-based service providers in the business intelligence applications domain such as data mining, data warehousing, OLAP and CRM. This can be attributed to the following reasons:

- The economic viability of paying for the usage of high-end software packages rather than having to incur the costs of buying, setting-up, training and maintenance.
- Increased demand for business intelligence as a key factor in strategic decision-making and providing a competitive edge.

Apart from the general factors such as economic viability and emphasis on business intelligence in organisations, data mining in particular has several characteristics, which allow it to fit intuitively into the ASP model. The features that lend themselves suitable for hosting data mining services are as follows:

- *Diverse Requirements.* Business intelligence needs within organisations can be diverse and vary from customer profiling and fraud detection to market-basket analysis. Such diversity requires data mining systems that can support a wide variety of algorithms and techniques. Data mining systems have evolved from stand-alone systems characterised by single algorithms with little support for the knowledge discovery process to integrated systems incorporating several mining algorithms, multiple users, various data formats and distributed data sources. This growth and evolution notwithstanding, the current state of the art in data mining systems makes it unlikely for any one system to be able to support all the business intelligence needs of an organisation. Application Service Providers can alleviate this problem by hosting a variety of data mining systems that can meet the diverse needs of users.
- *Need for immediate benefits.* The benefits gained by implementing data mining infrastructure within an organisation tend to be in the long term. One of the reasons for this is the significant learning curve associated with the usage of data mining software. Organisations requiring immediate benefits can use ASPs, which have all the infrastructure and expertise in place.
- *Specialised Tasks.* Organisations may sometimes require a specialised, once-off data mining task to be performed (e.g. mining data that is in a special format or is of a complex type). In such a scenario, an ASP that hosts a data mining system that can perform the required task can provide a simple, cost-efficient solution.

While the above factors make data mining a suitable application for the ASP model, there are certain other features that have to be taken into account and addressed in the context of Web-based data mining services, such as: very large datasets and the data intensive nature of the process, the need to perform computationally intensive processing, the need for confidentiality and security of both the data and the results. Thus, while we focus on data mining Web services in this paper, many of the issues discussed are relevant to other applications that have similar characteristics.

The potential benefits and the intuitive soundness of the concept of hosting data mining services is leading to the emergence of a host of commercial data mining application service providers. The current *modus operandi* for data mining ASPs is the “managed applications” model (Tiwana and Ramesh, 2001). The operational semantics and the interactions with clients are shown in figure 1.

Typically a client organisation has a single service provider who meets all the data

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/internet-delivery-distributed-data-mining/5194

Related Content

IPR Protection for Digital Media Distribution: Trends and Solutions in the E-Business Domain

Bill Vassiliadis and Vassilis Fotopoulos (2007). *International Journal of E-Business Research* (pp. 79-97).

www.irma-international.org/article/ipr-protection-digital-media-distribution/1894

Managing E-Business Change

Colin G. Ash and Janice M. Burn (2009). *Electronic Business: Concepts, Methodologies, Tools, and Applications* (pp. 33-43).

www.irma-international.org/chapter/managing-business-change/9273

Delivering Superior Customer Perceived Value in the Context of Network Effects

Fan-Chen Tseng, Ching-I Teng and David M. Chiang (2007). *International Journal of E-Business Research* (pp. 41-50).

www.irma-international.org/article/delivering-superior-customer-perceived-value/1874

Crowdfunding Platforms Dynamics: When Local Connections Matter

Amaya Erro-Garcés and Begoña Urien (2021). *Multidisciplinary Approaches to Crowdfunding Platforms* (pp. 109-133).

www.irma-international.org/chapter/crowdfunding-platforms-dynamics/272700

Enterprise Service Bus for Building Integrated Enterprises

M. Antonia Martínez-Carreras, Francisco J. García-Jiménez and Antonio F. Gómez-Skarmeta (2012). *Handbook of Research on E-Business Standards and Protocols: Documents, Data and Advanced Web Technologies* (pp. 618-636).

www.irma-international.org/chapter/enterprise-service-bus-building-integrated/63490