

Abbreviations Application in 108 VNPT Service Exploitation in Da Nang City

Nguyễn Nho Túy, Phan Huy Khánh

VNPT Da Nang - nhotuy68@gmail.com tuyenn.dng@vnpt.vn;

Da Nang University of Science and Technology, The University of Da Nang - khanhph29@gmail.com

1. ABSTRACT

Looking up and searching for information are always the basic functions of every database (CSDL) management and exploitation application. Information search performance depends not only on resource capacity of systems or searching algorithms but also on time of manipulation and processing on computers of users (NSD). The fact shows that time of typing keywords for information searching of users always take a lot of time, especially for Vietnamese words. It is necessary to give solutions to help users in order to shorten time of keyboard events and enhance operation performance of systems. In this article, we present a solution by developing Abbreviation generating functions (Abbreviation: CVT), re-establishing customer database (CSDL) in the station 108 VNPT Da Nang. The Switchboard 108 VNPT Da Nang has been received practical benefits in information searching for customers owing to practical application of the solution.

Key words : Abbreviation (CVT), database (CSDL), Abbreviation generating functions, Abbreviation data field, information searching

2. Introduction

Telephonists (ĐTV) in the Switchboard 108 VNPT¹ Da Nang (hereinafter referred to as “Switchboard 108”), in their own working shift every day, receive hundreds of calls from customers who request to provide information in all fields. Telephonists receive calls, listen to and

determine contents of requests, then conduct computer events in order to search information and respond the requests. The Switchboard 108 requires every telephonist to always handle all situations in the time as short as possible. However, service exploitation of the Switchboard is undesirable, often obstructed and overloaded. In fact, the telephonists spend over a half of transaction time conducting manipulation on computers: defining main contents of requests, conducting keyboard events (typing Vietnamese words), searching, considering and notifying results. Throughout this process, processing texts and updating search values (keys) are often repeated, often cause unclear names, addresses, geographic locations or regions, etc. On the other hand, almost all transactions, the telephonists have to fast type Vietnamese words. In order to correctly answer customers, the telephonists must type correct requested contents, correct Vietnamese spelling [1], [11] (for examples *sĩ* or *sỹ*, *dành* or *giành*, etc), correct marks, signs (for examples, *oà* or *òa*, *ỳ* or *y*), etc). Obviously, such details significantly increase transaction time that affects the general response performance of the whole system. As a result, the Switchboard 108 is overloaded, then its response capacity is low. Every day, customers often wait a long time for their turn that causes troublesome feeling and affects the psychology of both customers and employees of the Switchboard 108.

Under the pressure of enhancing service performance of the Switchboard 108, we propose a solution based on application of Abbreviation research results [4], [8], available Abbreviation database (CSDL) (<http://chuvietat.com> [2], [6]) and the

transaction database of the Switchboard 108 [7] (hereinafter referred to as database (CSDL) 108).

In addition to the introduction, this article consists of results on analysis of current operation situation of the Switchboard 108, the solution using Abbreviation generating principles in order to re-establish database (CSDL) 108, practical application and assessment of efficiency and benefits, and the last part of the article is conclusion.

3. Analysis of current operation situation of the Switchboard 108

3.1. Transaction activities via phones

The majority of incoming calls from customers require telephonists to search on phone contacts (DBĐT). The requested contents mainly include name of agencies, enterprises or individuals, ... corresponding address and phone number, social information, news, etc. The data about data exploitation to supply customers [7] are counted in the table 1 below.

Table 1. Statistics of Operation data in 2014

| No. | Operation contents of telephonists | Unit | Value |
|-----|--|-------------|--------|
| 1 | Total telephonists in each shift (8 hours) | Telephonist | 14 |
| 2 | Total number of incoming calls | Call | 59 046 |
| 3 | Total number of calls searching on DBĐT | Call | 37 788 |
| 4 | Total maximum time of transaction in a working day | Minute | 366 |
| 5 | Total time of handling for one call searching on phone contacts (DBĐT) | Second | 100 |
| 6 | Total time of information researching on database (CSDL) for a call | Second | 65 |
| 7 | The number of calls to be handled in a day | Call | 220 |

Based on the table 1, we analyze as following :

- A telephonist receives an average of over 160 calls a day but can handle over 220 ones, if the telephonist conducts rational manipulation.
- More than 60% of calls require to visit DBĐT to provide addresses, phone No., etc (37788/59046 calls, in the time of statistics).
- A telephonist spends over 50% of transaction time (183/366 minutes of working in a shift)

searching on DBĐT.

