An Implementation of Information Technology in Massive Questionnaire Survey for the Climate Index of SMEs

Xin Fu[1], Linping Xu[2(✉)], James Ho[3], and Zhiwei Zhang[3]

[1] Business School, Nanjing University Jinling College, Nanjing 210089, China
xin.fu@jlxy.nju.edu.cn
[2] School of Economics, Nanjing University, Nanjing 210093, China
xlp@nju.edu.cn
[3] School of Information Science and Engineering, Nanjing University Jinling College, Nanjing 210089, China
{james.ho,zhangzhiwei}@nju.edu.cn

Abstract. A combination of WeChat applet and web-based support system is designed and utilized in order to facilitate a massive questionnaire survey for Climate Index of small and medium-sized enterprises. This survey system could assist both in investigation stage and analysis stage. The WeChat mini program enables investigators to retrieve interviewed enterprise information and upload questionnaire answers automatically in a single online applet. The web-based support system provides reinforcement in extreme data exclusion, index calculation, statistical classification as well as result visualization. The whole system has been utilized in Climate Index of SMEs survey 2020 and verified its efficiency and reliability.

Keywords: Climate Index · Questionnaires · WeChat applet · Data processing

1 Introduction

A series of statistical yearbooks indicated that small and medium-sized enterprises, abbreviated as SMEs, played a significant role both in economic development and employment stability. Take SMEs in Jiangsu as an example. In 2016, the contribution rate of this category of enterprises to industrial growth reached 84.1%, and it absorbed more than 80% of new labors in the job market [1]. Meanwhile, SMEs in China have been confronting several difficulties including upsurging costs in raw materials and productivity, restricted fund-raising channels [8], and relatively inferior innovative capability [4].

Thus, the *Enterprises Ecological Research Center*, EERC, was set up by Business School of Nanjing University Jinling College in April 2014, in response to a demand for continuous observing and evaluating the operation status and performance of small and medium-sized enterprises. The institution introduced an integrated approach to assess

the circumstances for SMEs, which is called Climate Index of SMEs (hereinafter stand for Index) and have conducted series of yearly surveys in Jiangsu province since 2014.

During every summer vacation, teachers and students from Nanjing University are distributed to 13 prefecture-level cities in Jiangsu to investigate more than 3000 targeted companies through questionnaire surveys and face-to-face interviews with senior managers. Based on the accumulated effective statistics, the EERC is able to process and calculate the Index, reflecting various aspects of SMEs in terms of operating condition, cost structure and macroeconomic environment. The total index, alongside with several sub-indexes, are released and published in a cluster of research reports every September. The analysis drawn from the annual report attracts extensive attention from authorities as well as enterprises in decision-making [3, 5].

Following several years of the practice of the Climate Index survey, the flaws of manual data collection and process, especially in such massive survey with thousands of questionnaires, have been emerging obviously. The data collection which is conducted mainly by paper questionnaire could take weeks to be distributed and assembled, with inevitable errors or blank answers, even inconsistent information in one single questionnaire. Meanwhile, manual data upload and post process are also time-consuming. Results from enterprise respondents' answers were typed into Excel documents, where the data is calculated by a series of formulas. Any analysis or graphic presentations in terms of charts or tables could only be processed based on the complicated Excel filtering functions and not be classified automatically, which is quite inefficient.

In order to optimize the accuracy and efficiency, the article introduced an implementation of information technology which could digitalize and streamline both the survey phase and data analysis phase. In the data collection phase, an online questionnaire based on WeChat applet technology could facilitate investigators' online or offline interviews and eliminate incomplete or inconsistent answers by automatic discrimination mechanism. In the data processing and analyzing phase, a web-based system is proposed to provide whole procedure support including enterprise information retrieving, extreme data preprocessing, index calculation as well as further research with result visualization.

2 Systematic Requisitions

2.1 Enterprise Information Retrieve

In the past year surveys, investigators entered the basic information of interviewed enterprises manually, which included company scale in terms of revenues, labor force and assets. The respondents provided these sensitive details instinctively and indistinctly, while investigators could not verified the information in time. The WeChat applet should retrieve and validate the interviewed enterprises' basic details by comparing it with the database from a cooperated authorized enterprise information provider.

