

Research on Unified User Data Model Based on Multidimensional Electronic Channels of Internet Marketing Services

Xingxiong Zhu^{1, a}, Ying Tu², Baoxian Guo³, Renjie Li³

¹*State Grid Electronic Commerce Co., Ltd. State Grid Power Finance and E-Commerce Laboratory, Beijing, China*

²*Power Science Research Institute, State Grid Zhejiang Electric Co., Ltd. Hangzhou, China*

³*State Grid Electronic Commerce Co., Ltd, Beijing, China*

^a*zhuxx@pku.org.cn*

Keywords: user data; heterogeneous data; data model; unified data model; big data; multidimensional electronic channels

Abstract: This paper studies the characteristics of heterogeneous user data in different business channels of multidimensional electronic channels, data migration technologies, data fusion technologies, and visualization display technologies, establishes unified data models, unified configuration models, unified fusion models, and visual display models for all channels, and provides data foundation for multidimensional electronic channels unified identity authentication and service supervision system.

1. Introduction

At present, large organizations and groups have established rich electronic service channels. With the rich application of various electronic channels, users are gradually using it, and at the same time users have to experience problems such as the need to register separately and the inquiry service is not shared. Due to multiple data sources and data formats, the lack of unified management of various electronic service channels and lack of supervision over the services of various electronic channels, there is an urgent need to start from the source and to uniformly authenticate users. Specific research is conducted on topics such as electronic channel service supervision.

The user data types of various electronic channels are different and cannot be managed in a unified manner. The timeliness of data sharing between channels is declining, and the customer experience is also declining. Therefore, how to integrate and open up the major channels, study a unified model of standardized user data, establish a centralized and unified account management model to enhance customer satisfaction, is put forward.

Using e-service channels as a platform to study user data for different services under a variety of different e-service channels. Through the reasonable collection of user data of different services under different electronic service channels, the establishment of user data models for different

electronic service channels is conducted against different business scenarios, so as to achieve the establishment of a standardized user data model for major electronic service channels. It provides initial technical support for the next unified account, unified configuration, unified management and control, unified monitoring, data analysis, and strategic research ^[1].

Each e-channel user data type is different, and a standardized unified user data model needs to be established to complete user data fusion and interaction among e-service channels.

2. Methodology

Heterogeneous data, as its name implies, is data of different structures, that is, a collection of related multiple data. Heterogeneous database system is a collection of related multiple database systems, which can realize data sharing and transparent access. Each database system already exists before it joins the heterogeneous database system and has its own database management system ^[2]. While implementing data sharing, each database system still maintains its own application features, integrity control and security control.

Heterogeneous data is reflected in different levels. The computer architecture is heterogeneous, the physical storage of data comes from computers in different architectures, the operating system is heterogeneous, the storage of data comes from different operating systems, and the data format is heterogeneous ^[3]. Different data storage and management mechanisms can be either relational database systems or file data. Data is stored in discrete physical locations. The data is stored and maintained in different business logics, so that the data in the same sense is heterogeneous in performance. Heterogeneous data is often not one level of heterogeneity, but is heterogeneous on multiple levels.

Unified identity authentication and service supervision system construction, brings users "one-time registration, full channel application" convenient experience.

Research data migration technology, set up a unified configuration model to ensure uninterrupted service and to achieve unified identity authentication and service supervision ^[4]. Establish the unified configuration model, transform the historical data to the new production system, and complete the extraction, cleaning, loading of the mass data, and using big data analytics ^[5].

3. Multidimensional Electronic Channel Data Fusion

The business systems and data of the marketing service multidimensional electronic channel have the following features. Mixed data, including structured and unstructured; Large amounts of data, multiple channels of data, data volume are very large; The data quality is uneven, the main data is not uniform, and the data consistency problem is existed ^[6].

The first thing of any enterprise application needs to solve is how to integrate and unify heterogeneous data sources in the enterprise to form a standard, unified, and reliable data source as the basis of the application system. In order to achieve the sharing of heterogeneous data sources, the integration and conversion of heterogeneous data sources must be resolved first.

3.1 Data conversion and migration

The process of data conversion and migration can be divided into three steps: extraction, conversion, and loading. Data discrepancies analysis is a prerequisite for establishing a mapping relationship. The conversion step generally includes the process of data cleaning. The data cleaning is mainly directed at the source database, and performs corresponding cleaning operations on data that are ambiguous, duplicated, incomplete, or in violation of business or logic rules. Before the cleaning, data quality analysis is needed to find out the data that has problems. Data loading is to

load the extracted and converted result data into the target database through a loading tool or a self-written structured query language program.

