Application of Chinese Phonetic Initial Inquiry in Police GIS

Yuanyuan Han, Lianhe Yang

School of Computer Science and Software, Tianjin Polytechnic University, Tianjin, China

Email: hanyuanyuan125@163.com, yanglh@tjpu.edu.cn

Abstract: The application of GIS technology in police work has become one of the most important aspects in the present police information infrastructure construction, specially along with the gradual popular implementation of "golden shield project" and its depth application, the police service information became an important technical method to "the technical strong police" in our country. This paper introduces what Chinese Phonetic Initial (CPI) Inquiry is, how to make it come true as well as how applies it in the police GIS to provide the fast highly effective policy-making support for the police service work.

Keywords: GIS; police service information; Chinese Phonetic Initial Inquiry

1. Introduction

With the rapid development of urban modernization, the traditional management models and approaches have failed to meet the public security system development and modernization of information technology operational needs. But as the development of 3S technology (GIS (geographic information system), GPS (global positioning system), RS (remote sensing)), the characteristics of mass data storage, query, analysis, visualization applications and spatial information are consistent with the public security system and modernization of operational information needs [1].

Police GIS brings the traditional database into the visualization, this makes up the limitations of Public Security Organs in the application to the current the analysis data. Through the comprehensive utilization of GIS spatial analysis function and strong visual expression ability, it makes police data and space information as one. By monitoring the various elements of policing work in space distribution and real-time operation, analyzing its inner link, can make the reasonable configuration and dispatching resources. So it can supply many models to manipulate data and help policemen making decisions quickly.

110 Command System (110CS) brings "concentrated tandem, automatic shunt, category jurisdiction, center for monitoring" into practice in emergency call reception and handling system. It makes emergency call reception and handling system perfected and greatly shortens the time. At the same time, the 110CS ensures the efficient and sensitive command and control system. Greatly improve the efficiency of the Public Security Bureau and the rate of capturing.

Nevertheless, improve emergency call reception and handling to enhance the efficiency of alarm cannot completely meet the needs of the modernization of public security organ and the actual cases. So, in order to further enhance the Public Security Bureau in decision-making, scheduling and events, meet modern combat capability, gradually realize command center by "duty type" to "staff type" or "command type"[2]. Using the relevant application module of police GIS to ascertain the concrete location of alarm spot and find a nearest path from start point to alarm spot in time. In order to enrich the methods of the task of alarm receipt and disposal, improve the efficiency and accuracy, dispatch police resources rationally, and increase the rate of scene capturing. Make Public Security Organ severely punish the crimes according to law, service people, maintain social stability, more quickly in temporal and spatial reaction, more reasonable in police cooperation, higher content of science and technology, the better able to play an active role in building a harmonious socialist society[3] [4]. Police GIS should be continuously toward standardization, such as portability interoperability, scalability and general environment etc. It can greatly enhance the data application, implement information sharing.

Based on this actual application needs, research and development the police GIS with ability to storage and manage mass data and higher spatial visualization application that becomes a particularly urgent need, and has a very business and marketing value.

2. Introduce CPI Inquiry

A. The disadvantages of previous inquiry method

Chinese information retrieval is one of the most important functions of MIS (Management Information System), in information management application it often need to input Chinese characters to query corresponding data[5]. For example an inquiry according to name of somebody, name of place, name of book etc. The operation process commonly is: Switching to an input method, inputting Chinese characters, querying, and then closing the input

978-1-935068-40-2

method. This operation process is more complicated and inefficient. Mainly in the following respects:

1) Need overmuch keystroke. For example: inquiry the police named "xiangyanglou", you will keystroke 12 times at least. Although the existing Chinese character input methods provide phrases, and more and more intelligent. But if you input infrequent name of somebody or someplace, this built-in limited words obviously don't meet the needs of users, it requires users input the words one by one, resulting in low query efficiency.

2) For the input speed depends on the operator's proficiency. By a five-stroke input method, the speed of querying slightly faster, but it needs to memorize the rules of the input method. Therefore for ordinary users, they often choose the most popular input method, such as Sogou or ABC. Due to a lot of homophones words in Chinese characters, needs many times in the next page to find, especially for the infrequent words, more times. All of those can impact the speed of querying relevant information, bring inconvenience for jobs.

