

## Full Paper

- Struggling Towards a Knowledge-based Society ..... 87  
*Thaweesak Koanantakool*
- ATM Technology for IMT-2000 Mobile Communication Networks..... 96  
*S. Kamolphiwong, S.Sae-Wong, and M. Unhawiwan*
- ระบบระบุผู้พูดภาษาไทยโดยวิธีไดนามิกสไลม์วอร์ปิง ..... 108  
*จุฬารัตน์ ตันประเสริฐ, ชัย วุฒิวิวัฒน์ชัย และวารินทร์ อัจฉริยะกุลพร*
- แนะนำ Bio-Mirror ของฐานข้อมูลชีวภาพของเครือข่ายนักวิจัยภาคพื้นเอเชียแปซิฟิก ..... 119  
*สมนึก คีรีโต และจิตินต์ เขียนดวงจันทร์*
- New VLSI circuits for fast, low cost binary adders ..... 127  
*Vitit Kantabutra, Pasquale Corsonello and Stefania Perri*
- Programmers: A Critique on Communication and Software Engineering Issues in Optical ..... 135  
*Chatpong Tangmanee*

## Invited Short Paper

- Optical Soliton Equation and Its Application in Optical Devices ..... 143  
*Design For High Speed Communication*  
*C. Taeapanich and P.P. Yupapin*

## Short Paper

- Fast Fourier Transform Photoreflectance ..... 148  
*Jiti Nukeaw*
- Optimisation of Switched Mode Power Supply Feedback Compensation Using HSPICE..... 152  
*Chanin Bunlaksananusorn*
- การสร้างเครื่องเชื่อมโลหะอิเล็กทรอนิกส์ โดยใช้หลักการฟูลบริดจ์พีดับบลิวเอ็มแบบเลื่อนเฟส ..... 156  
*พร้อมสวิตช์ที่แรงดันศูนย์*  
*ณัชพงศ์ หัตถิ, ประสิทธิ์ จำปา และพงศ์พิชญ์ วิชาสุรณทล*
- จาก ISO1900 สู่ QS-9000: กลยุทธ์แห่งความสำเร็จ ..... 161  
*วรพจน์ ทั่นะเมธี และสุวรรณ ตรงจิตวิกรัย*
- Call for Papers and Tutorials ..... 166

# ATM Technology for IMT-2000 Mobile Communication Networks \*

เทคโนโลยีเอทีเอ็มสำหรับเครือข่ายสื่อสารเคลื่อนที่ IMT-2000

*S. Kamolphiwong, S. Sae-Wong, and M. Unhawiwan*

*Department of Computer Engineering, Faculty of Engineering,  
Prince of Songkla University, Hatyai, Songkla, Thailand 90112*

**ABSTRACT:** Recently, the mobile network technology is moving from second generation to third generation, known as International Mobile Telecommunications 2000 (IMT-2000). IMT-2000 not only provides conventional services offered by second generation mobile phones but also mobile multimedia services as well as the capacity needed for high traffic volumes. ATM (Asynchronous Transfer Mode) has been chosen as the switching and multiplexing technology for carrying all signals in the future B-ISDN (Broadband ISDN). ATM is expected to be a core of both wireline and wireless networks. In this paper, we present an investigation of ATM network technology for IMT-2000 mobile networks. Our study work will focus on internetworking architecture of IMT-2000 and ATM. We first show some limitations of AAL-1 (ATM Adaptation Layer 1) when it is used for carrying low bit rate voice. In contrast, when assessing the use of AAL-2 for carrying low bit rate voice, AAL-2 offers a number of advantages when compared with AAL-1. We briefly reviewed network architecture issues related to internetworking model. Some alternative ways to apply ATM as a switching technology for mobile networks are presented. We have shown that AAL-2 can carry voice traffic from CDMA (Code Division Multiple Access) mobile networks efficiently. We have addressed some challenge issues which may be concerned of future research topics.

**KEY WORDS:** ATM, IMT-2000, Mobile, Wireless, Internetworking

**บทคัดย่อ:** ปัจจุบันเทคโนโลยีเครือข่ายแบบเคลื่อนที่ที่กำลังพัฒนาจากรุ่นที่สองสู่รุ่นที่สามซึ่งรู้จักในชื่อ IMT-2000 (International Mobile Telecommunications 2000) ความสามารถของ IMT-2000 นั้น ไม่เพียงครอบคลุมการทำงานของเครือข่ายปัจจุบัน แต่ยังครอบคลุมถึงการบริการพาหุสื่อแบบเคลื่อนที่ และการรับส่งข้อมูลความเร็วสูง เอทีเอ็ม (ATM) เทคโนโลยีได้ถูกเลือกให้ทำหน้าที่การสวิตช์และมัลติเพล็กซ์ สำหรับการสื่อสารทุกรูปแบบของเครือข่ายแบบแบนกว้าง เอทีเอ็มเป็นเครือข่ายที่สามารถรองรับการทำงานของเครือข่ายแบบใช้สายและไร้สาย บทความนี้กล่าวถึงการค้นคว้าและวิจัยเกี่ยวกับการเชื่อมต่อระบบ IMT-2000 เข้ากับเครือข่ายเอทีเอ็ม โดยบทความนี้จะเริ่มจากการแสดงให้เห็นถึงข้อจำกัดการใช้งาน AAL-1 (ATM Adaptation Layer 1) ของ ATM ในการรับส่งข้อมูลเสียงความเร็วต่ำ ในขณะที่การใช้งาน AAL-2 สำหรับการรับส่งข้อมูลเสียงแบบความเร็วต่ำจะให้ประสิทธิภาพสูงกว่า AAL-1 ได้แสดงให้เห็นว่าสามารถบริการข้อมูลเสียงจากเครือข่ายเคลื่อนที่ CDMA (Code Division Multiple Access) ได้อย่างมีประสิทธิภาพ สถาปัตยกรรมของเครือข่ายทั้งสองสำหรับการเชื่อมต่อจะถูกกล่าวถึงโดยย่อ วิธีการต่อเชื่อมระหว่างเอทีเอ็มกับเครือข่ายชนิดเคลื่อนที่จะถูกแสดงไว้ หัวข้อที่น่าสนใจอื่นสำหรับการวิจัยดังกล่าวจะกล่าวไว้ในตอนท้าย

**คำสำคัญ:** เอทีเอ็ม ไอเอ็มที-2000 เครือข่ายไร้สาย เครือข่ายเคลื่อนที่ การเชื่อมต่อเครือข่าย

## 1. Introduction

It is clear that Broadband Integrated Service Digital Network (B-ISDN) is playing a central role in the evolution of future networks. B-ISDN is designed as a universal network to deliver a broad diversity of traffic types [1],[2]. ITU (International Telecommunications Union) has chosen

ATM (Asynchronous Transfer Mode) as the switching and multiplexing technology for carrying all signals in the future B-ISDN [3],[4],[5]. ATM uses fixed-length packets, called cells, and is a virtual connection-oriented. The cell length of 53 bytes (5-byte header + 48-byte information field) is an engineering compromise to accommodate conflicting requirements of a whole range of traffic types, be it computer data or real-time traffic such as voice or

\* supported by NECTEC research grant 2542, 3<sup>rd</sup> Generation Mobile Phone (IMT-2000) project.

video. ATM network consists of a set of ATM switches to multiplex/demultiplex traffic streams. Each ATM switch is connected by point-to-point ATM links. Each ATM link can accommodate several VPs (virtual path) and each VP may comprise a number of VCs (virtual channel)[6]. This allows the aggregation of dissimilar types of traffic streams easier to accomplish in one ATM link.

Over the pass few years, wireless mobile communications has been rapidly growing in terms of users, services offered, and geographical coverage areas. Several wireless technologies are already in use, including analogue systems and digital systems based on time division multiple access (TDMA) and Global System for Mobile Communications (GSM). Recently, the mobile network technology is moving from second generation to third generation, known as International Mobile Telecommunications 2000 (IMT-

2000)[10]. IMT-2000 not only provides conventional services offered by today mobile phones but also mobile multimedia services as well as the capacity needed for high traffic volumes, e.g. upto 2 Mbps, as shown in Figure 1 [11]. ATM is expected to be a core of both carrier networks and enterprise networks including wireline and wireless. It seems that ATM technology will be gradually interacted with wireless mobile networks [21],[23],[24],[25]. In this paper, we first briefly review network architecture issues related to internetworking model between ATM and IMT-2000. Some alternative ways to apply ATM as a switching technology for mobile networks are presented. We have shown that AAL-2 can carry voice traffic from CDMA (Code Division Multiple Access) mobile networks efficiently. We have addressed some challenging issues which may be concerned of future research topics.

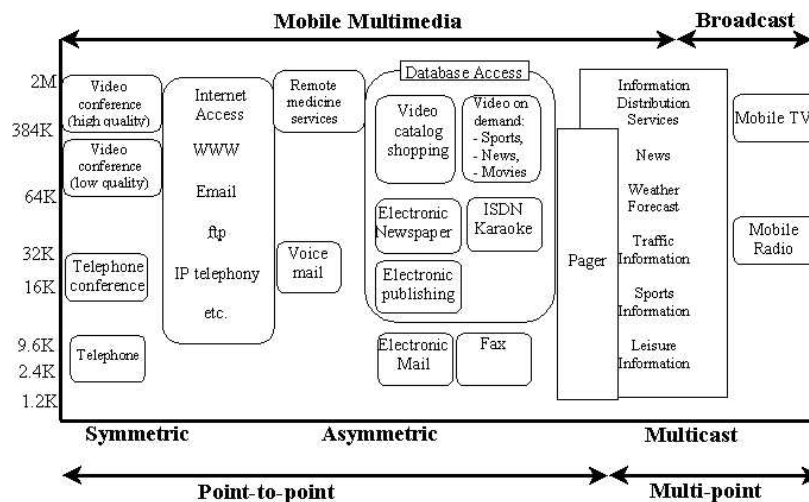


Figure 1 Various services offered by IMT-2000

The paper is organised as follows: Section 2 will describe ATM mechanism for carrying voice traffic. Section 3 will describe the internetworking architecture of IMT-2000 and ATM. We will present some challenging issues to deal with such architecture in section 4. Section 5 concludes our paper.

## 2. ATM Network Technology

ATM is expected to be an efficient transport of heterogeneous traffic with several key advantages as mentioned above. ATM reference model, as defined by ATM Forum [27], is described with reference to Figure 2.