3.2 User portrait modeling system

To build a user portrait platform, the following points need to be considered to combine the user's portrait platform with the actual business.

First of all, a user portrait platform should be set up to connect a data platform with a large amount of user data and a visual data tool platform. According to different user interaction scenarios, the value of the data platform is tapped. Research and development of production, user researcher, marketing, etc. are able to analyze user characteristics of different products at any time based on needs, to quickly gain insight into user needs. The core questions that the platform needs to answer are who the user is, what the user needs, and where the user is.

Then, a complete user portrait platform needs to consider a comprehensive model system. The data needed to build a user's portrait platform is divided into three types of entities: user, product, and channel.

Users, data dimensions include natural features, interest features, social features, and consumer characteristics. From the point of view of data characteristics, it can be divided into basic attributes and derivative tags. Basic attributes include objective fact data such as age, gender, region, and income.

The product, data dimension includes product positioning and product attributes. Product attributes are actual data such as features, colors, energy consumption, and prices of commodities. Product positioning, that is, the style and positioning of a product, needs to be matched with user tags.

Channels, users obtain information on the information channel and purchase goods on the purchase channel. Different types of users have different preferences for channels, and precise channels can improve efficiency and benefits.

4. The Unified User Data Model

The heterogeneous data of multidimensional electronic channels are collected into the unified model of data by using Extract-Transform-Load tools, the data is cleaned, weighed, fused and reconstructed, and forms a unified user basic information table and data, which forms the data base of the user's unified data model.

Consider a new approach to consolidation, based on service-logical data consolidation, rather than physical consolidation based on data sets. This way of integration is that it's not to insist on physical concentration, but to maintain the distribution of enterprise data. The data of each system through the interface packaged into services, registered to the enterprise service bus, through the enterprise service bus to provide a unified data services, so as to achieve the logical integration of data. The source data can be structured data or unstructured.

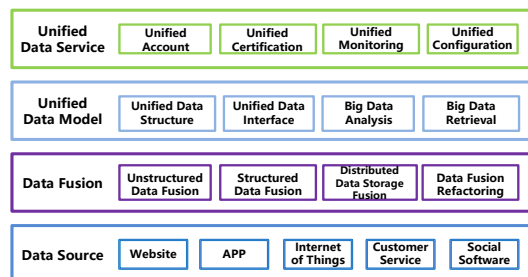


Fig.1. Unified System Model

There are four parts in the unified model, includes data source, data fusion, unified data model and unified data service, shown in Fig.1.

4.1 Process for structured data

The structured database needs open database interface, is used for metadata management system to extract data structure information from source database, and to protect the existence of meta system. The service generation module can query the metadata of each business system stored in the metadata system. The block of code that extracts data is automatically generated by simple actions such as checking and combining fields. It is packaged into a web service, stored in the services run module, and the service is registered with the Enterprise Service Bus (ESB), the service registration can be manual registration, if the ESB can support auto enrollment through the API is better, external data services.

4.2 Process for unstructured data

For NoSQL databases, because there is no unified data structure, it is not possible to automatically generate code blocks and publish them as services. However, web service can be generated by customizing service interfaces, integrated through ESB, and distributed to data integration platforms to provide external services. In this case, custom development of web services can only be performed for each interface.

4.3 User login authentication

Provide users with unified registration, registered users can use the platform products based on their permissions.

4.4 Operations management

Manage all users of the unified user identity authentication and service supervision system, including the user's audit, deletion, etc. Assign corresponding usage rights for the users identified by the unified user identity authentication and service supervision system.

4.5 Platform access management

To access the unified user identity authentication and service supervision system for the management of products.

4.6 Access product section

Increase the user authentication interface, and hook up with the unified user authentication and service supervision system authentication interface to implement user authentication.

5. The Technical Framework of Unified User Data Model

The whole technical architecture of unified identity authentication platform is divided into infrastructure layer, basic service layer, application layer, presentation layer.

5.1 Infrastructure layer

Infrastructure for carrying unified user data model systems, includes networks, servers, storage, etc. It meets the needs of multidimensional electronic channels.

5.2 Basic service layer

Basic service includes data fusion and unified data model service.

5.3 Application layer

The application layer is responsible for the processing of business logic such as unified user account, unified user certification, and unified monitoring.

5.4 Presentation layer

It presents the unified visual user model, and supports mobile and web sides.