B. The concept of CPI inquiry

The "CPI" we mean you input the first letter of Chinese characters in the textbox to research what you want. For example: If you want research some information about "xiangyanglou" in police GIS, you only need to input the CPIs "XYL" in the query textbox, that is, all pronunciation of the same string "XYL" will be listed in table, you will choose the target.

C. The advantages of CPI inquiry

The efficiency based on Chinese character query is low, because it needs overmuch keystroke. Regardless of what kind of keyboard input, take more trouble. Therefore it becomes the key issue of impacting querying speed. If don't input Chinese characters to query Chinese information, there is no doubt, this will greatly improve the inquiry efficiency. Taking this fact into consideration, the author raises the thought of CPI inquiry. Then this method is applied to police GIS, will improve the staff's work efficiency, quickly locate alarm spot and find a nearest path from start point to alarm spot in time.

3. Implementaion of CPI Inquiry

A. Machine code to generate CPIs contrast

Well known, Chinese characters were stored in computer with machine codes. In this way, every Chinese character corresponds to single machine code regardless of input mode. In addition, every machine code was arranged in sequential order [6]. Such as, Chinese character initial by "A" was ordered in the head, by "Z" was ordered in the end, others certainly were ordered between "A" and "Z".

In machine code, Chinese characters' order is: characters for level one were ranked by initials; however, characters for level two were ranked by radicals. Now, in most cases, people use the GB2312-80 Chinese characters for level one finding the head and the last characters machine codes. So, for the researching characters of level one, you only need to local the scope where the machine code is, and you will know which initial corresponds to the Chinese character. For example, the machine code's range of Chinese initial "A" is between HB0A1 and HB0C4, initial "B" is between HB0C5 and HB2C0. According the feature of machine codes, designing a function, if a machine code is equal or greater than HB0A1 and less than or equal to HB0C4, then this Chinese character initial is "A", and so on.

This method is easy, although can generate the CPIs, it has lots of shortage: firstly, the number of Chinese characters is only 3755, frequently-used Chinese characters outweigh this amount greatly, and it cannot satisfy demand; secondly, Chinese characters for level two is ordered by radicals rather than initials, by this measure, we cannot distinguish those characters, although they are unfrequented once they are in use will be more problems. Finally, this measure is unable to process the polyphone. Based on above, this author puts forward a new modified mind to generate CPIs.

B. Method of generate CPIs

We create a new phonetic dictionary *dict1.dbf*, all of inquiry is based on the dictionary, so how to create the *dict1.dbf* is the prerequisite of CPI. First, we get the GB2312-80 Chinese-phonetic letters Contrast [7] *dict.txt*, including 6763 Chinese characters, as shown in Fig. 1.

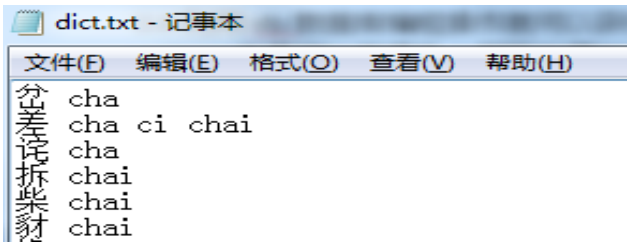

Figure1. Chinese-phonetic letters Contrast

Disposing the file of dict.txt in the Visual FoxPro, as follows these steps:

1) Create a new table *dict1.dbf*. There are five fields, hanzi, duoyinzi1, duoyinzi2, duoyinzi3, duoyinzi4, in this table, denoting Chinese character and, it's the first pronunciation, the second pronunciation, the third pronunciation, and the fourth pronunciation. All of the fields have the same type, character, and their widths are 2.

2) Table *dict1.dbf* is created, the content of their fields are null. So, we will realize from *dict.txt* to *dict1.dbf*. Programming to fill in the fields content automatically [8]., as shown in Fig. 2.