- For each call, telephonist spends little time (~ 40 seconds) listening, explaining, contacting with customers.

3.2. Database organization of the Switchboard 108

The Fig. 1 shows Database system of the Switchboard 108 [7] :

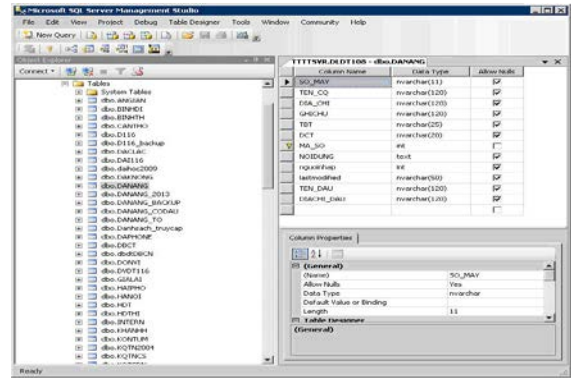


Fig. 1 : CSDL 108 Da Nang.

According to the statistical results, there are 24 groups of database lists as described in the table 2. For example:

- Phone contacts (DBĐT) Group including local database (Da nang and provinces/cities) have the same structure: SO_MAY, TEN_CQ, DIA_CHI...
- Group of database in social fields (tourist, economics, culture, etc) have separate structure. For example, database in tourist include the fields: MA_SO, TIEU_DE, NOI_DUNG, NGAY_NHAP...

Table 2. Statistics of operation database.

| No. | Group | Number of records | Capacity (MB) |
|-----|-----------------------|-------------------|---------------|
| 1 | Phone contacts (DBĐT) | 148 250 | 42.424 |
| 2 | Hot news | 5 914 | 36.155 |
| ... | ... | ... | ... |
| 24 | Monetary | 191 | 0.047 |
| | Total | 215 408 | 153 |

The table 2 is analyzed specifically as following:

- Database (CSDL) 108 often has to innovate contents, monthly there are about 1000 new records (making up 0.46%) that are supplemented, updated.

- Phone contacts (DBDT) group makes up 28% of database (CSDL) 108 (42,4/153MB)
- About 64% of customer's requests require to visit database (CSDL) Phone contacts (DBDT) , the remaining percentage focuses on other requirements.

The results on analysis of current database (CSDL) 108 exploitation situation show that:

- The volume of database (CSDL) 108 is very large, the frequent updating results in duplicated and redundant data.
- Repeated information searching makes telephonists boring, spends a lot of time and causes local transaction obstruction.
- Telephonists often make a mistake due to a short transaction time via phones with various types of customers, due to the difference between the key words for searching and database contents in the database (CSDL) 108 (spelling errors, local words,...).

3.3. Current situation of Abbreviation use in the database (CSDL) 108

Abbreviations are being popularly used today, especially from the explosion of communications services, Internet (messaging, chatting, forum [8], [9] ...). Owing to abbreviations, text manipulation becomes simple, a text is short but convey a larger amount of information. The fact that abbreviations are frequently used not only makes the system of abbreviations more diversified, abundant [4], but also results a lot of difficulty in proper use of languages, circumstances, standardization [9], etc.