2.2 Questionnaire Filling and Inspection

Since thousands of questionnaires were distributed and collected each year, manually entering and checking these paper documents is becoming more and more energy-consuming. Investigators were required to type the answers twice into the Excel document in order to check the inconsistent information in any grid. Table 1 illustrated the

Table 1. Scale of Climate Index Survey in 2014–2020.

	Number of Students involved	Number of Teachers involved	Number of Valid Questionnaires
2014	249	20	3508
2015	500	30	5488
2016	420	26	3221
2017	460	30	3904
2018	401	34	4112
2019	392	23	3983
2020	328	18	2029
Total	2750	181	26245

massive scale of the questionnaire survey of Climate Index of SMEs. With the support from the web-based survey system, surveyors should directly choose the options in the questionnaire one by one through the WeChat applet in their smartphones when they interview the enterprises' executives. Before being uploaded to the system database, all the answers should pass the automatic discrimination, namely, comply with several basic rules designed by the survey organizer to avoid incorrect or incomplete data.

2.3   The Extreme Data Preprocessing

The entered questionnaire data are transmitted from interviewers' WeChat applet to a networking database. Prior to any index calculation, the system should be capable to valid the effectiveness of each questionnaire by excluding the unreliable or vacant answers. Some respondents' judgement might be over optimistic or pessimistic, others might be contradictory even in different answers within one single questionnaire. It is considerably crucial to reject these sorts of questionnaire which would damage the robustness of further analysis.

2.4   Index Calculation

One of the vital attributes of the survey system is to calculate the Climate Index from collected data. Instead of traditional manual processing, the system should enhance the timeliness of calculation and declaration both for the total index and the sub-indexes of production, finance, marketing, policy situation and legal environment. An interval period for calculation should be taken into consideration as it is unpractical to update the proclaimed indexes over-frequently.

2.5   Multi-dimension Analysis, Statistic Categorizing and Result Visualization

The system is required to deliver a user-friendly interactive interface to economic or management researchers but quarantine the detailed database mechanism from the qualified

applicants. Data analysts who intend to study the SMEs in various dimensions should be empowered to easily categorize and calculate the sub-indexes in terms of areas, industries, enterprise scales and time. An automatic function of graphic presentation in bar charts, radar charts or matrix should also be proposed.

3   Overall Design and Structure of the Survey System

In order to satisfy the above requisitions, the survey system is designed as a combination of a WeChat applet and a web-based support system shown in Fig. 1. The WeChat applet enables users to implement data collecting and inspecting operations such as retrieving enterprise information, uploading and viewing questionnaires online timely and conveniently. The web-based support system facilitates users to preprocess data, calculate index as well as post analyze statistics.

3.1   WeChat Applet Design

The WeChat applet is introduced to promote work efficiency and data accuracy, which are both vital in data collection phase. It applies the C/S architecture to comply with WeChat official requirement while core codes are written in JavaScript with wxml and wxss for page rendering.

In detail, the applet system adopts the Model-View-Control Layer Frame, i.e. MVC. The View Layer operates on the WeChat and present user interface; the Control Layer operates on the server side and mainly controls the invocation of related logic; the Model Layer encapsulates the business logic of the applet and access to the database.