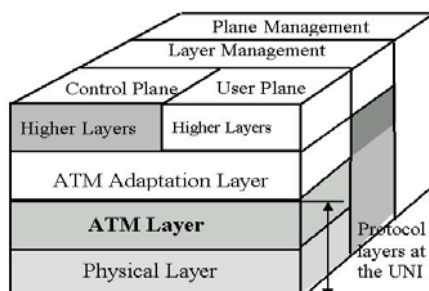


Figure 2 B-ISDN Protocol reference model.

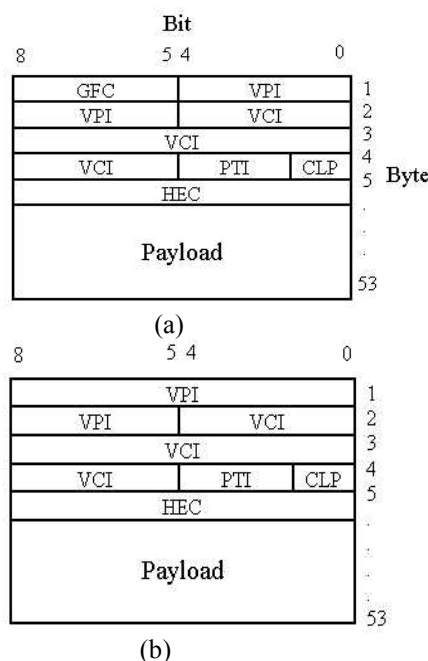
The user plane is concerned with transfer of user information, flow control, and recovery functions. The control plane deals with call connection establishment and release, while the ATM Adaptation Layer (AAL) is concerned with a number of processes necessary to transform the user data stream into a format suitable for ATM. More details of AAL will be described in next section. Among others, the user information is partitioned to form individual ATM cells at the source node, and reassembled again at the destination node.

ATM uses a fixed cell length, 53 bytes long. There are two different formats of ATM cell header, one for use at the User-to-Network Interface (UNI) and the other one for use at the Network-to-Network Interface (NNI) as shown in Figure 3.

Main advantages of ATM are:

- ATM is a connection-oriented network, with each connection setup being assigned with its QoS (quality of service) requirements, e.g. delay, loss and cell delay variation,

- With ATM, the incoming traffic channels are aggregated using statistical multiplexing into one communication link[7]. High system utilisation is easily obtained,
- ATM offers a very high speed of data transmission rate, from few hundreds megabits per second to gigabits per second per channel[8],
- ATM provides multi-priority services of traffic types [9],
- ATM enables the use of single management system providing end-to-end functions, e.g. fault isolation, performance monitoring,
- ATM is scalable technology.



GFC = Generic Flow Control    VPI = Virtual Part Identifier  
VCI = Virtual Channel Identifier    PTI = Payload Type Identifier  
CLP = Cell Lost Priority    HEC = Header Error Check

Figure 3. ATM cell headers: (a) UNI and (b) NNI.

### 3. ATM Mechanism for Carrying Voice Traffic

In this section, we investigate how efficient ATM can carry low-bit-rate voice information. AAL (ATM Adaptation Layer) service is employed to carry voice traffic, both for wireless and wireline networks. The structure of AAL is shown in Figure 4.

In ATM, AAL is considered as a mean of transport to carry all traffic types. There are 5 types of AAL services: AAL-0, AAL-1, AAL-2, AAL-3/4, and AAL-5. AAL-1 and AAL-2 are used for real-time traffic while AAL-0, AAL-3/4 and AAL-5 are used for non-time sensitive traffic. AAL-1 has been standardised in both the ITU-T [12] and ANSI since 1993, is incorporated in the ATM Forum specifications for Circuit Emulation Services (CES), and is offered by several ATM equipment manufacturers. AAL-1 provides the following services to

the AAL user:

- Transfer of service data units with a constant source bit rate and the delivery of them with the same bit rate,
- Transfer of timing information between source and destination,
- Transfer of structure information between source and destination,
- Indication of lost or error information not recovered by AAL-1, if needed.

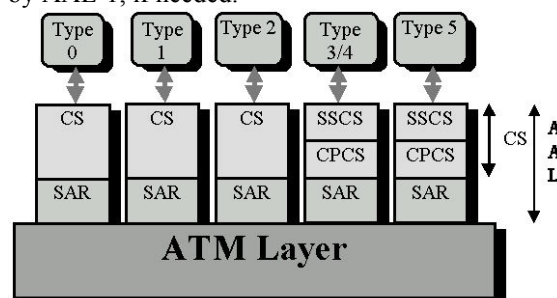


Figure 4. AAL structure and ALL service types.

However, although AAL-1 is offered by many vendors AAL-1 should not be considered an optimum solution for Voice and telephony over ATM (VTOA) with the following reasons:

- Only a single user of the AAL can be supported (per virtual channel),
- Reducing delay requires significant additional bandwidth (if its peak is high),
- Not allow for partially filled cells (always 47 bytes of payload), bandwidth is used even when there is no traffic,
- Voice is always 64kbps or bundles of 64kbps ( $N \times 64$ ), this means that each connection requires at least 64 kbps,
- No standard mechanism in the AAL-1 structure for compression, silence detection/suppression, idle channel removal, or CCS (Common Channel Signalling),
- The bandwidth cannot simply change to meet new application requirements.

As mentioned above, AAL-1 has some limitations and may not operate efficiently to carry low-bit-rate voice traffic. ATM Forum and ITU-T study Group (SG) 13 of a new service in AAL layer that can carry low-bit-rate voice information have been discussed [13]. The primary requirement on the new AAL is to provide efficient transport of small native packets over ATM networks in such a way that allows a very small transfer delay across the ATM network while still allowing the receiver to recover the original packets. To this end, AAL-2 with multiplexing capability is established. New AAL-2 provides the following advantages when compared with AAL-1 [13],[15],[16]:

- Efficient bandwidth usage through VBR service,
- Support for voice compression and silence detection/suppression,
- Support for idle voice channels detection,

- Multiple user channels with varying bandwidth on a single ATM connection,
- VBR ATM traffic class which allows the bandwidth changed.

Figure 5 shows the scenario of how new AAL-2 multiplexes several voice traffic into one channel (one virtual channel (VC) in ATM context). Each 64 kbps voice channel is compressed by using the CS-ACELP scheme [17] with silent suppression. During silent period there is no data to send. New AAL-2 is suitable for compressed wireline telephony, wireless telephony, and wireless data. Some performance results have been evaluated [18],[19],[20],[21].

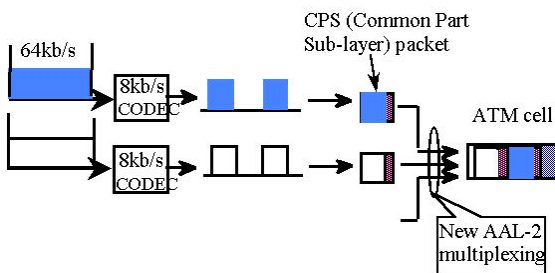


Figure 5. Voice and telephony multiplexing over ATM using AAL2.

To allow AAL2 to multiplex several voice and data traffic into one ATM channel, the cell structure has been standardised in ITU-T SG 13 in AAL-2 [13], as shown in Figure 6 [14]. Multiple voice channels can accommodate in one ATM cell and belong to the same VC.

Figure 7 shows in details how multiple voice channels can be multiplexed into one ATM cell. Each voice signal (indicated by logical link connection (LLC)) is assigned to a short cell (7 octets long). A short cell header, 3 octets long, is added to in front of each short cell. The short cell is referred to as a common part sub-layer (CPS) packet. A channel identifier (CID) is included in the CPS header to identify the original and destination of the CPS packet. A start field (STF), one octet long, is placed in the first byte of a standard ATM payload to indicate starting position of CPS packet in each ATM payload. The STF is needed because the cell payload may contain the remainder of a CPS packet from the previous cell.

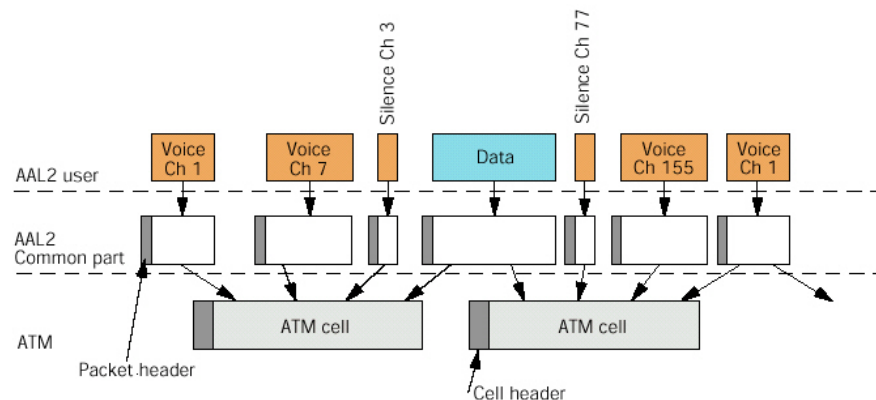


Figure 6. Voice and data multiplexing in AAL-2.

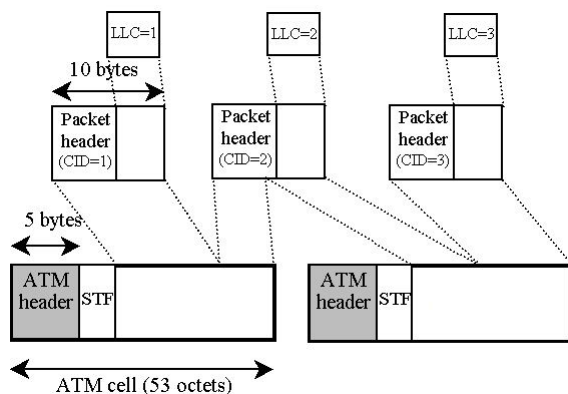
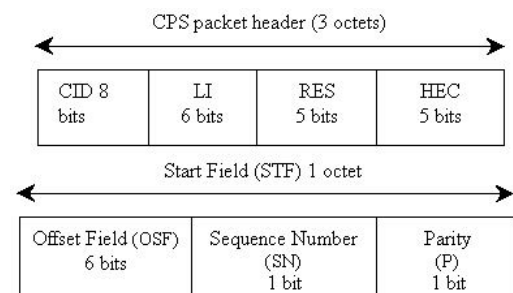


Figure 7. ATM cell structure for supporting AAL-2 multiplexer

Figure 8 shows structure of CPS packet header and start field. CPS packet header is 3 octets long and STF is one octet in length. CID is used to indicate channel number of the original voice source. Length indicator (LI) is the

length of the short packet and header error check (HEC) is used for error checking of each short packet's header. Each field in the STF is used as follows: Offset field (OSF) is the offset position in ATM cell. Sequence number (SN) is a modulus two of the short packet, and parity bit is used for parity check of the STF.



