Research on Data Mining of Personal Income Tax in Tax Collection and Administration Audit

Yiran Tao1,

1Chongqing College Of Architecture And Technology,  401331

Abstract. In the context of the rapid development of economy and information technology, the level of tax informatization is getting higher and higher. The tax department has obtained a large amount of tax management data. Taxation departments must analyze and apply these data in order to grasp the deficiencies in the current tax management work, so as to take effective measures to improve tax quality and efficiency. This is the focus of the work of the taxation department. In the process of personal income tax collection and management, full application of data mining technology can break through the traditional personal tax collection and management model and improve the level of personal tax collection and management. At the same time, this can also ensure the effectiveness of the tax data verification work, which is positively helpful to improve the tax data application capabilities. This can also provide an effective reference for tax management decision-making. At this stage, the full application of risk-oriented audit methods based on data mining technology can improve the effectiveness of personal income tax collection and management audits. However, it is necessary to pay attention to a comprehensive study of the sources of personal income tax data. At the same time, the taxation department should also discuss the data processing of personal income tax and the review of the rationality, legality, and compliance of personal tax collection, so as to improve the data mining level of China's personal tax collection and management.

1 Introduction

Taxation is the main component of China's fiscal revenue, and the level of tax collection and management will have a certain impact on the security of China's treasury funds. In the tax collection and management work, audit operations need to be strengthened. Tax collection and management should use the relevant content of the annual audit plan to ensure the effectiveness and reliability of the audit work. In the process of tax collection and management, personal income tax is an important tax in the local tax system. County-level administrative regions face a relatively large number of collection and management objects when carrying out individual tax collection and management work. In the traditional personal income tax audit process, the audit methods used are mainly based on account book basic audit and system basic audit, which can no longer meet the current needs of personal income tax collection and management audit. Therefore, tax administration departments need to fully mine personal income tax data and conduct personal income tax audits in accordance with relevant Chinese regulations. This can ensure the quality of individual tax collection management audit.

2 Personal Income Tax Data Acquisition and Processing

2.1 Source of Personal Income Tax

At this stage, the relevant data of personal income tax is mainly based on the relevant reports in the tax collection and management system to complete the information recording work. Subsequently, the staff of the tax department needs to directly import and import the information of all the tables into the SQL database. Then, the tax department staff will separate from the personal income tax data summary table to facilitate data mining. For example, when mining personal income tax data, the main methods adopted are: select * from collection table where tax type = 'individual income tax'.

2.2 Personal Income Tax Treatment Method

The taxation department must analyze and process the personal tax data information when obtaining personal tax data information in order to grasp the application value of the personal tax information and improve the efficiency of information utilization. First, the taxation department must judge the authenticity and accuracy of
the individual tax information, and ensuring that the individual tax information is complete and reliable is the basic link in the application of the individual tax information. After the taxation department separates the personal tax information, it must first judge the authenticity and completeness of the personal tax information in accordance with the relevant operating habits. The individual tax collection amount in the general paper tax collection report is the sum of the amount of the individual tax in the warehouse. When judging the authenticity and completeness of the data, the tax department needs to compare the data consistency of the two tables. The same data indicates that the tax data information is complete. Second, the tax department needs to analyze and compare the data between the tax data tables. In the process of taxation data management, the taxation department must focus on basic tables for analysis. The cross-checking relationship of the basic table already existed at the beginning of tax collection and management. When comparing the differences between the tables, the tax department should judge whether there is an audit risk based on the comparison result. If there is a discrepancy between the tables, it indicates that there is an audit risk, and the taxation department must conduct an audit and verification work. If the tax amount is the same as the entry amount in the individual tax entry form, it means that all individual taxes are entered into the inventory. If there are differences, the tax department needs to analyze the specific differences. For example, the tax collection amount is larger than the storage amount, and the tax department needs to use the following method to find the tax receipt number of the tax that has been collected but not stored in the storage: select * from a tax collection table where whether it has been stored in the storage 'has been stored in the storage'.

