

Software Tools Developed and Employed for Survey on Central Taiwan Science Park and Data Analyses.

F.-Y. Leu, Y.-T. Huang

▶ **To cite this version:**

F.-Y. Leu, Y.-T. Huang. Software Tools Developed and Employed for Survey on Central Taiwan Science Park and Data Analyses.. In International Conference of Territorial Intelligence, Besançon 2008., Oct 2008, Besançon, France. pp.10. halshs-00516475

HAL Id: halshs-00516475
https://shs.hal.science/halshs-00516475

Submitted on 29 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Software Tools Developed and Employed for Survey on Central Taiwan Science Park and Data Analyses

Fang-Yie Leu and Yao-Tien Huang

leufy@thu.edu.tw

Professionnal adress

Dept. of Computer Science and Information Engineering, Tunghai University, Taiwan

Summary: In this paper, we would like to study environmental change and living impacts on resident lives before and after Central Taiwan Science Park (CTSP) started its construction in 2003. To collect the information of concern, we interviewed people who live around the CTSP, including Situn District and Daya Township, by dispatching students who brought questionnaires designed in this study to visit the residents home by home. The surveyed data are recorded in a database and analyzed by using two software tools, iData Analyzer and ESIR ArcGIS which are data mining tool and GIS packages, respectively. The former is used to find out the connotative information embedded in raw data so users can catch the deep knowledge from what we have collected, whereas the latter is employed to spatially present the raw data and analyzed data on an electronic map from which users can realize how the data are spatially distributed and what their spatial semantics are.

After mining the data, we found that people living in Daya Township have no significant changes on their everyday lives. But Situn District has been quickly developing. Most employees of the CTSP and their families come to live in Situn since this district is located on the side near Taichung city which is a huge modern city with plenty of people and living facilities, e.g., restaurants, supermarkets, hotels and others, which have been prosperously constructed in the past few years in Situn District. Even this, from some special viewpoints, we can see that Situn's development will soon approach to a bottleneck due to its limited hinterland. The consequence is many people will move to Daya which will then migrate to a busy and business town from a current quiet agricultural town.

During the interview, we did not record the interviewed people's GIS positions. Thereafter, we spent a lot of time to locate their addresses one by one on the electronic map. In fact, if we can develop a software tool to automatically record the position, including longitude and latitude, for each interviewed home/resident, it would be easier for us to show up the addresses by sending the position information to a GIS system. Therefore, in this research, we develop a positioning system by employing a GPS. When a visitor finishes his/her interview to a person and steps out the house, what he/she needs to do is pressing a few key buttons on his/her notebook computer. This will record the geographical position of the house to the corresponding record in the database. The tool can also transform the latitude and longitude to TM2 which is the format that our GIS system provides. Besides, for the team members' convenience, the tool saves the data/file in TXT format so the data can be

conveniently invoked by Microsoft Office, ESRI ArcGIS and other software tools without any transformation. To further the utilization of the positioning and GIS tools, operation guides for them have been written.

Keywords: GPS, GIS, Central Taiwan Science Park, environmental change, impacts on resident people.

1. INTRODUTION

The main objective of this study is to provide software tools for territorial intelligence [1, 8-9, 11, 16-21] analysis.

Territorial knowledge, territorial methods and tools, and territorial governance are three main research topics concerning territorial intelligence analysis. Among them, territorial knowledge is the one that emphasizes spatial analysis. In this study, we combined different disciplines to develop a new analytical tool, and made use of several existing tools. However, to successfully study territorial intelligence, in addition to basic tools, communication and coordination among government, organizations and citizens are very important in territorial development and management.

An academic cross-discipline research center, named Methodologies and Technologies of Information Applied to the Human and Social Sciences (Centre MTI@SHS), led by Dr. Jean-Jacques Girardot, University of Franche-Comte, comprises members from different departments. One of the main goals of the center is analyzing territorial intelligence by using information technology and geographical information system. They have successfully applied computer technology to assist to social scientists to do society research, and created several systems to help governments and organizations to plan and evaluate social policies. Of course, the system can also be used to control and prevent regional disasters. The center has developed a system, called Catalyse which integrating fundamental notions, information technologies and applications, consists of four subsystems, PRAGMA, ANACONDA, NUAGE and SITRA. Several European nations and non-profit organizations have used the Catalyse to control and balance their regional development. The center also requests all its researchers learning essential information technology, e.g., geographical information systems and data analysis techniques, …etc, so they can continue developing and improving the Catalyse. Dr. Jean-Jacques also welcomes anyone and any non-profit organizations to use the freeware tool to develop their territorial management systems.