CID=Channel Identifier  
RES=Reserved  
LI=Length Indicator  
HEC=Header Error Check

Figure 8. Structure of packet header and start field.



Figure 9 [22] demonstrates the performance comparison of AAL-1 and AAL-2 to carry voice traffic on T1 link in various time delays. We see that AAL-1 can carry up to 20 voice sources even we increase time delay. AAL-2 with voice compression, 64k and 32k, can carry more voice channels when delay time increases, almost 4 times higher than AAL-1. Figure 10 shows a number of users supported by AAL-1, AAL-2, and AAL-5 when voice coding rate is varied. We can notice that AAL-2 gains a much higher number of users when voice coding rate is low.

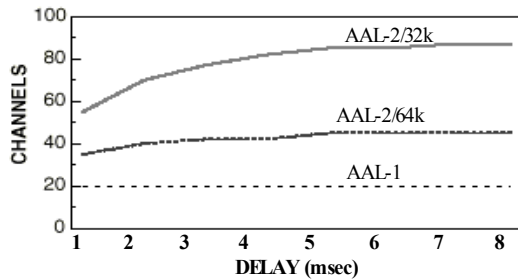


Figure 9. Number of voice channels carried by AAL-1 and AAL-2 when delay is varied.

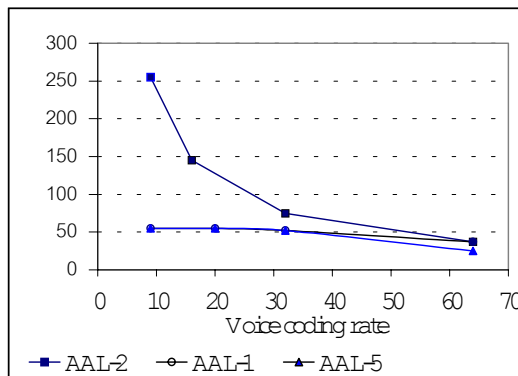


Figure 10. Number of users supported by AAL-1, AAL-2 and AAL-5 when voice coding rate is varied.

#### 4. IMT-2000 Architecture

It is expected that IMT-2000 mobile phones will offer broad-range of services to users; which include high quality audio, real-time video, video conferencing, facsimile, information retrieval such as file transfer and Internet access [29],[30],[31],[32]. The third-generation air interface standardisation for the schemes based on CDMA seems to focus on two main types of WCDMA (Wide-band CDMA): network asynchronous and network synchronous, as shown in Figure 11 [31]. In the former schemes (WCDMA 7 proposed by ETSI and ARIB, and TTA-II 8 WCDMA proposed by Korea) the base stations are not synchronised while in the later schemes (TR45.5 (or cdma2000) proposed by the United States and TTA-I 9 proposed by Korea) the base stations are synchronised to each other within a few microseconds. All schemes are geared towards the IMT-2000 radio transmission technology selection process in ITU-R TG8/1.

The main objectives for the IMT-2000 air interface can be summarised as follows:

- To provide data rates at least 144 Kb/s in a vehicular environment, 384 Kb/s in a pedestrian environment, and 2048 Kb/s in an indoor office environment,
- High spectrum efficiency compared to existing systems,
- High flexibility to introduce new services.

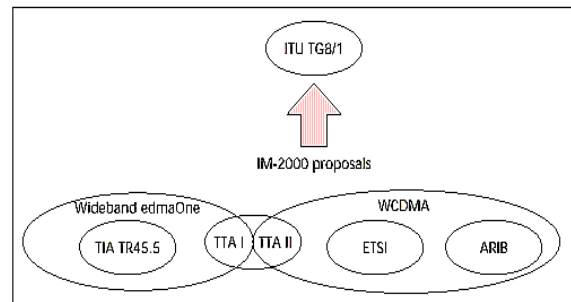


Figure 11. Relationship between WCDMA schemes and standards bodies.

The fundamental differentiation between 2<sup>nd</sup> and 3<sup>rd</sup> generation mobile phones is the much improved data services. The most significant improvements over 2<sup>nd</sup> mobile phone on the signalling are: Change of reverse link design from non-coherent to coherent demodulation, adoption of turbo code by both cdma-2000 and WCDMA developers for high-speed data transmission, and use of ISMA with capture for short data bursts. On the network side, the world of 2nd generation wireless is divided into those supporting roaming protocols based on both GSM/MAP and IS-41. Harmonised IMT-2000 standard is expected to support both as a compromise. Figure 12 shows the coverage area, mobility and information bit rate of various mobile systems.

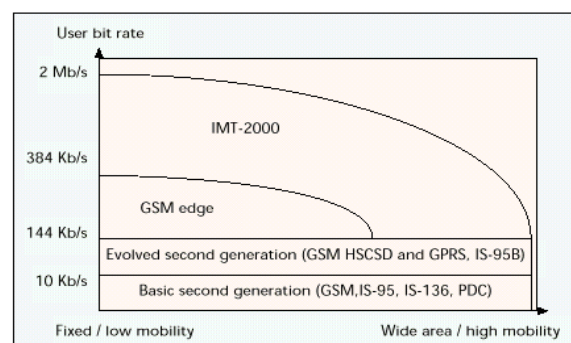


Figure 12. Bit rate versus coverage area/mobility of various mobile phones

#### 5. Internetworking architecture of IMT-2000 and ATM

ATM network is designed for transmission media which is very low bit error rates, e.g. upto  $10^{-10}$ . It was questioned whether ATM will work in the highly noise wireless environment. This problem can be solved by extending media access layer, called *radio access layer*. To do so,

the conversion unit has to be setup between mobile and ATM networks. This means that ATM networks not only are able to carry traffic between mobile and mobile networks but also mobile networks to other kind of networks, e.g. wireline networks, computer/data networks as shown in Figure 13 [25]. Applying ATM technology for IMT-2000 infrastructure networks is efficient. Both are statistical multiplexing and short-packet based-technology. In order to handle various signals from difference mobile phones, direct transport of standard ATM cell may not work effectively when the traffic is low bit rate voice signal. A number of integration of wireless and wireline networks has been proposed [25],[26],[28]. However, there are several points to be considered regarding to mobile and ATM protocol architectures. In this section, we will concern with data link layer, e.g. ATM Adaptation Layer (AAL) and ATM Layer (or cell layer).

Figure 14 shows a possible way to apply ATM as a switching technology for mobile networks. WCDMA (Wide band CDMA) and ATM are both statistical multiplexing and short-packet based-techniques. As a

result, the system takes a high bandwidth utilisation. Therefore, AAL-2 is employed as multiplexing/de-multiplexing to improve transmission efficiency in mobile networks.

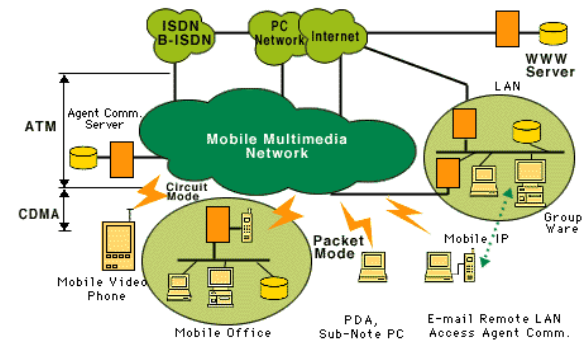


Figure 13. Mobiles, wireline, legacy LANs, and ATM internetworking model

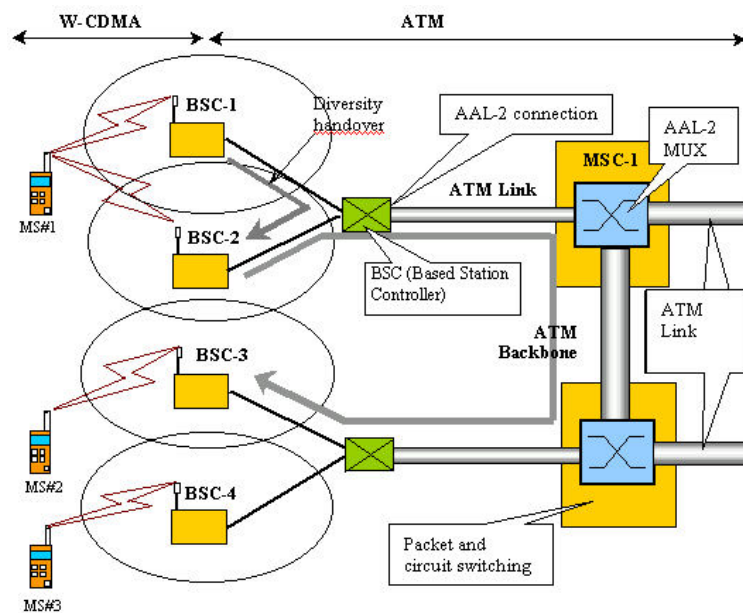


Figure 14. Applying ATM as a switching technology for IMT-2000

Figure 15 demonstrates how to employ AAL-2 as a switching technique for IMT-2000. Layout of protocol stacks of internetworking between IMT-2000 and ATM of Figure 14 can be depicted by Figure 16. We can notice

from the protocol stack that multiplex and demultiplex of WCDMA signals are coupling to AAL-2 layer which also has a multiplexing capability.