The database server is chosen to be settled on Alibaba Cloud, using MySQL with knex.js and koa2 as the backend development framework. A set of tables have been established in the database after the questionnaire data was analyzed and refined. In addition, the Entity-Relation diagram in Fig. 2 illustrates the primary key and foreign key
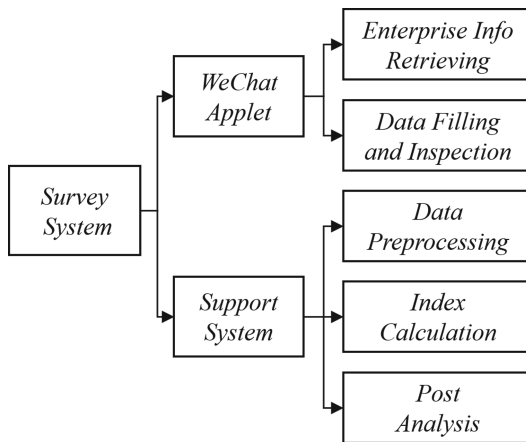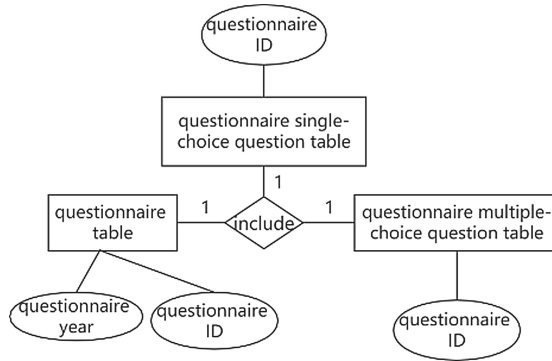

Fig. 1. Survey system structure.

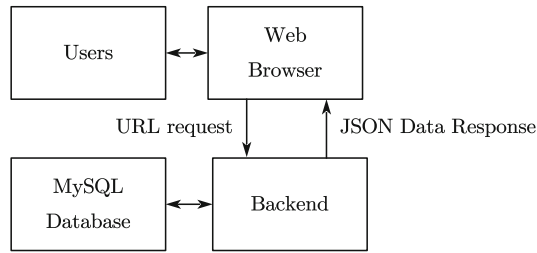**Fig. 2.** ER diagram of database tables in WeChat applet.

Fig. 3. Infrastructure of frontend and backend.

of each table alongside with the relationship between them. Communications between the applet and the server bank on the JSON data.

3.2 Support System Design

The support system adopts a B/S architecture to isolate the front and back ends. In this manner, researchers could streamline the data application through the web browser acting as the communication interface while the backend stored and packaged all the services logic. The JSON data format is deployed for information exchange between the frontend and backend, presented in Fig. 3.

MySQL is preferred as the most appropriate system data foundation and established on the commercial Alibaba Cloud for internet access. All the three related implementing modules defined in Fig. 1 are supported by this database. Relations between a series of tables in the database are demonstrated in a concise ER diagram as Fig. 4. Enterprise information, together with questionnaire answers are the two essential tables which formulate the fundamental statistics of Climate Index and post analysis. All the data gathered from onsite or online interviews with enterprise executives would be transmitted simultaneously to the questionnaire answer table, triggering subsequent preprocessing and index calculation. Additionally, efficient Climate Index calculation also demands for categorized optimistic and pessimistic matrix table.
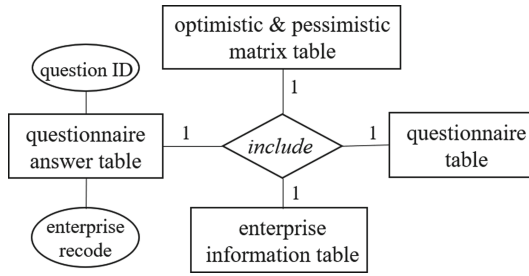
Fig. 4. Brief ER diagram of the database.

4 Functionalities Implementation

4.1 Enterprise Information Retrieve

Accurate and authentic enterprise information constructs the base of the Climate Index survey. EERC chose to collaborate with an authorized enterprise information provider who could verifies the up-to-date details of the interviewed enterprises. Investigators could search the enterprise identity name through the WeChat applet, which activates a third-party API to the data provider for acquiring and retrieving related business intelligence. The return data would be spontaneously loaded into the particular questionnaire when the enterprise identity corresponds to the provider's database. If the API failed to return the expected enterprise information, investigators would also be cued to modify any input message or manually fill in business basic data for the specific enterprise. Figure 5 demonstrates the precise retrieving procedure.

4.2 Questionnaire Answers and Inspection

Apparently, questionnaire answers are the essence of the Climate Index survey. Based on the onsite conversation or online video-chat with enterprise executives who are quite familiar with their own business, investigators are able to choose the proper options in the questionnaire one after another until all the blank questions are responded [7].