Nevertheless, database (CSDL) 108 uses few abbreviations, abbreviations mainly are used in the following cases:

- Name of data field, for example, *TEN_CQ* means name of an agency, enterprises, etc.
- Private names or terms, phrases popularly used, for examples, *TNHH* means limited, etc.
- Code of lists referring to database, for example, *DNG* means Da Nang city, *QNM* means Quang Nam city...
- Contents of database (CSDL) 108 (unmethodical and ununited), for example *TP*

(city), *K* (exhaust), titles such as *TS.*, *PGS...*

4. Solution to abbreviation use

4.1. Establishing abbreviation generating functions

According to theoretical researches into search problems, relevant practical results [3], and efficient use of abbreviations [12], [13], we would like to propose the following principles:

- Abbreviations are only English letters (rather than Vietnamese letters with marks, signs), and digits 0...9.
- Special characters are not used: punctuation marks, space (SP).
- Abbreviations are short and reminiscent: Telephonists can imagine abbreviations immediately after listening to customer's request.

Establishing exercises:

V means a set of Vietnamese letters.

$A = \{A .. Z\}$, means a set of English letters.

$P = \{., ; : ! ? SP\}$ means punctuation marks, space (SP).

$N = \{0 .. 9\}$.

$\Sigma = V \cup A \cup P \cup N$, Σ^* is a set of all sentences (strings) that are established by combining \cup /Concatenation in turn:

- Single word $W \in (\Sigma - \{SP\})^*$ (excluding SP).
For example, $W = 'H' \cup 'ô' \cup 'i' = "Hôi"$.
- A sentence $S \in \Sigma^*$ includes single words W_i , $i=1..L$, in which L means the number of words of S or the length of S excluding space.

Establishing basic functions:

- 1) $Con(W_1, W_2, ... W_L) : \Sigma^* \times ... \times \Sigma^* \rightarrow \Sigma^*$
consecutive concatenation of single words W_i , $i=1..L : Con('N', 'V') = "NV"$.
- 2) $Concat(S) : \Sigma^* \rightarrow (\Sigma - P)^*$
returning a compound sentence including all single words of S after eliminating punctuation marks from S in P :
 $Concat("Ủy ban Nhân dân")$
 $= "ỦybanNhândân"$ (People's committee)
- 3) $ConcatFirstLetter(S) : \Sigma^* \rightarrow (\Sigma - P)^*$
returning a compound sentence including all

first letters of single words of S after eliminating punctuation marks in P.
ConcatFirstLetter(“Hội đồng Nhân dân”) = “HĐND” (People’s Council).

- 4) FirstWord(S) : $\Sigma^* \rightarrow \Sigma^*$
Returning the first single words in S.
FirstWord(“Ủy ban Nhân dân”) = “Ủy”.
- 5) Last(S) : $\Sigma^* \rightarrow \Sigma^*$
returning the last sentence after eliminating the first single words from S.
- 6) LastWord(S) : $\Sigma^* \rightarrow \Sigma^*$
returning the last single word in S.
LastWord(“Ủy ban Nhân dân”) = “Dân”.
- 7) First(S) : $\Sigma^* \rightarrow \Sigma^*$
returning the last part of a sentence after eliminating the last single word from S.
- 8) Upper(W) : $\Sigma^* \rightarrow A^*$
changing W into Capital (without mark) :
Upper(“Hội”) = “HỘI”.
- 9) Unsigned(S) : $\Sigma^* \rightarrow A^*$
returning S but eliminating all marks, “Đ” into “D”
Unsigned(“Đại biểu”) = “Dai bieu”
- 10) Functions for handling addresses:
WordNum(S) : $\Sigma^* \rightarrow A^*$
withdrawing a Word from S, Word is the first word in S including digits 0...9, eliminating digits after “/”:
WordNum(“K266/5, HOANG DIEU”) = “266”
DelNum(S) : $\Sigma^* \rightarrow A^*$
returning S' eliminating Word including digits 0..9:
DelNum(“K266/5, HOANG DIEU”) = “HOANG DIEU”

Based on the established basic functions, we establish abbreviation processing functions whose parameter are S sentence, any content of database (CSDL) 108. S sentence may be:

- Private name (name of people, agencies, enterprises...).
- Addresses.
- Contents of any text.