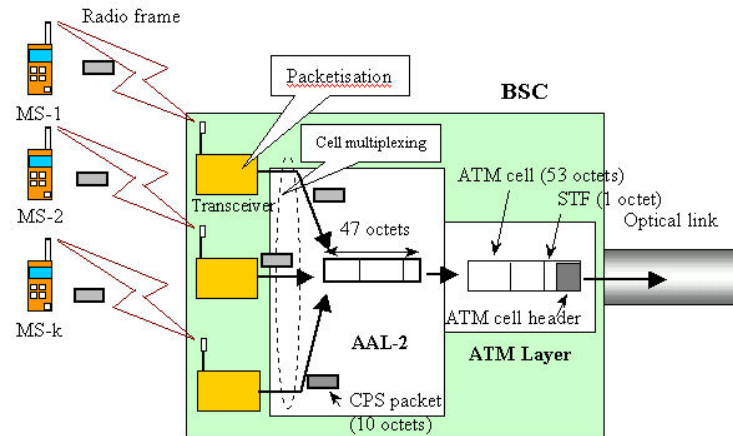


Figure 15. Applying switching AAL-2 in ATM for IMT-200 networks

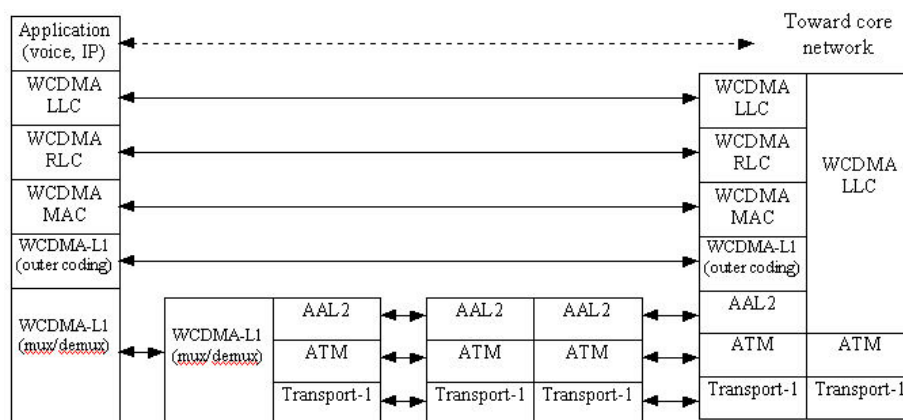


Figure 16. Protocol stacks for internetworking between ATM and IMT-2000

To demonstrate how efficient AAL-2 can carry voice traffic from CDMA mobile networks to a core network, some numerical results have been carried out [34],[35]. In [34], CMDA rate set 2 with a vocoder full rate of 8.1 kbps, 50% at the full encoding rate of 14.4 kbps and 50%

at the rate of 1.8 kbps (silence), is used. The analytical results are summarised in Table 1 with various kinds of carrier services. It shows that AAL-2 performs best. Note that, even frame relay network performs similar results, however, the efficiency gains are not considered.

Table 1 Number of voice calls supported for CDMA rate set 2.

| Transmission facility (Mbps) | Max delay variation (ms) | Number of voice calls supported |             |     |             |
|------------------------------|--------------------------|---------------------------------|-------------|-----|-------------|
|                              |                          | AAL-2                           | Frame relay | TDM | AAL-1/AAL-2 |
| T1 (1.536)                   | 20                       | 123                             | 125         | 24  | 72          |
| T1 (1.536)                   | 5                        | 104                             | 108         | 24  | 72          |
| T3 (44.7)                    | 20                       | 4,090                           | 3,500       | 672 | 2,108       |
| T3 (44.7)                    | 5                        | 3,964                           | 3,023       | 672 | 2,108       |

- **Diversity handover control.** In soft handover, a mobile station is connected to more than one base station simultaneously. Soft handover is used in CDMA to reduce the interference into other cells and to improve performance through macro diversity. Softer handover is a soft handover between two sectors of a cell. Base stations in WCDMA need not be synchronised, and therefore, no external source of synchronisation is needed for the base stations. Asynchronous base stations must be considered when designing soft handover algorithms and when implementing position location services.

In WCDMA, the diversity handover technique is used to improve the quality of service. In order for diversity handover to work efficiently, ATM transmission is required to access network between active BSCs. Sending and receiving signalling messages associated with the handoff procedure can be time-consuming (if using *virtual trees* fashion), leading to large handoff latency delays, and consequently, increasing probability of cell loss. To increase the efficiency of handover, a base station may have ATM switching capacity allowing for ATM connection can be routed directly between adjacent base stations [29].



Figure 7 shows how diversity handover works. In this scenario, radio frames come from BS-1 and BS-2 (Figure 17 (a): up-link). BSC picks the same radio frame number from both BSs. BSC makes a comparison using quality information and then selects the highest quality frame before forwarding to the core network. In the down-link, Figure 17 (b), BSC adds frame number to the user data before putting in ATM cell. The ATM cell is made into two copies before sending to each BS. We note that soft handover can occur among up to three cell sites simultaneously, in which case it is called three-way soft handover.

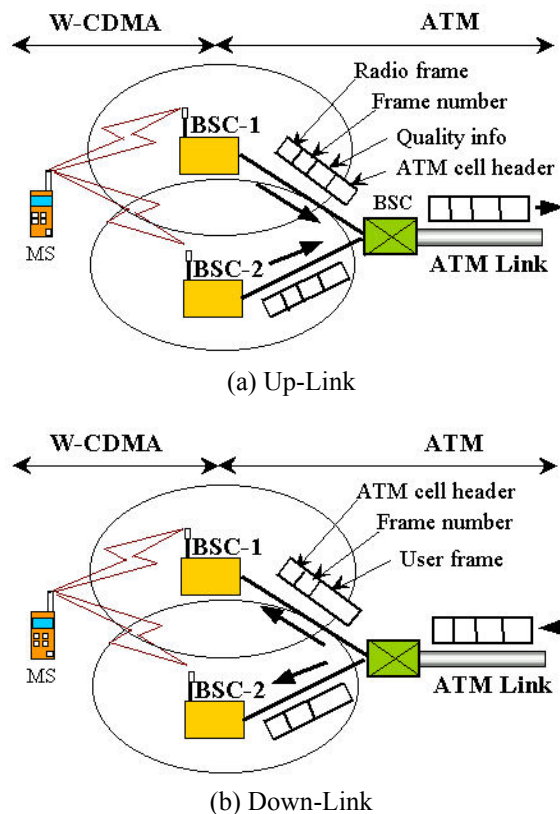


Figure 17. Diversity handover process

Figure 18 illustrates an example of connecting ATM route paths during a mobile moving. In this scenario, mobile phone  $MS_k$  which belongs to cell site  $C_0$  has a conversation to  $MS_i$  which belongs to cell site  $C_5$ . During conversation,  $MS_k$  moves from cell site  $C_0$  to  $C_1$ ,  $C_2$ , and  $C_3$  respectively. A connection between base station and ATM switch (VCC) needs to be established appropriately to an active base station. It may connect to other switches, e.g. switch A to E. If this situation exists, new optimal path may be needed to re-compute, e.g. as shown by dot line.

Figure 19 illustrates an example configuration of the whole network. Each base station may connect to local ATM switch. A number of base stations connected to ATM switch may be an engineering design choice. A group of local switches is connected to a public ATM switch which may connect to other networks.

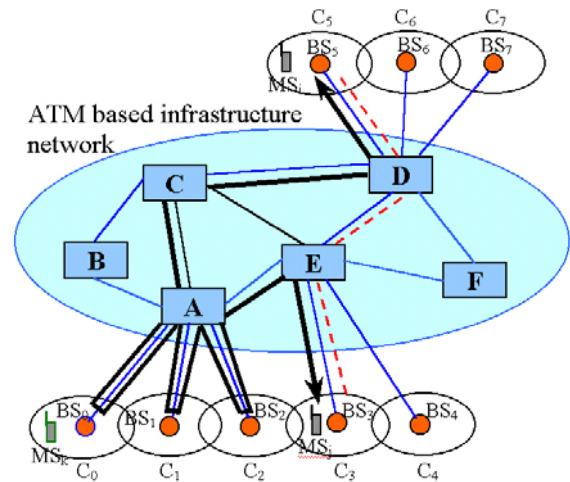


Figure 18. Connection path between mobile stations and ATM switches during mobile station moving.

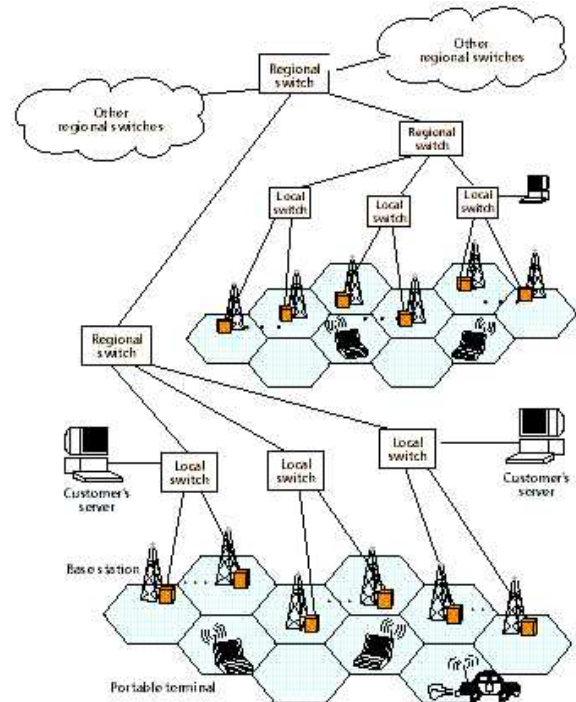


Figure 19. Typical example of internetworking between wireline and wireless networks.

## 6. Challenging Issues

- 6.1. The Internet Engineering Task Force (IETF) has proposed IP Mobility (IPM) based network [33], so called *Packet Telephony* [36] or *Internet Telephony* [37]. This Mobile IP concepts and multimedia services could be [42]. To meet that end, it could be point-to-point, point-to-multi-points, circuit-switch and packet-switch capabilities must be provided. This could be a challenging issue to investigate when IPM used by IMT-2000 travels across ATM networks. Mobility support for IP over wireless ATM is presented in [38],[39].

6.2. Even the future mobile network, such as IMT-2000, is expected to be able to serve heterogeneous traffic, the main traffic may still be voice. *Voice and Telephony over ATM* (VTOA) is a part of an ongoing convergence to a single network technology [15],[40],[41]. When voice and telephony are combined with other services, e.g. electronic mail, voice mail, mobile radio – as shown in Figure 1-, multiple simultaneous access per MS (Mobile Station) must be capable. To that end, an appropriated model need to be employed.