Territorial methods and tools are used to diagnose and analyze territorial intelligence. Its objectives focus on serving citizen requests, promoting society development, and solving territorial disputes by using specific approaches. It emphasizes how to attract users to participate in territorial activities and what territorial actors should behave. They share data with regional organizations. The target is to encourage territorial administrators and actors to analyze data with Catalyse tools. Now, for each local government and nation, territorial development has been a basic standard. Government officers often follow the standard to manage their territory.

Previous territorial governance experience has proved that territorial knowledge has a strong influence on democracy progress and citizen society activities. In Centre MTI@SHS, territorial knowledge is continuously accumulated and developed. The researchers of this center have proposed a new management method which is developed under citizen participation. This method can effectively improve the communication between government and citizens, and guarantee balanced development between the democracy and territory. To achieve the goal, the research method tries to popularize computer technology, which can further show up how territorial intelligence influences citizen societies, a nation's territorial policies, and society culture, and how it affects the development and evolution of specialized aspects particularly in social science.

Centre MTI@SHS analyzes territorial intelligence variables and individuals, through which they can

understand behaviors of a society and its members. After that, researchers can find out the phenomenon and process about territory, so they can apply social welfare to target societies, and let government officers who are authorized to make policy decision [12, 14] realize how to assist specific groups that need help. This can also promote territorial administrators to find out factors that cause specific problems and troubles. Thereafter, they can make right decision to solve the problems and troubles then.

An important way to do territorial intelligence research is collecting more information and observing society behavior on territory. However, quantity of territorial historical data is huger day by day. This implies that researchers will face a very large database from which they can retrieve much more data and many more samples. Now, how to analyze the huge amount of data is quite a problem. The simplest way is using software tools, i.e., information technologies.

We hope that information technologies can not only be popularly used by people who take part in citizen activities, but also be promoted to territorial globalization and balance, so all conflicts of interests and cooperation can be soon solved.

In this study, several software tools have been employed and developed as a basis of an established analytic platform for territorial intelligence. The platform provides territorial intelligence actors and other users with powerful analytical tools to process and analyze data. For example, researchers can input data to database, and invoke a data mining tool [2-3, 5-7, 10, 15] to find out deep knowledge. A GIS system [13, 22-26] is also provided to present relations between/among results generated by executing different policies made by a government. GIS can also assist users to monitor and prevent disasters. We introduce them as follows.

2. A DATA MINING TOOL AND ANALYSIS RESULTS

We used a data mining tool, iData Analyzer, to analyze question 34 to question 38 in the questionnaires designed for survey on people who live in Situn District and Daya Township.

Question 34: *About 4 to 5 years ago (before CTSP started its construction), where had you gone shopping?*
Question 35: *In the passed one year, where have you gone shopping?*
Question 36: *About 4 to 5 years ago (before CTSP started its construction), where had you eaten out?*
Question 36a: *In the passed one year, where have you eaten out?*
Question 37: *About 4 to 5 years ago (before CTSP started its construction), where had you gone to spend your leisure time?*
Question 37a: *In the passed one year, where have you gone to spend your leisure time?*
Question 38: *About 4 to 5 years ago (before CTSP started its construction), where was your office or business location?*
Question 38a: *In the passed one year, where was your office or business location?*

After using the data mining tool, we found some interesting phenomena. The analysis results show that after

Central Taiwan Science Park started its construction in 2003, 86.67% of Situn District people have changed their jobs, and 92.86% of people still work in Great Taichung region, including Taichung county and Taichung city.

In Daya Township, the analyses are as follows.

(1) 75.68% of interviewed residents had no job. We analyzed the reason, and found that visitors visited these residents in business/working hours. People who earn money for families have gone to work. So most interviewed residents are housewives or farmers.

Questions 36 and 36a focus on people's eating-out area, whereas questions 37 and 37a concern people's leisure area.

(2) The analyses results of questions 36 and 36a are:

(2.1) Eating-out areas for no-job interviewed residents are 100% in Daya Township.