6. The Core Technology of Unified User Data Model System

The core technologies include the followings such as multi-dimensional electronic channel data analysis and modeling; Customer-oriented multi-dimensional electronic channel service; Marketing-oriented multidimensional heterogeneous electronic channel service supervision; Data migration, data integration, high concurrency, high performance, distributed storage, security assurance support technologies, unified identity authentication and service monitoring system^[7].

6.1 Highly-compatible mobile application data automatic acquisition technology

The collected data has complicated indexes, diverse collection channels, and different technologies and framework protocols. How to solve the above differences and study the high-compatibility and high-performance acquisition technologies is key to accurate collection of intelligent service supervision at the data collection level and to diversification analysis.

6.2 Business monitoring index and application analysis technology for multidimensional electronic channel user and application service

The collected data comes from multiple electronic channels and requires detailed quantitative analysis of these data. Uses data mining techniques to analyze the internal relationships between these data and user experience, application service health operation, and fault warning. Studies the corresponding monitoring indicators and analysis techniques. Through the research of the content, technical support is provided for the improvement of user experience and stable operation of application services.

6.3 The new mode of unified identity authentication with multiple electronic service channels

To open and share accounts between multiple channels, the unified account authentication system mainly starts from the three aspects of unified access, unified account management, and unified rights management. Realizes the mutual recognition of accounts among systems, the data structure among systems, and the unification of communication methods. Achieves uniform external services, ensuring the unification of customer experience.

6.4 The unified account management

Unified account management mainly manages the account information of all channels, the relationship between channel accounts and users. Manages the menus of all channels, the function items of all channels, and the interface services provided to all channels.

6.5 The application integration methods

For information exchange and data sharing between systems involves a small amount of quasi-real-time data transmission, message transmission, can be achieved through application integration.

The interface is integrated, and the business system completes the integration of functions such as registering, logging in, merging users, modifying user information, binding user information, and authorizing third-party logins through interface integration.

Interface and unified user Application software development kit (App SDK) integration, the authorization login is integrated using the App SDK. The unified user platform provided App SDK ensures that the access styles of all access users are unified.

Web integration and interface integration, unified user platform provides a complete set of user authentication and single sign-on integration specification. The business system is in accordance with the specification to achieve the authentication interface and user information query interface. This type of integrated business system enables full-service single sign-on integration of different domain names between business systems.

7. Conclusion

This paper studies the characteristics of heterogeneous user data, data migration technology, data fusion technology and visualization display technology across channels. Establishes unified user data model, unified configuration model, unified fusion model and visual display model of each channel service. Provides unified identity authentication, big data analysis and supervision service.

Acknowledgment

This work was supported by the Science and Technology Project of State Grid Corporation of China under the Grants No. 52110417001D, Research on key technologies of unified identity authentication and service supervision based on "Internet +" marketing service electronic channels.

References

- [1] Zaharia M, Xin R S, Wendell P, et al. *Apache spark: a unified engine for big data processing [J]. Communications of the ACM*, 2016, 59(11): 56-65.
- [2] Zhou Q. *Research on heterogeneous data integration model of group enterprise based on cluster computing [J]. Cluster Computing*, 2016, 19(3): 1275-1282.
- [3] Baumann P, Mazzetti P, Ungar J, et al. *Big data analytics for earth sciences: the earthserver approach [J]. International Journal of Digital Earth*, 2016, 9(1): 3-29.
- [4] Xingxiong Zhu, Qingsu He, Shanqi Guo. *Application of block chain technology in supply chain finance [J]. China's circulation economy*, 2018, 32 (03): 111-119.
- [5] Jess, T., Woodal, P., & McFarlane, D. (2014). *A framework for identifying suitable cases for using market-based approaches in industrial data acquisition. International Data and Information Management Conference (IDIMC) 2014*, 113-124.
- [6] Drovandi, C., Holmes, C., McGree, J., Mengersen, K., Richardson, S. T., & Ryan, E. *Principles of Experimental Design for Big Data Analysis. Statistical Science*, 2017,32 (3), 385-404.
- [7] Han Xiao, Di Fangchun, Liu Guangyi, Zhang Yi, Chen Jinxiang, Liu Fengcheng, Sun Xin. *Application framework and practice of big data application for unified data Model of smart grid [J]. Power grid Technology*, 2016,40 (10):

3206-3212.

Author Introduction

Xingxiong Zhu (1975), male, graduated from Peking University, master degree in software engineering, one of the main drafters of China National Standard GB/T 25656-2010 "Information Technology Chinese Linux Application Programming Interface (API) specification". Email, zhuxx@pku.org.cn