ATM has been chosen as a promising transportation technology in IMT-2000 [42]. Their internetworking, between IMT-2000 and ATM, are in progress. As we note above, MS is able to serve a variety kind of traffic simultaneously, two categories services, AAL2 and AAL5, have been propose [42]. However, some technical issues may not be clearly defined [43],[44], especially in AAL5 (which is the best effort service). Moreover, ABR (Available Bit Rate) service which has been chosen by ATM Forum [45] to delivery non-time sensitive/lose sensitive traffic in ATM networks needs to be studied on its performance/algorithm when interoperated with IMT-2000.

6.3. Figure 20 shows the coverage and bandwidth of existing and forthcoming wireless technologies. *Wireless ATM* could be considered as a next generation of mobile networks. IMT-2000 will provide voice and data transmission rate upto 2 Mbps while wireless ATM will provide access rate upto 25 Mbps [46]. Both systems will have some interfacing parts to exchange their data. This issue may be concerned.

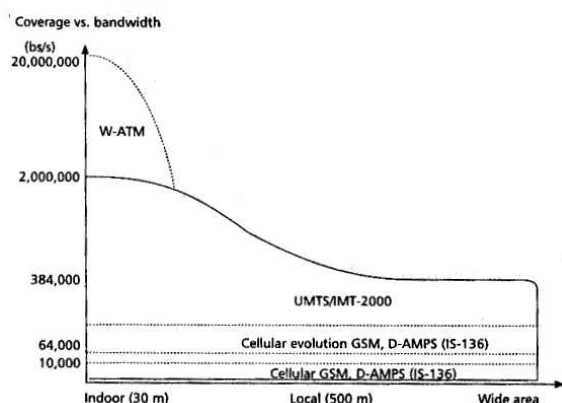


Figure 20. Coverage and bandwidth of wireless systems

6.4. Global roaming and internetworking between various wireless systems via more than one kind of intermediate/core networks is challenging. For example, Figure 21 shows the interconnectivity of the IMT-2000 systems and identifies all possible network-to-network interfaces (NNIs) in a global roaming environment [47]. In some cases of interconnection may contain more than one physical representation of each type of generic functional

networks. In addition, the physical representations may not be directly interconnected, making it necessary to use a transit or intermediate network. This may create a high complexity protocol.

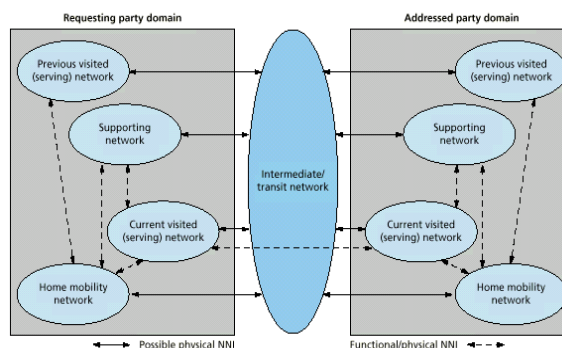


Figure 21. IMT-2000 interconnectivity in a two-domain global roaming environment.

6.5. In ATM networks, we need to consider an overhead used by each service. Table 2 shows that even AAL-2 can carry the largest number of voice channels compared with other AALs, however, it has the biggest overhead; almost 40 percent of used bandwidth. It is interested how to reduce the overhead of this scheme while maintaining a similar performance. Enhanced mechanism may be proposed.

Table 2 comparison of overhead and number of voice channels between various AAL services

| Service Types    | Overhead % | Voice Channels |
|------------------|------------|----------------|
| AAL-1            | 11         | 72             |
| AAL-2 (Original) | 15         | 72             |
| AAL-2 (New)      | 38         | 123            |
| AAL-3/4          | 32         | 72             |
| AAL-5            | 24         | 72             |

## 7. Conclusion

In this paper, we have shown the performance of ATM networks to carry low bit rate voice traffic. AAL-2 (ATM Adaptation Layer 2) protocol with channel multiplexing capability has been designed to support low bit rate delay-sensitive services to delivery the required QoS and maintain high efficient resource utilisation. We have found that ATM offers a number of advantages over other technologies, e.g. frame relay, TDM. Applying ATM technology for IMT-2000 infrastructure networks is efficient. Both are statistical multiplexing and short-packet based-technology. As a result, a network gains a high system utilisation and packetisation delay could be eliminated. Furthermore, ATM offers a fast mobility and soft handover supported in W-CDMA based-mobile networks. However, some works are on going in progress, some challenging issues have been addressed in this paper.

## References

- [1] K. Asatani and S. Nogami, "Standardization of Network Technologies and Services," *IEEE Communications Magazine*, pp.82-90, August, 1995.
- [2] S.L. Sutherland and J. Burgin, "B-ISDB Interworking," *IEEE Communications Magazine*, pp.60-63, August, 1993.
- [3] J. Kim, T. Suda, and M. Yoshimura, "International Standardization of B-ISDN," *Computer Networks and ISDN Systems*, vol. 27, 1994, pp. 5-27.
- [4] ITU-T Recommendation I.150, *B-ISDN Asynchronous Transfer Mode Functional Characteristics*, 1993.
- [5] ITU-T Recommendation I.261, *B-ISDN ATM Layer Specification*, 1993.
- [6] R. Händel, M. N. Huber, and S. Schröder, *ATM Networks: Concepts, Protocols, Application*, 2<sup>nd</sup> Edition, Addison-Wesley Pub. Co., Inc., 1994.
- [7] K. Sriram, "Methodologies for Bandwidth Allocation, Transmission Scheduling, and Congestion Avoidance in Broadband ATM Networks," *Computer Networks and ISDN Systems*, V.26, pp.43-59, 1993.
- [8] K. Genda, N. Yamanaka, Y. Arai, and H. Kataoka, "A High-Speed-Retry Banyan Switch Architecture for Giga-Bit-Rate B-ISDN Networks," *Communication System*, V.7, pp.223-229, 1994.
- [9] G. Gallassi, G. Rigolio, and L. Fratta, "Broadband Assignment in Prioritized ATM Networks," *IEEE GLOBECOM*, pp.852-856, 1990.
- [10] ITU-T Draft Recommendation Q.FNA (Ver.9), *Network Functional Model for IMP-2000*, Sept 1997.
- [11] H. Nakamura, "Applying ATM to Mobile Networks Infrastructure," *IEEE Communications Magazine*, January, 1998, pp.66-73.
- [12] ITU-T Recommendation I.363.1, 1993.
- [13] ITU-T Recommendation I.363.2, "B-ISDN ATM Adaptation Layer Type 2 Specification", 1997.
- [14] Jan Höller, "Voice and telephony networking over ATM," *Ericsson Review* No. 1, 1998.
- [15] M. S. Chambers, H. Kaur, T. G. Lyons, and B. P. Murphy, "Voice over ATM," *Bell Labs Technical Journal*, October-December, 1998, pp.176-190.
- [16] D.W. Petr, and et al., "Efficiency of AAL2 for Voice Transport: Simulation Comparison with AAL1 and AAL5," *IEEE INFOCOM'99*, pp.896-901.
- [17] ITU-T Recommendation G.729, "Coding of Speech at 8 kbps using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP)," Nov. 1995
- [18] H. Saito, "Performance Evaluation of AAL-2 Switch Networks," *IEICE Trans on Communications*, V.E82-B, No.9, Sept 1999, pp.1411-1423.
- [19] G. Mercankosk, J. E. Siliquini, and Z. L. Budrikis, "Provision of Real-time Services over ATM using AAL type 2," *Mobicom'99*, pp.83-90.
- [20] C. Liu, and et al., "Packing Density of Voice Trunking using AAL-2," *GLOBECOM'99*
- [21] G. Mercankosk, J. E. Siliquini, and Z. L. Budrikis, "Provision of Real-time Services over ATM using AAL type 2," *WONMOM'98*, pp.83-91.
- [22] M. McLoughlin and K. Mumford, "Adapting Voice for ATM Networks: A Comparison of AAL1 Versus AAL2," *General DataComm*, 1997.
- [23] E. Ayanoglu, "Wireless Broadband and ATM Systems," *Computer Networks*, V.31, 1999, pp.395-409.
- [24] A. T. Campbell, K. Rauhala, "Mobile ATM: State of the Art and Beyond," *IEEE Networks*, January 1998.
- [25] NTT DoCoMo, "IMT-2000 Mobile Communication Networks," [http://www.nttdocomo.co.jp/corporate/rd/tech\\_e/netw11\\_e.html](http://www.nttdocomo.co.jp/corporate/rd/tech_e/netw11_e.html).
- [26] M. H. Bergren, and et. al, "Wireless and Wireline Convergence," *Bell Labs Technical Journal*, Summer 1997.
- [27] ATM Forum. *ATM User-Network Interface Specification, Version 3.0*. Prentice Hall, NJ, 1993.
- [28] G. Bostelmann, B. Depouilly, and A. Weber, "Network Architectures for Multimedia Mobility," *Alcatel Telecommunications Review*, 3<sup>rd</sup> Quarter 1997, pp.222-231.
- [29] E. Ayanoku, and et al., "Mobile Information Infrastructure," *Bell Labs Technical Journal*, Autumn 1996, pp.143-164.
- [30] A. Sasaki, "Recent Standardization Activities on IMT-2000 Radio Transmission Technology in Japan," *IEICE Trans. On Fundamentals*, V.E80-A, N12, Dec 1997, pp.2340-2346.
- [31] R. Prasad and T. Ojanperä, "An Overview of CDMA Evolution Toward Wideband CDMA," *IEEE Communications Surveys*, <http://www.comsoc.org/pubs/surveys>, Fourth Quarter 1998, Vol. 1 No. 1.
- [32] ITU-R Recommendation M687, "Future Public Land Mobile Telecommunication Systems (FPLMTS).
- [33] Internet Engineering Task Force (IETF), "Draft-IETF-Mobileip-IPM-arch-00.txt", September 1998.
- [34] H. John. Baldwin, et.al., "AAL-2 – A New ATM Adaptation Layer for Small Packet Encapsulation and Multiplexing," *Bell Labs Technical Journal*, Spring 1997, pp.111-131.
- [35] N. Gerlich and M. Ritter, "Carrying CDMA Traffic over ATM using AAL-2: A Performance Study," *Technical Report No. 188*, University of Wurzburg, 1997.
- [36] S. R. Ahuja and K. G. Murti, "Packet Telephony," *Bell Labs Technical Journal*, Spring 1997.
- [37] H. Schulzrinne and J. Rosenberg, "Internet Telephony: Architecture and Protocols – an IETF Perspective," *Computer Networks*, V.31, 1999, pp.237-255.
- [38] A. Acharya, J. Li, et al., "Mobility support for IP over Wireless ATM," *IEEE Communication Magazine*, April 1998, pp.84-88.
- [39] A. Acharya, J. Li, et al., "Mobility Management in Wireless ATM Networks," *IEEE Communication Magazine*, November 1997, pp.100-109.
- [40] Z. Chen, K.P. Das, and B.Y.C. Tang, "Voice and Multiservice Network Design over ATM and IP Networks," *Bell Labs Technical Journal*, October-December, 1998, pp.255-272.