(2.2) 64.66% of working residents whose working areas are in Daya Township.

(3) Analysis results concerning leisure area for Daya Township people are:

(3.1) 75.68% of no-job interviewed residents have not changed their leisure areas after CTSP started its construction.

(3.2) 88.72% of working interviewed resident did not change their leisure areas after CTSP started its construction.

(4) There is another interesting analysis result, which is 75.68% of no-job interviewed residents have not changed their eating-out area and leisure area after CTSP started its construction.

The result represents that most Daya Township residents did not change their eating-out and leisure areas since 2003. We bring up three issues to explain these phenomena and reasons: 1) Buildings for office and living (e.g., apartments) have not been sufficiently constructed around CTSP; 2) Educational level of Situn District residents is higher than that of Daya Township people; 3) CTSP was started its construction in 2003, but we visited CTSP in July 2006. Living change takes time. So the Daya Township residents have not significantly changed their living habits yet.

In Situn District, the places around CTSP have been well constructed to support CTSP's running, and a lot of buildings are now under construction or has been finished its construction. Today, CTSP is almost completely constructed. But both Situn District and Daya Township residents have not shown big changes on their living habits. We believe and predict that the residents, due the trend of science park development, will change surrounding environments and habits several years later.

3. APPLICATION OF ARCGIS

We keyed in surveyed data to database whose database management system is MySQL, and classified data into two categories, according to where interviewed residents live, i.e., Situn District and Daya Township. After that, we used ArcGIS [4] to spatially present the surveyed data collected for question 34 to question 38 of the questionnaires, and plot the regions that residents eat out, shop, and spend leisure time. Finally, we showed the regions on two maps, one for Situn District and the other for Daya Township.

We calculated statistic data, including maximum, minimum, and average values of question 34 for Daya Township, and as shown in Figure 1 displayed the data on an electronic map in which light green means that no resident was interviewed in this village.

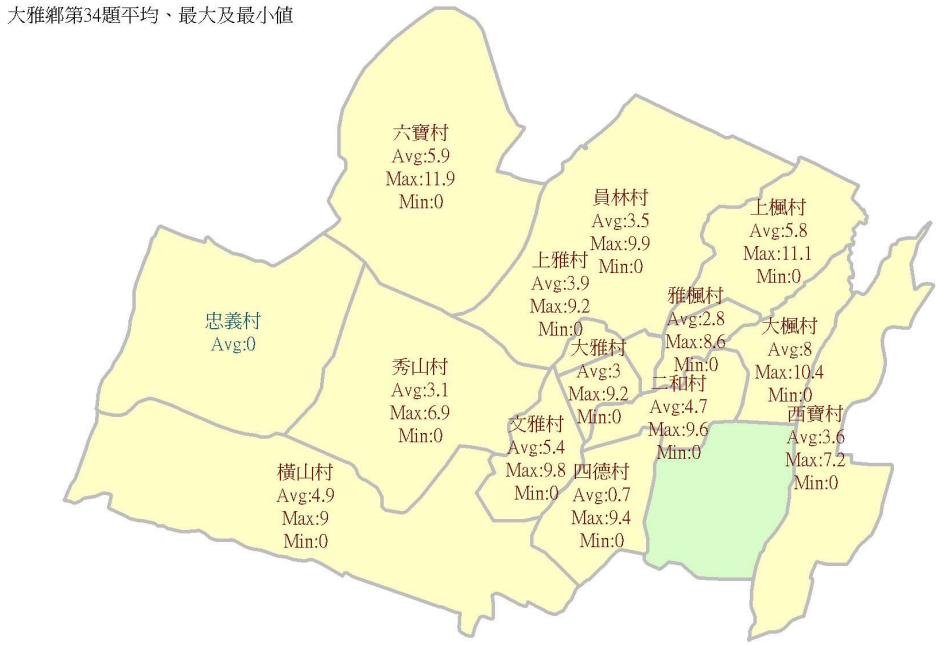


Figure 1: *The surveyed data concerning question 34 for each Daya village includes its maximum, minimum, and average values which are shown by an GIS system.*

Figure 2 shows the similar results of the same question concerning Situn District. The light blue, light green and pink bars, respectively, represent average, maximum, and minimum values. The shadowed villages represent their minimum value is 0, i.e., no resident is interviewed.