Dict1

Hanzi	Duoyinzi1	Duoyinzi2	Duoyinzi3	Duoyinzi4
岔	cha			
差	cha	ci	chai	
诧	cha			
拆	chai			
柴	chai			
豺	chai			

Figure2. Generated *dict1.dbf*

With having been created table *dict1.dbf*, we can program to generate the CPIs in Visual FoxPro (without including the polyphone), key codes as follows:

```
cizu = allt(thisform.text1.value)
t= len(cizu)/2
use dict
pym=""
aa=""
tt=""
for i=1 to t
   bb=subs(cizu,i*2-1,2)
   loca for hanzi=bb
   do while .not.eof()
      if found()
      aa=aa+left(duoyinzi1,1)
      pym=pym+allt(duoyinzi1)
      cont
      endi
   endd
endfor
tt=upper(aa)
thisform.text2.value=allt(tt)
thisform.text2.refresh
```

The map data of Tianjin includes 27 layers, they are: police substation, schools, banks, government agencies, drugstore, Post and Telecommunications etc. Each layer has 6 files. Now we illustrate with police substation layer, it consists of *paichusuo_point.dbf, paichusuo_point.prj, paichusuo_point.sbn, paichusuo_point.sbx, paichusuo_point.shp, paichusuo_point.shx*. The *paichusuo_point.dbf* is Attibute table. It stores all the names of police substation. Append *yintou* field in this table [9]. Theoretically, according to the method above, it can automatic generate CPI data for the police substation. Fill the CPIs to *yintou* field, as shown in Fig. 3.

Paichusuo_point

Zifuchuan	Yintou
和平公交治安派出所	HPGJZAPCS
和平刑侦支队第八大队	HPXZZDDBDD
马场派出所	MCPCS
和平区防暴巡警支队	HPQFBXJZD
和平区新兴城管执法中队	HPQXXCGZFZD
和平区法院司法警察大队	HPQFYSFJCDD
新兴派出所	XXPCS
天津市公安局刑侦局	TJSGAJXZJ
学府街派出所	XFJPCS
嘉州路交警大队	GZLJJDD
和平区南营门城管执法中队	HPQNYMCGZFZD
南营门派出所	NYMPCS
甘肃路派出所	GSLPCS
和平区刑侦支队第七大队	HPQXZZDDQDD

Figure3. Table of paichusuo_point

Practically, the issue is much more complicated than in theoretical, because there are lots of polyphones in Chinese. We must consider the Chinese polyphones. For example, the CPIs of "yinhang" will be "YX", in face it is "YH", and the CPIs of "shanghang" will be "SX", in fact it is "SH". So, the CPIs need to be amended.

All the pronunciations of polyphones found in GB2312-80 have been listed in [4], and every polyphone's pronunciations are arranged according to its usage frequency. The first pronunciation is high-frequency, the second and latter frequency is lower and lower. For the most part, when a polyphone appears in a place name, it pronounces the first. Furthermore, although some polyphone has different phonetic letters, they have the same CPIs. In this case, it is unnecessary to amend it. In conclusion, the amendment of CPI is only the non-high-frequency pronunciation.

Just as the name suggests, "polyphone" means that a Chinese character has different pronunciations in different context. In fact, through using 2-character-word, 3-character-word or 4-character-word, the pronunciation of a polyphone can be determined. So, the author creates a table with non-high-frequency pronunciation polyphones, *point.dbf*. There are only two fields in this table, *cizu* and *yintouxg*. The field *cizu* denotes the non-high-frequency pronunciation polyphone, *yintouxg* denotes the word's CPIs. Both of them have the same type, character, and their widths are 8 and 4 respectively. In this table stores all place name of Tianjin including polyphone. The algorithm of CPIs amendment is:

```
For place name in Names
Match every word in point.dbf
  If matched
    Replace yintou with yintouxg
```

Up to now, through modifying structure, generating CPIs, amending polyphones initials, CPI inquiry function has been implemented in police GIS.

C. The issue of coincident codes

Although CPI inquiry needs not to input Chinese characters, because of a character says with CPI, a lot of Chinese characters will have the same CPI, so must think about the coincident code problem. However, when the number of Chinese characters is more than 4 or 5, the rate of coincident code will reduce. In conclusion, CPI inquiry is very practical and general in MIS.

4. Application of CPI Inquiry

In ArcGIS software, we create a new map, named *Tianjin.mxd*, add 27 layers data, modify the *Attribute table* respectively, append *yintou* in last column, and fill the Chinese Phonetic Initial we have got. The *Attribute table* of *paichusuo_point* is shown in Fig. 4.