When the investigators have successfully retrieved and signed in the enterprise identity, all the questions in the questionnaire would be directly loaded and presented in the WeChat applet in their smartphones, as exhibited in Fig. 6. Taking the advantage of thoughtfully designed single or multiple-option questions in several business-related aspects, surveyors could unveil the particular enterprise's operation status in terms of production, marketing, finance, external policy situation and legal environment through the responses to these questions.

Several elementary rules designed by the survey organizer, EERC, must be obeyed when the options of answers are inspected in order to prevent any incorrect or incomplete data from undermining the reliability of calculated index. For example, blank responses would not be tolerated in those questions with non-omitting label, or minimum choices should be fulfilled in the multiple-option questions with pre-requirements. And those questionnaires with almost the same responses, whatever optimistic or pessimistic, would
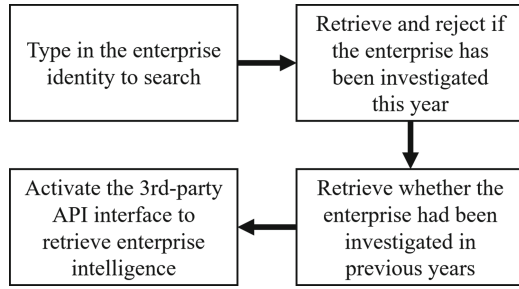
Fig. 5. Enterprise identity and information retrieving procedure.
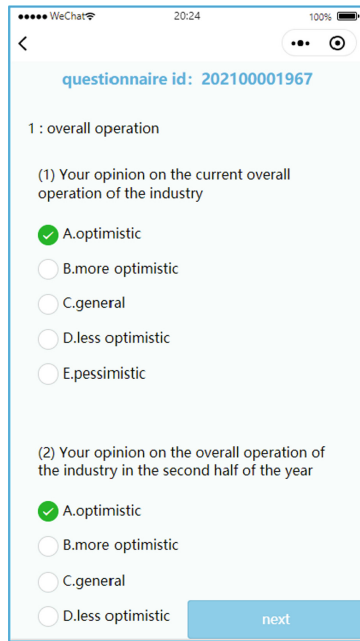

Fig. 6. WeChat applet questionnaire example.

be rejected as invalid samples either. Once the answers passed the inspection, they would be "locked" for reviewing only and no more adjustment is permitted.

4.3 Questionnaire Data Preprocessing

According to the interviewees' subjective assessment to the economic environment and business operation condition, the responses to the questions in the questionnaire are assigned value from 1 to 5, respectively from pessimistic to optimistic. Any vacant answers are assigned value 3 so that it would not trigger any statistical deviation in the calculated index.

Aiming to differentiate the over-estimated replies, an accumulated score, which is the simple aggregation of all the answer values, is marked for each particular questionnaire. An exclusion mechanism is introduced in this stage, so that a specific proportion of the lowest or/and highest scores, representing most negative or positive estimation, could be eliminated in the consequent calculation. The survey organizer and other researchers could schedule the exclude proportion in advance or easily modify it afterwards when any macroscopic elements might be considered.

4.4 Index Calculation and Declaration

The survey support system is mainly originated for efficient index calculation and prompt result declaration. The real-time feedback of the survey is visible and accessible to the investigators and researchers for the period that the survey is under way.

Index calculation absolutely banks on the data from Climate Index survey. Technically, all the responses in the questionnaire are composed of two replies: spot and expected estimation to the business environment. The enterprise executives need to reply to the two sub-questions from value 1, the most pessimistic perspective, to 5, the most optimistic perspective, based on their subjective estimation to recent and near future business condition. These replies to the questions will be regarded as the basic influential components to the Climate Index of SMEs.

In the first place, the answer value 1 to 5 to each specific question will be shifted to the calculation foundation from most pessimistic pe_1, pe_2 to most optimistic op_1, op_2, where footmark 1 and 2 are corelating to spot and expected estimation. Subsequently, all the valid questionnaires are deconstructed into cities where jurisdiction could be imposed over urban and rural areas. Thus, the Climate Index components matrix could be assembled in this manner.