On the other hand, when visitors visited residents, they did not gather and record the residents' GPS information. To show up the information on an electronic map, we must manually estimate each interviewed resident's location by using inner interpolation method, and then key in the data one by one, resulting in poor data collection efficiency. To solve this problem, we developed a positioning tool that can not only gather GPS information, but also read GPS files and import them into ArcGIS all automatically with very few human interference.

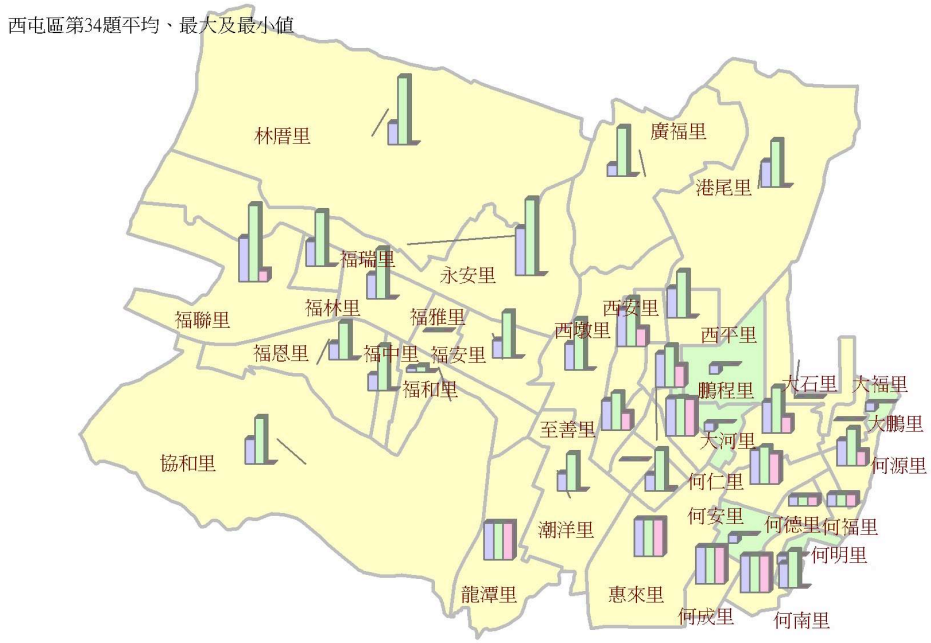


Figure 2: M*aximum, minimum, and average values of question 34 for each town in Situn District (max: blue; min: pink; avg: green).*

4. A POSITIONING TOOL



Figure 3: *The positioning interface of our GPS tool.*

Figure 3 shows the tool's interface. After a user clicks down "□□GPS□□(reading GPS signal"), it starts receiving GPS signal and decoding the signal. GPS signal consists of interviewed resident's home longitude and latitude which will be sent to ESRI ArcGIS. So, the signal should be transformed to X-Y coordinates which are shown on "□□□□X□□ (transformed X coordinate) " and "□□□□Y□□(transformed Y coordinate) ". We should also key in

resident's ID to "□□□□(record information)". Then, after the user clicks down "□□□□□□(save the record) ," the tool will record the ID with its corresponding X-Y coordinates in a database file.

In order to make the data of concern readable to Personal Digital Assistant (PDA) devices, we save the resident ID and X-Y coordinates in TXT format. Users can then easily use or edit this data by using Microsoft Excel and other software packages.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we used a data mining tool, iData Analyzer, to analyze questionnaires concerning Situn District and Daya Township, and mine interesting features that are hidden in questionnaires. ArcGIS is employed to show analyzed results concerning question 34 to question 38 for Situn District and Daya Township on electronic maps. When visitors visited these residents, they did not record their GPS information immediately since at the beginning of the interview there is no tool can help them to do that. So we developed a positioning tool to record GPS information, which provides TXT format and is able to import GPS files into ArcGIS automatically to show up interviewed residents' positions.

In the future, we will develop and integrate some required software packages, including a data mining tool, a decision support system, SPSS, a data classification tool (e.g., K-means algorithm), and an GIS system. The data mining tool as stated above is iData Analyzer developed by information Acumen Corporation. The classification tool is a decision tree which organizes and orders important attributes that are used to classify data. Entropy will be the factor to determine which level an attribute should be located. K-means algorithm can cluster existing objects into K groups. It is a useful clustering tool. The GIS system as stated above is ArcGIS which provides spatial vision with maps, such as buffering, geographic statistic distinction, etc.