In the process of searching and auditing, the taxation department must focus on designing individual tax receipts that have not been put into the warehouse. Especially in the case of audit doubts, the taxation department shall assist in the application of other audit methods to ensure the effectiveness and reliability of the verification work. The taxation department needs to conduct an in-depth analysis and verification of the individual tax entry form Jining when comparing and analyzing the entry data information of the remaining individual tax collection and levy forms. If there is still a difference, data mining must be used to find the specific reason for the difference. Taxation departments can group and compare individual tax receipt numbers, tax types, tax items, and tax rates, and search and analyze the details of the tax amount in the warehouse. The main method is: select a.*, b.*, a. Collection amount-b. Deposit amount from (select tax ticket number, sum (actual tax amount) collection amount from a tax collection table where the tax period starts like group by tax receipt number) a full join (select tax receipt number, sum (incoming tax), the amount of storage, the name of the collector from the storage table where the tax period starts like group by tax receipt number, the name of the collector ) B on a. Tax receipt number = b. Tax receipt number where abs (a. Collection amount-b. Storage amount)>0.

In the process of tax collection and management audit, the taxation department can use the above-mentioned data mining method to find out the specific situation of the inconsistency between the tax receipt amount and the amount in the warehouse in a timely manner. When the taxation department audited the tax collection and management work of a certain county, there were discrepancies between the levied amount of dozens of tax receipts and the amount in the treasury. However, the tax receipt status is warehousing, and the tax receipt is collected by the same collector, and the amount that has not been warehousing is tens of thousands of yuan. The taxation department can basically judge that the tax collector is suspected of embezzling taxes. The tax department can use this as a clue to actively implement the situation, and transfer the collection officer to the relevant unit Jining for processing. Taxation departments can obtain taxpayer information tables, collection tables, warehousing forms, declaration forms and other data through comparative analysis of individual tax collection and storage tables. In addition, the tax department can also use a variety of comparison methods for data mining to improve the effectiveness of individual tax audit work [1].

3 Verification of Personal Income Tax Data Information

When examining the rationality, legitimacy and compliance of individual tax collection, the taxation department shall make judgments based on individual tax collection and management tax rates and tax items. The tax items of individual tax mainly include income from wages and salaries, income from production and operation of individual industrial and commercial households, contracted operations of enterprises and institutions, and income from leased operations. When the income tax on wages and salaries is collected, tax audits are carried out at seven levels of excess progressive tax rates and other tax items [2]. In the actual verification work, the taxation department can adopt the method of first totaling and then dividing according to the specific situation of individual tax collection, and use data mining strategies to obtain important audit doubts. The taxation department must first complete the individual tax item and tax rate review of the individual tax collection table, and then generate the individual tax rate table. The main method is: Select tax type, tax item, tax rate from collection table where tax type = 'personal income tax' group by tax type, tax item, tax rate order by tax type, tax item, tax rate.

The tax department can use the above sentence to obtain the tax items and tax rate types of all individual taxes. Taxation departments can use this as a basis to screen some tax items and tax rates that do not meet the individual tax collection regulations in a simple way. For example, the taxation department used this method to find out more than ten types of tax items and tax rates that did not exist in a certain tax collection and
management audit process. Taxation departments use these non-compliant tax items and tax rates as keywords, and can dig out specific tax situations that do not comply with the regulations. In a certain tax collection and management audit process, there are as many as hundreds of individual tax receipts that do not meet the requirements. In-depth analysis of these data by the tax department can be used as audit risk points and audit doubtful points for the next step of verification work. Generally, there are many reasons for this problem. Tax officials do not have a comprehensive grasp of tax rates and tax items. In the collection and management of small and many tax items, they have not carried out work according to relevant procedures, and there is a black box operation in tax collection and management. These behaviors will cause tax items and tax rate types that do not meet the requirements [3].