- [41] L. Yanoff, "Multiservice Networking," Alcatel Telecommunications Review, 1<sup>st</sup> Quarter 1998, pp.23-31.
- [42] A. Sasaki, M. Yabusaki, and S. Inada, "The Current Situation of IMT-2000 Standardization Activities in Japan," IEEE Communication Magazine, September 1998, pp.145-153.
- [43] B. Lee, "IMT-2000 Transport Network Using AAL Type 2," *Proceedings ICOIN13*, 1999.
- [44] H. Kim and S. Lim, "Applying DSS-2 ATM Parameters to the A-Interface of IMT-2000 Network having Heterogeneous Link Facility," *Proceedings ICOIN13*, 1999.
- [45] F. Bonomi and K.W. Fendick, "The Rate-Based Flow Control Framework for the Available Bit Rate Service," *IEEE Network*, March/April, 1996, pp. 25-39.
- [46] ETSI Project Broadband Access Networks (EP BRAN), <http://www.etsi.org/BRAN/BRAN.htm>
- [47] M. Torabi and R. E. Buhrke, "Third-Generation Mobile Telecommunications and Virtual Home Environment: A Prioritization Analysis," Bell Lab Technical Journal, July-September 1998, pp.50-62.

## Abbreviations

|          |  |
|----------|--|
| AAL      | ATM Adaptation Layer                         |
| AC       | Authentication Centre                        |
| ADPCM    | Adaptive Digital Pulse Code Modulation       |
| AMF      | Authentication Management Function           |
| ATM      | Asynchronous Transfer Mode                   |
| B-ISDN   | Broadband ISDN                               |
| BS       | Base Station                                 |
| CCAF     | Call Control Agent Function                  |
| CCF      | Call Control Function                        |
| CCS      | Common Channel Signalling                    |
| CDMA     | Code Division Multiple Access                |
| CDV      | Cell Delay Variation                         |
| CID      | Channel Identifier                           |
| CN       | Core Network                                 |
| CODEC    | Coder/Decoder                                |
| CPCS     | Common-part Convergence Sub-layer            |
| CPS      | Common Part Sub-layer                        |
| CS       | Convergence Sub-layer                        |
| GLR      | Gateway Location Register                    |
| GMSC     | Gateway MSC                                  |
| GSM      | Global system for mobile communications      |
| HEC      | Header Error Control                         |
| IETF     | Internet Engineering Task Force              |
| IMT-2000 | International Mobile Telecommunications 2000 |
| IP       | Internet Protocol                            |
| IPM      | Internet Protocol Mobility                   |
| ISDN     | Integrated Service Digital Network           |
| ITU      | International Telecommunications Union       |
| LI       | Length Indicator                             |
| LLC      | Logical Link Control                         |
| LMF      | Location Management Function                 |
| MAC      | Medium Access Control                        |
| MCF      | Mobile Control Function                      |

|         |  |
|---------|--|
| MRTR    | Mobile Radio Transmission and Reception    |
| MSC     | Mobile Switching Center                    |
| MT      | Mobile Terminal                            |
| PCM     | Pulse Code Modulation                      |
| QoS     | Quality of Service                         |
| RACAF   | Radio Access Control Agent Function        |
| RAN     | Radio Access Network                       |
| RFTR    | Radio Frequency Transmission and Reception |
| RNC     | Radio Network Controller                   |
| SACF    | Service Access Control Function            |
| SAR     | Segmentation and Reassembly                |
| SCP     | Service Control Point                      |
| SDP     | Service Data Point                         |
| SIBF    | System Access Broadcast Information        |
| SNCF    | Satellite Network Control Function         |
| SRF     | Special Resource Function                  |
| SSCS    | Service-specific Convergence Sub-layer     |
| STF     | Start Field                                |
| TDMA    | Time Division Multiple Access              |
| VBR     | Variable Bit Rate                          |
| VLR     | Visited Location Register                  |
| VOCODER | Voice Coder                                |
| VTOA    | Voice and Telephony over ATM               |
| WCDMA   | Wide-band CDMA                             |

## Biographies



Dr. Sinchai Kamolphiwong received the B.Sc.(EE) and M.E.(EE) degrees in Electrical Engineering from Prince of Songkla University (PSU), Thailand, in 1984 and 1988 respectively. In 1999, he received the Ph.D. (EE) degree from the University of NSW, Australia.

He was with the Department of Electrical Engineering, PSU, Thailand

in 1984 where he tough in microprocessors and system designs. In 1992, he was with the Department of Computer Engineering at the same university where he involved in PABX and system control development projects. He is now an Assistant Professor in the Department of Computer Engineering, PSU, Thailand. He is a team leader of IP telephony development project. He is a research member of 3rd Generation Mobile Phone (IMT-2000) project. His main interest research areas are: flow control in ATM networks, IP networks and packet telephony, high speed networks, computer network protocols, network embedded systems, mobile networks, modelling and simulation in computer networks. He is a member of IEEE, ACM, ComSoc, and Computer.



มัลลิกา อุนหวิวรรณ์ สำเร็จการศึกษาจาก คณะวิศวกรรมศาสตร์ สาขาโทรคมนาคม ณ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2542. เป็นผู้ช่วยวิจัยโครงการพัฒนาโทรศัพท์เคลื่อนที่ยุคที่ 3 ปัจจุบันกำลังศึกษาปริญญาโท ภาควิชาเทคโนโลยีสารสนเทศ และวิศวกรรมไฟฟ้า คณะสถาบันเทคโนโลยีนานาชาติสิรินธร มหาวิทยาลัยธรรมศาสตร์ งานวิจัยที่สนใจ การเข้าถึงข้อมูลโดยการเข้ารหัส ระบบโครงข่ายเคลื่อนที่ไร้สาย



สุธน แซ่หว่าง วิศวกรรมศาสตรบัณฑิต (เกียรตินิยมอันดับ 1) คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ เมื่อปี พ.ศ. 2542 ปัจจุบันเป็นอาจารย์ประจำภาควิชาวิศวกรรมศาสตร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์

เป็นผู้ช่วยวิจัยโครงการพัฒนาโทรศัพท์เคลื่อนที่ยุคที่ 3 และเป็นผู้วิจัยหลักในโครงการพัฒนาระบบไอพีเทเลโฟนนี้ งานวิจัยที่สนใจ: ไอพีเทเลโฟนนี้ เครือข่ายพหุสื่อ การส่งข้อมูลแบบทันเวลาจริง



## ระบบระบุผู้พูดภาษาไทยด้วยวิธีไดนามิกสไลม์วอร์ปิง

วารินทร์ อัจฉริยะกุลพร<sup>1</sup>, ชัย วุฒิวิวัฒน์ชัย<sup>2</sup>, จุฬารัตน์ ต้นประเสริฐ<sup>3</sup>

หน่วยปฏิบัติการวิจัยและพัฒนาวิศวกรรมภาษาและซอฟต์แวร์

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

ชั้น 22 อาคารมหานครยิบซัม ถ.ศรีอยุธยา เขตราชเทวี กรุงเทพฯ 10400

อีเมล: <sup>1</sup>varin@nectec.or.th, <sup>2</sup>chai@nectec.or.th, <sup>3</sup>chulak@nectec.or.th

**ABSTRACT** -- This paper proposes a closed set, text-dependent speaker identification system for Thai language. Speaking text used in this system is Thai digit 0-9. From our preliminary experiments, a pattern matching technique namely Dynamic Time Warping (DTW) has shown a very good performance, especially for text-dependent system. This work consequently focuses on DTW based system with several comparative studies. Changes of number of sound references and some significant DTW parameters, e.g. time-alignment window, are evaluated. Three algorithms for practical implementation of DTW are proposed and compared. Various kinds of popular speech features are conducted in experiment. Furthermore, experiment on longer speaking-text using concatenation of isolated digits is performed with a new efficient decision technique, proposed especially for concatenated speech system.

**KEY WORDS** -- Thai language speaker identification, Dynamic time warping

**บทคัดย่อ** – บทความฉบับนี้นำเสนอระบบระบุผู้พูดสำหรับภาษาไทยแบบระบบปิด และกำหนดคำพูดตายตัวคือเสียงตัวเลขภาษาไทย 0-9 จากการทดลองเบื้องต้นพบว่าวิธีการเทียบเคียงแบบไดนามิกสไลม์วอร์ปิง (Dynamic Time Warping : DTW) เป็นวิธีที่มีประสิทธิภาพสูง และเหมาะสมสำหรับระบบระบุผู้พูดชนิดกำหนดคำพูดตายตัว ในบทความนี้จึงเป็นการวิจัยต่อเนื่องเน้นหนักเฉพาะการใช้เทคนิค DTW โดยมีการทดลองเปรียบเทียบผลการระบุผู้พูด เมื่อเปลี่ยนแปลงจำนวนเสียงอ้างอิง เปลี่ยนแปลงค่าตัวแปรสำคัญในกระบวนการ DTW เช่นค่ากรอบของการเทียบจุด (Time-alignment Window) เปลี่ยนแปลงกฎการตัดสินใจ พร้อมทั้งนำเสนอและเปรียบเทียบเทคนิคการพัฒนา DTW ในทางปฏิบัติ 3 แบบ นอกจากนี้ยังได้ทำการทดลองเปรียบเทียบผล เมื่อใช้คำลักษณะสำคัญของเสียงแบบต่างๆ และการทดลองสำหรับเสียงพูดที่ยาวขึ้น โดยใช้เสียงของตัวเลขโคดต่อกัน สำหรับการทดลองกับเสียงตัวเลขต่อเนื่องนี้ ได้นำเสนอวิธีการตัดสินใจแบบใหม่ซึ่งให้ผลดีกว่าเดิมอีกด้วย