For one particular city, say ith of the 13 cities (take Jiangsu Province for example), there are M pieces of questionnaires are referring to, and a particular question, say jth of the N questions in each questionnaire, the *spot* and *expected values* are interpreted below, correspondingly,

$$[spot\ value]_{i,j} = \frac{\sum_{k=1}^{M} op_{1,i,j,k} - \sum_{k=1}^{M} pe_{1,i,j,k}}{\sum_{k=1}^{M} op_{1,i,j,k} + \sum_{k=1}^{M} pe_{1,i,j,k}} \tag{1}$$

$$[expect\ value]_{i,j} = \frac{\sum_{k=1}^{M} op_{2,i,j,k} - \sum_{k=1}^{M} pe_{2,i,j,k}}{\sum_{k=1}^{M} op_{2,i,j,k} + \sum_{k=1}^{M} pe_{2,i,j,k}} \tag{2}$$

Eventually, all the calculated figures assemble two $[13 \times N]$ matrixes, named spot and expected climate index matrix, which components are referring to every single city and scheduled question. The Total Climate Index, TCI, could be calculated as follow,

$$TCI = w_1 \sum_{i,j}^{13,N} [spot\ value]_{i,j} + w_2 \sum_{i,j}^{13,N} [expect\ value]_{i,j} \tag{3}$$

where w_1 and w_2 are coefficients of the spot and expected climate index assigned by researchers Xu, Ho and Zhang [6].

Owing to the answered questionnaires continuously transmitted into database while the survey is still under way, the index calculation should be renewed and revised regularly. Strictly speaking, the revise task is pointless for over-frequency or absorbing no new questionnaire data. The activating point could be arranged to a least figure as new Q questionnaires are uploaded into database, and additionally the point is user adaptable. In the meantime, the declaration interval period could be set at half an hour or above. The instant Climate Index can be exhibited on both WeChat applet and web browser pages.

4.5   Post Analysis and Result Visualization

Statistical categorizing and deep data analysis will be initiated as soon as the index calculation has been completed. The web-based support system chose vue.js as the user interface (UI) solution to facilitate surveyors' and researchers' manipulating experience. The UI always presents results for user's application immediately and is isolated from backend sophisticated processing, due to the progressive vue.js framework.

The questions in the survey are categorized into two levels of sub-indexes. Level 3 sub-index concentrates more on detailed business operation status such as raw material supply, capital-raising channels, innovation capacity and taxation burden. Meanwhile, 5 dimensions are categorized in Level 2 from higher perspectives, which were indicated in previous chapter as production, marketing, finance, external policy situation and legal environment. Both levels of sub-indexes are enabled to be calculated out of the spot and expected climate matrix.

Aiming to visualize the indexes and compare the results clearly, the support system also acquired ECharts from Baidu, an open-sourced Java Script plug, into vue.js

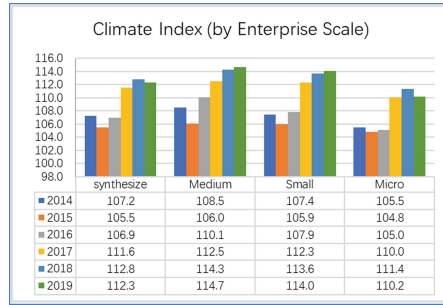**Fig. 7.**   Interface of sub-indexes analysis.

Fig. 8. Bar chart: Climate Index by enterprise scale.

infrastructure. The system empowers the registered researcher to interact on the web and generate any bar charts or radar maps automatically to satisfy the study purpose, as shown in Fig. 7. An example of the Total Climate Index in bar chart based on different enterprise scales is exhibited in Fig. 8.

5   Verification and Utilization

In the 2020 summer vacation, 318 students and 18 teachers from Nanjing University, were organized by EERC to participate in yearly survey for Climate Index of SMEs. They were trained to utilize the WeChat applet as available investigation devices, interviewed enterprise executives about their evaluation referring to the business development situation, recorded the responses in the questionnaire, collected the data and transmitted back to the system. Additionally, typing the statistics into Excel documents in weeks was replacing by automatic inspecting and uploading mechanism timely, which reduced survey workload.