When the tax department analyzes the personal income tax, it can understand that there are differences in the tax declaration methods, and the data information will also be different. There is a certain difference between individual declaration and unit withholding. In data mining, it is necessary to compare and analyze different personal income tax payment methods. In the case of withholding by the unit, the taxpayer is mainly the unit. However, when an individual declares tax, the subject of the taxpayer is the individual's name. When processing relevant information, the taxation department can effectively distinguish between different types of personal income tax collection methods based on the unit name is 3 characters, and the personal name is usually 2 to 3 characters, and then deal with this issue. First, the tax department faces many situations when analyzing individual tax declarations. The main reason is that individuals lack unified management and procedures when reporting taxes, which leads to confusion in the tax information declared by individuals. For example, differences in tax items, differences in the treatment of some special circumstances in tax items, calculation of different excess progressive factors, etc., will increase the complexity and quantity of tax information declared by individuals. If the taxation department cannot improve the tax management method and deal with these situations in traditional and conventional thinking, it will increase the difficulty of the tax examination and approval work. Even if a single tax receipt is used, there will be many different tax items and tax rates. This situation is quite special. When the tax department analyzes this situation, the writing of sentences may be affected by the complexity of the data. This leads to insufficient integrity and logic of data mining, and affects the results of data analysis. In this case, the tax department can deal with it in the form of tax invoices. The tax department needs to record specific situations separately, complete the search operation, and then conduct a single situation recording and review work, which can improve the efficiency of individual tax review. Although this working method will increase the overall workload, it can ensure the reliability and credibility of the audit results. For example, the tax department needs to verify the adjustment of the level difference between the labor compensation when screening personal labor compensation. Only in this way can you obtain tax receipts that have not been processed in accordance with tax regulations. In China's tax management standards, the individual tax collection and management of labor compensation is less than 20,000 yuan, and the tax rate is 20%; 20,000 to 50,000 yuan, and the tax rate is 30%. In the quick calculation process, 2,000 needs to be deducted; for labor services over 50,000 yuan, part of the tax rate is calculated at 40%, and 7,000 needs to be deducted in the quick calculation. According to this tax collection management method, it is possible to obtain records of all items that are not levied on labor remuneration tax in accordance with regulations. In addition, the tax department can also conduct joint management of the obtained record entries, and conduct centralized processing with other records that have not been conducted in accordance with the requirements of collection and management. Regarding the processing results of these record items as audit risk points, the next audit verification operation can be carried out. Second, when the taxation department analyzes the collection of individual tax withheld on behalf of the unit, the processing process is roughly the same as that of individual tax declaration. However, it should be noted that when the unit withholds tax on its behalf, it is necessary to withhold the personal tax of the internal employees of the unit. In the single tax receipt processing method, it is relatively difficult to strip the data, which affects the effect of data mining and leads to more content involved in data mining. This will greatly increase the difficulty of tax collection and management audit work. In order to strengthen the effectiveness of individual tax collection and management audit work and prevent the increase in difficulty, auditors can coordinate and communicate with the tax system before the tax audit. The auditors can record and store the personal tax information collected by the unit after completing the tax collection and management work. This is conducive to the professional construction of the tax system itself, and at the same time can improve the efficiency of tax auditing. For example, in the auditing unit's unified withholding of wages and salaries and individual tax collection and management, auditors can conduct unified management on employees with the same individual tax rate. Auditors only need to count and record the number of employees, which can effectively reduce the difficulty of audit work and improve audit efficiency [4].

4 Application of Personal Income Tax Data Mining Technology

4.1 Overview of Data Mining Technology

When analyzing data mining technology, the tax department mainly defines it technically and commercially. Among them, data mining is defined technically. It refers to the process of extracting information and knowledge that people don't know
beforehand with certain use value from a large amount of fuzzy and random practical application data. The commercial interpretation of data mining can define the data mining process as a data processing technique for obtaining valuable business information. In the process of data mining, the taxation department must come up with effective and practical data information from a large number of commercial databases. The tax department must ensure that the data information obtained is the key data that can improve the level of business decision-making. Analyze it from the data mining technology itself. Data mining mainly refers to the comprehensive application of scientific methods such as digital statistics, artificial intelligence, and neural networks. At the same time, data mining technology can combine these scientific methods with genetic algorithms, cluster analysis, and decision trees. In this way, it is possible to dig out hidden data information with a certain reference value for decision-making in a large number of complex databases, and to establish relevant models of auxiliary decision support based on the acquired data information. In the process of data mining, the data stored in the database needs to rely on the relevant algorithm standards of data mining. Only in this way can we obtain valuable data information from the database. After that, the tax department must preprocess the data. Taxation departments need to use specific data usage requirements as the basis, comprehensively apply algorithm regulations, select the most suitable data mining algorithm, and finally complete the data extraction process [5].