**คำสำคัญ** -- การระบุผู้พูดสำหรับภาษาไทย, ไดนามิกสไลม์วอร์ปิง

### 1. บทนำ

ระบบรู้จำผู้พูด (Speaker Recognition) เป็นระบบหนึ่งที่น่าสนใจในการนำมาใช้กับการระบุบุคคล (Person Identification) นอกเหนือไปจากการตรวจสอบลายนิ้วมือและการตรวจสอบรูปแบบม่านตา โดยส่วนใหญ่มีจุดประสงค์หลักเพื่อพัฒนาระบบรักษาความปลอดภัย ระบบรู้จำผู้พูดสามารถแบ่งได้เป็น 2 ลักษณะ [1][2] คือระบบระบุผู้พูด (Speaker Identification : SID) ซึ่งมุ่งเน้นที่จะให้ระบบจดจำให้ได้ว่าใครเป็นผู้พูด และระบบยืนยันผู้พูด (Speaker verification : SV) ซึ่งมีจุดมุ่งหมายที่จะตรวจสอบว่าผู้พูดเป็นบุคคลที่ระบุไว้หรือไม่ สำหรับ

SID ยังสามารถแบ่งออกเป็นระบบปิด (Closed Set) และระบบเปิด (Open Set) [1][3] ซึ่งการทำงานของระบบเปิด ไม่เพียงแต่ตรวจสอบว่าใครเป็นผู้พูดเท่านั้น ยังสามารถป้องกันบุคคลภายนอกกระทำได้อีกด้วย นอกจากนี้หากมองทางด้านคำพูดที่ใช้ในการรู้จำผู้พูดแล้ว ยังสามารถจำแนกระบบออกได้เป็น แบบกำหนดคำพูดตายตัว (Text Dependent หรือ Fixed text) และแบบไม่กำหนดคำพูด (Text Independent หรือ Free Text) สำหรับบทความนี้ได้นำเสนอระบบระบุผู้พูด ซึ่งเป็นระบบปิดและกำหนดคำพูดตายตัว

โดยทั่วไป ระบบระบุผู้พูดจะประกอบด้วยส่วนประกอบหลัก 2 ส่วน [2] คือ ส่วนการสกัดค่าลักษณะสำคัญ (Feature Extraction) และส่วนการเทียบเคียงรูปแบบเพื่อรู้จำ (Pattern Matching) สัญญาณเสียงที่ผ่านการประมวลผลขั้นต้นเพื่อปรับสัญญาณให้เหมาะสมกับระบบ จะถูกนำมาสกัดค่าลักษณะสำคัญ (Feature) ซึ่งเป็นกลุ่มตัวเลขจำนวนไม่มากนักที่ใช้เป็นตัวแทนสัญญาณเสียงได้ เนื่องจากวัตถุประสงค์หลักคือการแทนสัญญาณเสียงด้วยกลุ่มตัวเลขจำนวนไม่มากและต้องคงไว้ซึ่งลักษณะสำคัญของเสียงด้วย จึงมีการนำหลักการทางกลศาสตร์มาประกอบกับเทคโนโลยีในการเข้ารหัสสัญญาณ (Coding) เพื่อใช้เป็นวิธีในการสกัดค่าลักษณะสำคัญ เมื่อพิจารณาวิธีการสกัดค่าลักษณะสำคัญที่มีการนำเสนอมาแล้ว จะสามารถแบ่งแยกออกได้เป็น 2 กลุ่ม [3][4] ได้แก่ค่าลักษณะสำคัญทางจันท์ลักษณะ (Prosodic Feature) ได้แก่ค่าความถี่มูลฐาน (Fundamental Frequency) ค่าความถี่ฟอร์แมนท์ (Formant Frequency) และค่าพลังงาน (Energy) เป็นต้น และค่าลักษณะสำคัญบนพื้นฐานของสเปกตรัม (Spectral-Based Feature) ซึ่งได้รับการพิสูจน์ว่าให้ผลการรู้จำดีกว่าค่าลักษณะสำคัญในแบบแรก ค่าลักษณะสำคัญกลุ่มนี้ได้แก่ สัมประสิทธิ์การประมาณพหุเชิงเส้น (Linear Prediction Coefficient : LPC) [5] และสัมประสิทธิ์เซปสตรัม (Cepstral Coefficient) ซึ่งมีทั้งสัมประสิทธิ์เซปสตรัมแบบคำนวณมาจาก LPC (Linear Predictive Coding Derived Cepstrum : LPCC) [2] [6][7] และแบบที่คำนวณบนพื้นฐานของชุดตัวกรอง (Filter Bank) และสเกลของความถี่แบบไม่สม่ำเสมอ (Nonuniform Frequency Scale) เช่น สัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient : MFCC) [2][6][8] เนื่องจากสัมประสิทธิ์เซปสตรัมสามารถให้ผลการรู้จำได้ดีกว่า LPC [7][8] จึงมีการพัฒนาค่าลักษณะสำคัญในกลุ่มของเซปสตรัมอีกมากมาย อาทิเช่น เซปสตรัมที่หักค่าเฉลี่ยของตัวมันออก (Cepstral Mean Subtraction : CMS) [8] และเซปสตรัมผ่านการกรองภายหลัง (Postfiltered Cepstrum) (PFL) [8][9][13] เป็นต้น นอกจากนี้ยังมีการคำนวณค่าการเปลี่ยนแปลง (Derivative หรือ Delta) ของค่าสัมประสิทธิ์เหล่านี้เพื่อใช้เป็นค่าลักษณะสำคัญอีกด้วย [1][8]

ส่วนประกอบที่สองของระบบระบุผู้พูด คือส่วนของการเทียบเคียงค่าลักษณะสำคัญของสัญญาณเสียงที่ต้องการรู้จำ กับค่าลักษณะสำคัญของสัญญาณเสียงอ้างอิง มีวิธีการในการเทียบเคียงรูปแบบจำนวนมากที่ถูกนำเสนอและให้ผลการรู้จำที่ดี อาทิเช่น ไดนามิกสไทม์วอร์ปิง (Dynamic Time Warping : DTW) [3][6] ควอนไทซ์เชิงเวกเตอร์ (Vector Quantization : VQ) [10] แบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov Model : HMM) [6][10] และโครงข่ายประสาทเทียม (Artificial Neural Network : ANN) [11] เป็นต้น การเทียบเคียงรูป

แบบ มีพื้นฐานของการคำนวณระยะห่างของรูปแบบที่จะเทียบเคียงกับรูปแบบอ้างอิง การเลือกใช้วิธีการเทียบเคียงแบบใดนั้น ขึ้นอยู่กับงานที่ต้องการจะพัฒนาค่อนข้างมาก เช่น DTW และ ANN เหมาะกับระบบที่กำหนดค่าพาดตายตัว ในขณะที่ VQ และ HMM เหมาะกับระบบแบบไม่กำหนดค่าพาด [1][10] ส่วนจะกำหนดหรือไม่กำหนดค่าพาดที่ใช้ในระบบก็ขึ้นอยู่กับลักษณะของงาน

เนื่องจากงานวิจัยในบทความนี้มุ่งเน้นที่จะพัฒนาระบบระบุผู้พูดแบบกำหนดค่าพาดตายตัว โดยใช้เสียงตัวเลขภาษาไทย 0-9 จึงได้ทำการทดลองขั้นต้น [12] และได้ผลสรุปว่า DTW เป็นวิธีที่มีประสิทธิภาพมาก ในบทความฉบับนี้ ได้ทำการทดลองเจาะลึกเฉพาะระบบระบุผู้พูดโดยใช้วิธี DTW และทำการทดลองเปรียบเทียบผลการรู้จำ เมื่อเปลี่ยนแปลงจำนวนเสียงอ้างอิง เปลี่ยนแปลงค่าลักษณะสำคัญที่ใช้ในระบบ เปลี่ยนแปลงขั้นตอนกระบวนการของ DTW และขั้นตอนการตัดสินใจ (Decision Rule) ในรายละเอียดปลีกย่อย และทำการทดลองสำหรับคำพูดที่ยาวขึ้น โดยใช้เสียงตัวเลขโคตคู่กันอีกด้วย

ในบทความนี้จะกล่าวถึงการสกัดค่าลักษณะสำคัญของเสียงในส่วนที่ 2 กล่าวถึงรายละเอียดของวิธีไดนามิกสไทม์วอร์ปิงในส่วนที่ 3 สำหรับรายละเอียดของระบบระบุผู้พูดที่นำเสนอจะแสดงในส่วนที่ 4 ในส่วนที่ 5 และ 6 จะกล่าวถึงผลการทดลองและบทสรุปตามลำดับ

## 2. การสกัดค่าลักษณะสำคัญของเสียง

ในบทความนี้ได้ทำการทดลองเปรียบเทียบประสิทธิภาพของการใช้ค่าลักษณะสำคัญแบบต่างๆ ซึ่งได้แก่ LPC, LPCC, PFL และ MFCC โดยหลักการสกัดค่าลักษณะสำคัญแต่ละแบบมีรายละเอียดดังนี้

### 2.1 สัมประสิทธิ์การประมาณพหุเชิงเส้น (Linear Prediction Coefficient : LPC) [5]

กำหนดให้  $X$  แทนสัญญาณเสียง และ  $X(n)$  แทนลำดับ (Sequence) ของสัญญาณเสียงที่ผ่านการแปลงเป็นดิจิทัลแล้ว หลักการของ LPC เริ่มต้นจากการกำหนดให้  $\hat{X}(n)$  หาได้จากผลรวมเชิงเส้นของ  $X(n)$  ก่อนหน้าจำนวน  $p$  ตัว (อันดับเท่ากับ  $p$ )

$$\hat{X}(n) \approx \sum_{k=1}^p a_k X(n-k) \quad \dots(1)$$

โดยที่  $a_k$  คือค่า LPC อันดับที่  $k$  ดังนั้น ค่าความผิดพลาด  $e(n)$  จะคำนวณได้จาก

$$e(n) = \hat{X}(n) - \sum_{k=1}^p a_k X(n-k) \quad \dots(2)$$

การนำ LPC มาประยุกต์ใช้กับงานวิจัยเกี่ยวกับเสียงนั้น สัญญาณเสียงจะถูกแบ่งเป็นเฟรม เฟรมละ  $m$  สัญญาณ เมื่อพิจารณาเฟรมที่  $n$  จะได้ค่าความผิดพลาดรวมดังแสดงในสมการที่ (3)

$$E_n = \sum_m \left[ \tilde{X}_n(m) - \sum_{k=1}^p a_k X_n(m-k) \right]^2 \dots(3)$$