REFERENCES

[1] Mitchell, T.M. (1997), "Does Machine Learning Really Work?" AI Magazine, vol.18, no.3, pp.11-20.

[2] McCarthy, V. (1997), "Strike It Rich," Datamation, vol.43, no.2, pp.44-50.

[3] Baltazar, H. (2000), "NBA Coaches' Latest Weapon: Data Mining," PC Week, pp.69-69.

[4] ESRI - The GIS Software Leader, http://www.esri.com/.

[5] Garcia-Holina, H., ullman, J.D. and Widoma, J. (2000), *Database System Implementation*, Prentice Hall.

[6] Roiger, R.J. and Geatz, M.W. (2003), *Data Mining*: A Tutorial-Based Primer, AddisonWesley.

[7] Adriaans, P. and Zantinge, D. (1996), *Data Mining*, Addison Wesley.

[8] Breiger, R.L. (ed.). (1990), *Social Mobility and Social Structure*, Cambridge, University of Cambridge Press.

[9] Camagni, R. (1991), *Innovation Networks: Spatial Perspectives*. London, BelhavenPress.

[10] Fayyad U., Grinstein, G.G. and Wierse, A. (2002), *Information Visualization in Data Mining and Knowledge Discovery*, Morgan Kaufmann.

[11] Florida, R. (1995), "Toward the Learning Region," *Futures,* 27(5): 527-536.

[12] Klijn, E. (1997), "Policy Networks: an Overview," in Kickert, W.J., E. Klijin, and J. Koppenjan (eds.), *Managing Complex Network: Strategies for the Public Sector*. London, Sage Publications.

[13] O'Looney, J. (2000), *Beyond Maps: GIS and Decision Making in Local Government,* Redlands, CA, ESRI Press.

[14] Rhodes, R.A.W. (1997), *Understanding Governance: Policy Networks, Governance Reflexivity and Accountability*. Buckingham, Open University Press.

[15] Roiger R.J. and M.W. Geatz (2003), *Data Mining-A Tutorial-Based Primer.*

[16] Sassen, S. (1994), *Cities in a World Economy*. Thousand Oaks, CA, Pine Forge Press.

[17] Wellman, B. and Wortley, S. (1990), "Different Strokes From Different Folks: Community Ties and Social Support," *American Journal of Sociology,* 96: 558-88.

[18] Western, M. and Wright, E.O. (1994), "The Permeability of Class Boundaries to Intergenerational Mobility among Men in the United States, Canada, Norway and Sweden," *American Journal of Sociology* 59: 606-629.

[19] Harris, J.M., Timothy A.W., Kevin P.G.., and Neva R.G. (ed.). (2001), *A Survey of Sustainable Development: Social and Economic Dimensions.* Washington, DC, Island Press.

[20] Redclift, M. (1987), *Sustainable Development: Exploring the Contradictions*. London, Routledge.

[21] World Commission on Environment and Development (WCED). (1987), *Our Common Future*. Oxford, Oxford University Press.

[22] Hariharan, R., Shmueli-Scheuer, M., Li, C., and Mehrotra, S. (2005), "Quality-driven approximate methods for integrating GIS data," *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, p.p. 97-104.

[23] Reibel M. (2007), "Geographic Information Systems and Spatial Data Processing in Demography: a Review," *Population Research & Policy Review*, vol. 26, issue 5/6, pp. 601-618.

[24] Constantin D.L., and Radu L.N. (2008), "Location and networking: Neural network-based models and geographic information systems," *ECONOMIC COMPUTATION AND ECONOMIC CYBERNETICS STUDIES AND RESEARCH*, vol. 42, issue 1-2, pp.131-142.

[25] Karadimas N.V., and Loumos V.G. (2008), "GIS-based modeling for the estimation of municipal solid waste generation and collection," *WASTE MANAGEMENT & RESEARCH*, vol. 26, issue 4, pp. 337-346.

[26] Hillier A.E. (2007), "WHY SOCIAL WORK NEEDS MAPPING," *Journal of Social Work Education*, vol. 43, issue 2, pp. 205-221.