**Fig. 1. Schematic Diagram of Data Mining Process**

### 4.2 System Design Points

In the process of individual tax collection management audit, taxation departments can use data mining technology to design individual income tax management and supervision systems. Using this system to carry out personal tax supervision can improve the service level of personal income tax declaration, management, evaluation, and inquiry. In the system design process, the following functions should be realized. Firstly, the taxation department should establish a one-family file of personal income tax declaration in the system to summarize the specific information and data of withholding agents and taxpayers to lay the foundation for subsequent management work. Secondly, the tax department should use one-family files to screen high-income individuals, high-income industries, key withholding agents, and key taxpayers. The tax department mainly conducts tax management work based on the filtered information. Thirdly, the taxation department should set different types of statistical indicators, and the use of accurate statistical results to carry out individual tax management can improve the effectiveness of individual tax management. In the meantime, the taxation department can provide taxpayers with data support services such as inquiries and statistics in accordance with China's taxation policies. Make full use of the data mining system to provide good tax management services. Fourth, the tax department can accurately determine whether the tax information declaration is accurate and true after analyzing and processing the tax information of the taxpayer withholding agent. Meanwhile, the taxation department can carry out effective taxation supervision work to ensure the reliability of tax collection and inspection work. Fifth, the tax department should build a data interface with strong compatibility to ensure that data can be shared with different departments and provide accurate and complete information for other tax collection and management. Sixth, when the tax department processes multiple personal income and tax information, the application of this system can upload personal information to the collection and management system, so as to further review the personal income tax.

Seventh, the taxation department can provide tax payment certificate printing services to provide taxpayers with tax payment basis[6].

### 4.3 Data Mining Process

When applying the personal income tax analysis and evaluation system, taxation departments need to build a complete and reliable personal taxation system based on the personal income tax collection and management system. Only in this way can it provide effective data resources for the tax assessment system. At the same time, the tax department can build intelligent software modules by using data mining and information data. This module mainly takes tax management as the core to carry out tax evaluation and management. In this way, the efficiency and level of tax management can be improved. In the process of applying this module, the taxation department can supervise and manage the taxation behavior declared by taxpayers, which has positive significance in promoting the modernization and informatization of taxation work. In the taxation data mining process of the taxation department, its specific mining process includes the following aspects. Firstly, the taxation department must analyze and compare the taxpayer's tax-related information and relevant index information, and use this as a basis to judge whether the taxpayer's tax declaration is complete and accurate. Secondly, the General Administration and Provincial Administrations can assess the quality of tax collection and management of different levels of tax authorities. The local tax bureaus of different regions, cities, and counties need to assess the taxation status of taxpayers and withholding agents. Simultaneously, the General Administration and Provincial Bureaus also need to evaluate their tax collection agencies. During the operation of the system, it mainly includes basic taxpayer data, decision trees, construction algorithms,
and object evaluation processing. During the operation of the system, the tax department can achieve the goals of data mining, predictive analysis and evaluation. In this way, effective verification of individual tax collection and management can be carried out, the efficiency of the verification work can be improved, and the authenticity and reliability of the verification results can be guaranteed.

5 Conclusion

In the current personal tax collection and management audit process, the full application of computer database technology can greatly improve the level of personal tax collection and management data mining. With the widespread application of computer audit methods in the tax collection and management audit process, more and more innovative audit methods are constantly being developed. This also has higher requirements for the auditor's work ability. Auditors must continuously enrich their professional skills based on the specific development of individual tax collection and management data mining. Meanwhile, the auditors must master advanced computer operation technology, so as to see more, think more, check more, and ask more. Only in this way can we quickly adapt to the innovative development of data mining technology for tax collection and management. At present, there are more and more data mining methods. In order to promote the effective application of data mining in tax collection and management, tax authorities can apply it to data mining of other types of taxes. This can improve the level of China's tax collection and management audit work.

References

2. Wang Liying. Research on national audit risk in local tax collection and management audit [D].
4. Li Jing. Design and implementation of personal income tax management system based on data mining technology [D]. Xidian University.