978-1-935068-40-2 © 2010 SciRes.

Figure4. Attributes table

In police GIS, open the existed map *Tianjin.mxd*, choose the menu "CPI INQUIRY", will call the Chinese Phonetic Initial inquiry function. For example: inquiry the place name, CPIs "XYLPCS", First, choose the corresponding layer, *paichusuo_point*; Second, input the CPIs "XYLPCS", only one record is listed in table, as shown in Fig. 5.

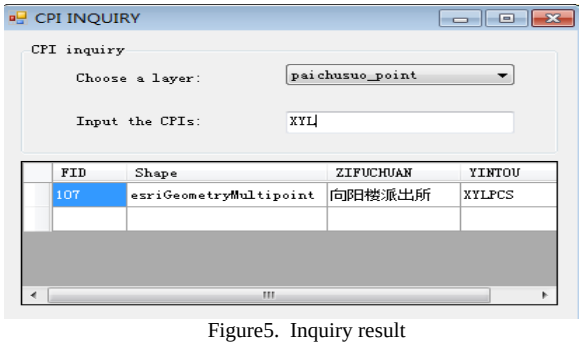

Figure5. Inquiry result

Double click the record of destination name, this place will be shown in center screen and twinkle in police GIS, just as shown in Fig.6.

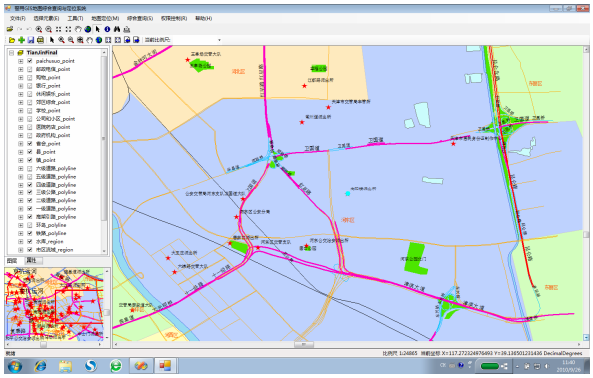

Figure6. Show in map

5. Conclusion

Compared with the traditional method of acquiring Chinese Phonetic Initial by Chinese code, this paper introduced method stores Chinese Phonetic in data dictionary, it gives full play to Visual FoxPro's advantages in disposing and storing mass data. In program implementations, C#, SQL Server 2005 and Visual FoxPro integrate that is the most popular language technology in GIS development system and implementation of Chinese Phonetic Initial inquiry, the portability is also good.

Because development of GIS needs to dispose mass basic geographic information and data, including vector-data and viewdata, the processing speed is slow. So in the future, we need to be further perfecting the data storage method in Chinese Phonetic Initial for enhancing the inquiry speeds.

References

[1] Li Rong, "Several kinds of main modes of Police GIS construction.", Police GIS information technology and application,Vol.3, Pp. 10-13, 2010.

[2] Su Li, Qian Mo, Jianjun Zhou, "A Component-Based GIS System for Police Applied to Alarm Receipt and Disposal", Computer Science and Software Engineering, 2008 International Conference on. Page(s) 151 -154.

[3] Fan Wen-you, Chen Ye, "Design and Study on Police Intelligence Analysis System Based on GIS", Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on.Page(s) 1- 4.

[4] Chen Zhiguo, "The exploration and practice of Police GIS constrction", Journal of Zhejiang Police College, Vol.111, Pp.75-78, Feb, 2009.

[5] Yin Bangzhi, "Implemention of the Using Chinese Phonemic Notation Initial to Query SQLServer 2005 Database with VC#2008", Computer Knowledge and Technology,Vol.11, Pp. 650, 2007.

[6] Li Qi, "Analysis and implement of Query Based on First Spell Alphabet of Chinese Characters", Journal of Sichuan Institute of Light Industry and Chemical Technology, Vol.16, Pp. 72, 2003.

[7] GB2312-80 Chinese-phonetic letters Contrast. http://blog.csdn.net/heiyeshuwu/archive/2007/05/23/1622397.aspx

[8] Shi Jimin, Visual FoxPro and its application system development, Beijing, Tsinghua university Press,2000.

[9] Shi Wei, ArcGIS Geography Information System Details, Beijing: Mechanical Industry Press, 2009.