For example: Developing private name processing functions:

Name of agencies, organizations or enterprises:

Using the function: ConcatFirstLetter(S). For

example,:

ConcatFirstLetter(“Ủy ban Nhân dân”) = “UBND”

Name of people (including first name, middle name and last name):

PeopleName(S) = Unsigned(ConcatFirstLetter(First(S)), ConcatFirstLetter(LastWord(S)))
PeopleName(“Nguyễn Văn Hoàng”) = “NVHOANG”

Title before name of people :

Title(S) = Concat(S)
Title(“PGS. TS.”) = “PGSTS”
Title(“Bác sĩ”) = “BS”

Developing address processing functions:

+ In city:

S = “K266/5, HOANG DIEU”
Address(S) = WordNum(S) + ConcatFirstLetter(DelNum(S)) = “266HD”

+ In countryside:

S = “Tổ 96 Hòa Cường Bắc, Quận Hải Châu”
Address(S) = WordNum(S) + ConcatFirstLetter(DelNum(S)) = “96HCBQHC”

+ Others : For unclear addresses by simple instruction, for example “Cạnh nhà hàng Hải Vân”, we use the function: ConcatFirstLetter(ConcatFirstLetter(“Cạnh nhà hàng Hải Vân”) = “CNHHV”

4.2. Ambiguity processing

Duplicated names or addresses frequently occur, thus there is surely ambiguity of returned results of the abbreviation generating functions. In the application development, we divide the cases to handle as following:

- Applying the abbreviation generating functions for various sentences with the same returned results. For example:
PeopleName(S₁) = PeopleName(S₂) = ...
- The same key words with various returned results (many records of database).

4.3. Establishing Database of Abbreviation

We re-establish database (CSDL) 108 by

supplementing CVT data fields as new key words. The supplementation is as following: S is a data field of database, S is a content of F in a record, then, the data field F1 to be supplemented will have contents of ABBRFUNC(S), ABBRFUNC is a abbreviation generating function as established above.

In fact, we supplement the fields:

- TBT (Abbreviated subscriber) is corresponding to TEN_CQ (name of agencies, enterprises, individuals...).
- DCT (Abbreviated address) is corresponding to DIA_CHI.

After applying the ABBRFUNC functions, the longest abbreviation includes 18 letters, for example, CTTHHNHTMCPPTTPHCM means “Công ty tài chính Trách nhiệm hữu hạn Một thành viên Ngân hàng Thương mại cổ phần phát triển thành phố Hồ Chí Minh”. Owing to the relatively strong server configuration (OS: Windows Server 2008 R2 Enterprise; MySQL Server 5.6; Core: 8; RAM=4GB; capacity of hardware: 200GB, BW: 1Gbps), expenses for calculation on the ABBRFUNC functions and data updating are minor. As a result, we have the database as established in the table 2:

Table 2 : Established database.

Capacity of the database: If there is no TBT, DCT field, DCT will be 42,424MB; if there are TBT, DCT fields, DCT will be 56,117MB, increasing by 13.693MB, equivalent to 132% compared to the last one.

4.4. Application development

Principles of application development ensure that:

- Phone contacts (DBĐT) Search window is designed in a manipulation screen.

- Hot keys are used to fast activate search functions or duplicated manipulation.
- Many search functions (Fig. 3).



Fig. 3 : Search functions

There are some functions as following:

- Finding contacts according to Abbreviated addresses: If a customer request “to find phone number or name of an agency, at K626/11 Trung Nu Vuong”, telephonists shall press F6 to activate the search function, then type “626TNV”, the system will return a list of agencies/ individuals in such address. Telephonists shall respond to the customers based on such list.
- Finding Abbreviated subscriber: telephonists shall press F5 and conduct similar manipulation to find addresses.
- Finding based on the list that was used previously if telephonists wish to obtain more correct information.