Rather than filtering manually by researchers in September and October in previous years, the WeChat applet and support system quickly discriminated all the collected data in the summer, and 2029 valid questionnaires were refined from originals. The samples are in very optimized quality since the Cronbach's alpha reached 0.937, which is the highest confidence since the series surveys were initiated.

Following the completion of the onsite and online survey, researchers deployed the web-base support system to calculate the Total Climate Index and a series of sub-indexes for deep analysis from different dimensions. The 2020 Jiangsu SMEs Climate Index and Ecological Environment Evaluation Report were released in December, involving dozens of monographic projects or research which were conducted on account of the system support.

A radar map in Fig. 9 with the comparison between 2019 and 2020 indicated that 3 aspects, which are production, marketing and finance, have dramatically shrunk due to the COVID-19 outbreak in the early 2020. The pandemic has truly affected companies' internal operation severely, however, the faith recovery shown in other 2 sub-indexes of policy and legal environment reflected that government and authorities had launched mass bailout policies and resumption regulations which were fully perceived by SMEs and consolidated economic resilience [2].
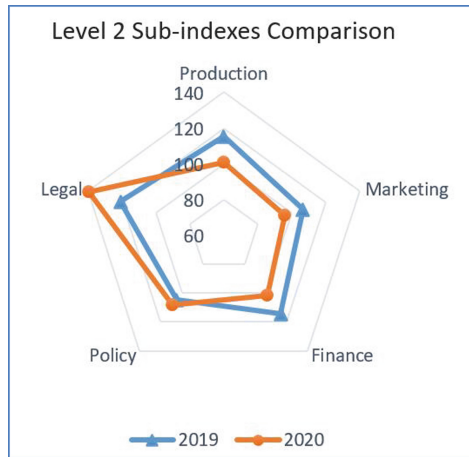
Fig. 9. Radar map: comparison by sub-indexes in 2019–2020.

Owing to the survey system combined with WeChat applet and support system, not only the filling or calculating period but also the raw information inspection have been reduced from weeks to hours. Researchers' efforts would be shifted more to depth analysis by interacting conveniently with the web interface and generating visualized results automatically.

6   Conclusions

In this paper, a dedicated implementation of information technology, which combined both WeChat applet and web-based survey support system, is proposed to facilitate the Climate Index of SMEs in both investigation stage and subsequent processing stage. Compared to the traditional methods in previous years, the system which was utilized in 2020 obviously demonstrated its efficiency, accuracy and reliability. There was an essential enhancement not only in the timeliness of questionnaire collection and inspection, but also in the assistance of data processing and further analysis.

References

1. Enterprises Ecological Research Center (EERC). (2017). Enterprises Ecological Evaluation Report of Jiangsu SMEs 2016, Nanjing University Press, Nanjing
2. Enterprises Ecological Research Center (EERC). (2021). Enterprises Ecological Evaluation Report of Jiangsu SMEs 2020, Nanjing University Press, Nanjing
3. Gu S, Xu L, Chen X (2018) Small and micro enterprise ecological environment and enterprise innovation performance–an empirical study based on survey data in Jiangsu Province. J Mod Econ Res 4:52–60
4. Lun X, Liu Y, Shen K (2017) The role of the government and the development of small, medium and micro enterprises-based on the survey data of 4574 enterprises in 13 prefecture-level cities in Jiangsu Province. J Econ Theory Econ Manag 4:82–96

5. Wu F (2017) Research on the government financing support index of SMEs in Nanjing. J Contemp Econ 31:72–75
6. Xu L, Ho J, Zhang Z (2021) CIS4: climate index of SMEs survey support system. In: 2021 International Conference of Social Computing and Digital Economy, pp 79–82
7. Xu L, Guan W, Ho J, Zhang Z (2021) Design and implementation of the questionnaire WeChat applet for the climate index of SMEs. In: 6th International Conference on Smart Grid and Electrical Automation, vol 2, pp 481–484
8. Zhang X, Xu L (2020) Rising house price and financing costs for SMEs: evidence from a survey data in Jiangsu Province. J China Soft Sci 1005:35–45