5. Result assessment

The table 3 includes operation data after the use of the CVT solution:

Table 3: Statistics of data about the use of CVT solution

| No. | Contents of operation on average | Unit | No using CVT | Using CVT |
|-----|--|--------|--------------|-----------|
| 1 | Searching TBT on Phone contacts (DBĐT) | Call | 22 673 | 22 673 |
| 2 | Handling time/call/telephonist | Second | 55 | 30 |
| 3 | Time for searching of name, address on Phone contacts (DBĐT) /call | Second | 100 | 75 |
| 4 | Calls that a telephonist can be handled per 1 shift | Second | 220 | 293 |
| 5 | Saving time due to keyboard events/call/telephonist | Second | 0 | 25 |
| 6 | Increased calls/call/telephonist | Call | 0 | 73 |

After by comparison of the table 1 and the table 3, here are results:

Data capacity of the new database (additional two TBT, DCT fields) increases to 132% compared to the previous capacity, which is entirely consistent with the current server capacity.

In comparison with the previous system, time of keyboard events of each telephonist significantly reduces, the telephonist only spends 30 seconds for keyboard events rather than 55 seconds. In each 8-hour shift, on average, a telephonist spends 6 hours listening to requests from customers and searching information on database, spends only 1.8 hours conduct keyboard events rather than 3.3 hours, and only 220 calls is handled. 14 telephonists often work in each shift, then saved time will be increased more than 14 times.

6. Conclusions

A solution to information searching by using abbreviations helps to reduce obstruction of calls, reduce time of waiting of customers, bring comfort to telephonists, increase performance, reputation and service quality.

The obtained achievements during establishing and exploiting database by using abbreviations allow us to continue our research in order to perfect and step by step to put various-database exploitation service in use such as fixed contacts, email contacts, especially, the Switchboard 108 will extend the scope of service 108 in the whole The middle and Central Highlands including more than 20 provinces/cities in 2016.

The solution can be further extended by increasing information searching in minority

languages of Vietnam such as Cham, Ede, Thai, Kh'mer, etc according to the needs of the Switchboard 108.

ACKNOWLEDGEMENT

We, authors of the research, would like to express my gratitude to the Switchboard 108, Information Technology Center – VNPT Da Nang and our workmates who facilitated and assisted our re-establishment of CSDL 108 exploitation by using abbreviations.

REFERENCES

- [1] Nguyen Tai Can. *Ngữ pháp tiếng Việt*. NXB Đại học và THCN, Hà Nội 1981.
- [2] Phan Huy Khanh, Nguyen Nho Tuy, *Nghiên cứu xây dựng cơ sở dữ liệu chữ viết tắt cho dịch vụ 1080 Bưu điện Đà Nẵng*. Summary record of National Scientific Conference “*Một số vấn đề chọn lọc của CNTT&Truyền thông*”, 2006.
- [3] Hoang Kiem, Huynh Thuy Bao Trân, *Một mô hình search engine cho tiếng Việt*. Source: www.nsl.hcmus.edu.vn/greenstone/collect/bckh2002/.../CNTT11.pdf
- [4] Nguyen Thanh Viet, Do Kim Bang. *Thuật ngữ viết tắt Viễn thông*. NXB Bưu Điện, 1999.
- [5] M. Zahariev *Abbreviations*. Simon Fraser University, Jun 2004.
- [6] Phan Huy Khanh, Nguyen Nho Tuy. *Setting Up the Database of Abbreviation for Service 1080*. RIVF'07, Vietnam, 2007.
- [7] The Switch 108 VNPT Da Nang. *Biểu mẫu thống kê, hệ thống CSDL quản lý giám sát nghiệp vụ, sản lượng, chất lượng dịch vụ*. Only for internal use, 08/2015.
- [8] Tran Tu Binh, *Viết tắt chữ Việt trong ngôn ngữ @*. Source : <http://vietpali.sourceforge.net/binh/VietTatChuVietTrongNgonNgu-ACong.htm>
- [9] Manh Khang. *Giật mình với lăm kiêu viết tắt của giới trẻ*. Source : <http://tuoitre.vn/tin/nhip-song-tre/20141218/>
- [10] <http://chuvietnhanh.sourceforge.net>
- [11] <http://ngonngu.net>.
- [12] <http://www.Abbreviationfinder.com>
- [13] <http://www.viettat.com>