หลักการหา  $a_k$  ที่ทำให้  $E_n$  มีค่าต่ำที่สุดทำได้โดยอาศัยการหาอนุพันธ์ของ  $E_n$  เทียบกับ  $a_k$  จะได้สมการที่ (4)

$$R(i) = \sum_{k=1}^p a_k R(i-k), 1 \leq i \leq p \dots(4)$$

โดยที่  $R(i)$  เป็นค่าอัตโนมัติสหสัมพันธ์ (Autocorrelation) หาได้จาก

$$R(i) = \sum_m X(m-i)X(m) \dots(5)$$

เมื่อแก้สมการที่ (4) ซึ่งประกอบด้วยสมการเชิงเส้น  $p$  สมการ จะได้สัมประสิทธิ์  $a_k$  ที่ใช้ในการประเมินค่าสัญญาณจำนวน  $p$  ค่า โดยแปลงสมการที่ (4) ให้อยู่ในรูปของเมทริกซ์ของค่าอัตโนมัติสหสัมพันธ์ ซึ่งจะได้เป็นเมทริกซ์ Toeplitz ซึ่งมีลักษณะสมมาตร และทุกๆสมาชิกในแนวทแยงมุมมีค่าเท่ากัน ดังนี้

$$\begin{bmatrix} R(0) & R(1) & \Lambda & R(p-1) \\ R(1) & R(0) & \Lambda & R(p-2) \\ M & M & O & M \\ R(p-1) & R(p-2) & \Lambda & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ M \\ a_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ M \\ R(p) \end{bmatrix}$$

การแก้สมการเพื่อหาค่าสัมประสิทธิ์  $a_k$  ทำได้โดยวิธี Levinson-Durbin [7] ดังนี้

$$k_m = \frac{R(m) - \sum_{i=1}^{m-1} a_{m-1} R(m-i)}{E_{m-1}}$$

$$a_m(m) = k_m$$

$$a_m(i) = a_{m-1}(i) - k_m a_{m-1}(m-i); 1 \leq i < m$$

$$E_m = (1 - k_m^2) E_{m-1} \dots(6)$$

โดยที่  $E_0 = R(0)$  และ  $a_0 = 0$  ในแต่ละรอบของ  $m$  ค่าสัมประสิทธิ์  $a_m(i)$  เมื่อ  $i = 1, 2, 3, \dots, m$  จะเป็นค่า LPC อันดับที่  $m$  ที่เหมาะสมที่สุด

## 2.2 สัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency

### Cepstral Coefficient : MFCC)

นอกจากเซปสตรัมปกติแล้ว ซึ่งหาได้จากการแปลงฟูริเยร์ย้อนกลับ (Inverse Fourier Transformation) ของค่าลอการิทึม (Logarithm) ของสเปกตรัมสัญญาณเสียง แต่วิธีการที่ให้ประสิทธิผลดีกว่าคือการปรับสเกลของสเปกตรัม ให้อยู่บนสเกลที่เหมาะสมสำหรับการรับฟังของ

มนุษย์ เนื่องจากสัญญาณเสียงพูดในช่วงความถี่ต่ำจะมีความสำคัญมากกว่าช่วงความถี่สูง จึงมีการออกแบบสเกลของสเปกตรัมให้สามารถเก็บรายละเอียดของสัญญาณเสียงช่วงความถี่ต่ำได้มากกว่า เช่น สเกลเมล (Mel scale) และสเกลบาร์ก (Bark scale) [2][6] เป็นต้น โดยการผ่านสัญญาณเข้าไปยังชุดตัวกรอง (Filter Bank) ที่มีการกระจายความถี่บนสเกลเมลตามสมการที่ (7)

$$Mel = 1000 \log_2 \left( \frac{f}{1000} + 1 \right) \dots(7)$$

จะได้ค่าพลังงานเสียง  $\tilde{S}_k$  สำหรับตัวกรองที่  $k$  และเมื่อนำค่าลอการิทึมของพลังงานนี้ มาผ่านการแปลงโคไซน์แบบไม่ต่อเนื่อง (Discrete Cosine Transformation) จะได้ค่า MFCC ลำดับที่  $n$  ดังแสดงในสมการที่ (8) [2][6][8]

$$c_{mel}(n) = \sum_{k=1}^K \log(\tilde{S}_k) \cos(n(k-0.5)\frac{\pi}{K}) \dots(8)$$

อย่างไรก็ตาม การออกแบบชุดของตัวกรองและรูปร่างของตัวกรอง เป็นจุดสำคัญที่ส่งผลกระทบต่ออัตราการเรียนรู้ด้วยเช่นกัน

## 2.3 สัมประสิทธิ์เซปสตรัมแบบคำนวณมาจากการประมาณ

### พันธะเชิงเส้น (Linear Predictive Coding Derived

### Cepstral Coefficient : LPCC)

นอกจากการคำนวณค่าเซปสตรัมโดยตรงดังได้กล่าวในหัวข้อ 2.2 แล้ว ยังสามารถคำนวณค่าสัมประสิทธิ์เซปสตรัมง่าย ๆ ได้จากค่า LPC [2] [7][8] อีกด้วยโดยอาศัยสมการต่อไปนี้

$$c_{lpc}(n) = \begin{cases} a_n, & n = 1 \\ \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k c_{lpc}(n-k) + a_n, & 1 < n \leq p \\ \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k c_{lpc}(n-k), & n > p \end{cases} \quad (9)$$

## 2.4 สัมประสิทธิ์เซปสตรัมแบบผ่านตัวกรองภายหลัง

### (Postfiltered Cepstral Coefficient : PFL)

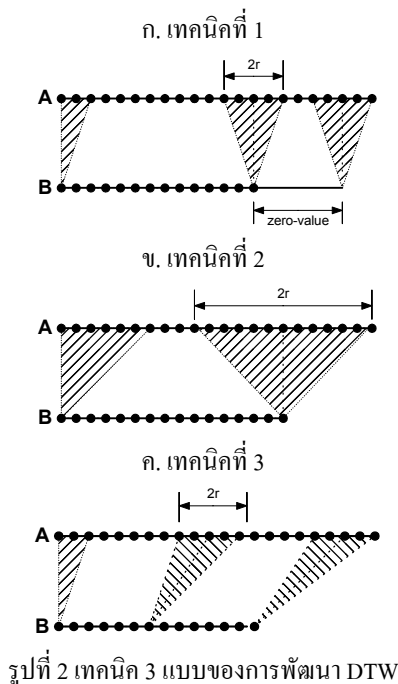
PFL เป็นวิธีการหนึ่งที่พัฒนาเพิ่มเติมต่อจาก LPCC และให้ประสิทธิภาพที่ดีขึ้นทั้งสำหรับการรู้จำผู้พูดและการรู้จำคำพูด (Speech recognition) [13] หลักการของ PFL คือการเน้นสัญญาณเสียงในเชิงความถี่เฉพาะบริเวณที่เป็นความถี่ฟอร์แมนต์ ด้วยหลักการนี้ทำให้สัญญาณเสียงที่ถูกลดทอนด้วยสัญญาณรบกวนยังคงแสดงลักษณะสำคัญของเสียงนั้นออกมาได้ PFL สามารถคำนวณได้ตามสมการที่ (10)

การพัฒนา DTW ตามหลักการที่กล่าวในหัวข้อที่ผ่านมา จะพบว่าเกิดปัญหาขึ้นบางส่วน ปัญหาแรกคือการเลือกค่า  $r$  และจำนวนจุดของการก้าวแต่ละครั้ง (ตามเงื่อนไขข้อที่ 4) ในหัวข้อ 3.1 ให้เหมาะสม การกำหนดค่าที่เหมาะสม จำต้องพิจารณาการทดลองเปรียบเทียบ ปัญหาที่สองคือการกำหนดค่า  $r$  ไว้ตายตัว จะส่งผลให้การจับคู่ลำดับสองลำดับที่มีความยาวต่างกันมาก ให้เป็นไปตามเงื่อนไขที่ 2) ในหัวข้อ 3.1 ได้ไม่เหมาะสมจึงได้นำเสนอเทคนิคเพิ่มเติม 3 แบบ เพื่อให้

สามารถจับคู่ลำดับสองลำดับใดๆ เข้าหากันได้ โดยไม่ขึ้นกับความแตกต่างของความยาวของลำดับทั้งสอง ดังต่อไปนี้

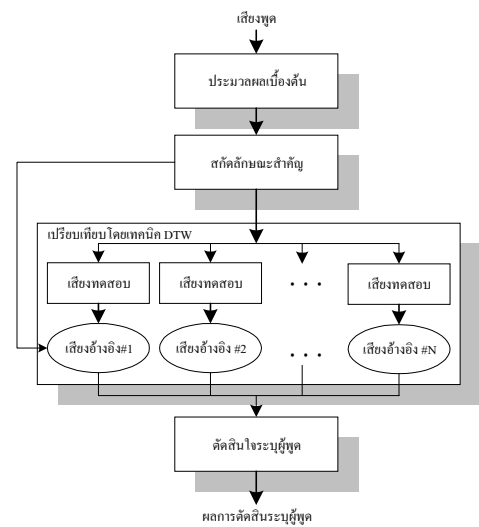
- **เทคนิคที่ 1** กำหนดค่า  $r$  เป็นค่าคงที่ค่าหนึ่ง และทำการเพิ่มจุดที่มีค่าเป็น 0 ในลำดับที่สั้นกว่าเพื่อให้การจับคู่ลำดับตรงตามเงื่อนไขได้
- **เทคนิคที่ 2** กำหนดค่า  $r$  ให้เปลี่ยนแปลงได้ โดยมีค่าเท่ากับผลต่างของความยาวของลำดับสองลำดับที่กำลังจับคู่กัน
- **เทคนิคที่ 3** นับเป็นวิธีที่สมเหตุสมผลที่สุด โดยการขีดจุดบนสองลำดับที่ควรจะถูกเทียบกันก่อน หลังจากนั้นจึงปรับละเอียดเพื่อหาจุดที่เหมาะสมในกรอบ  $r$  ข้างๆ จุดที่กำหนดตอนแรก

เทคนิคการเทียบเคียงด้วย DTW ทั้ง 3 เทคนิคที่กล่าวข้างต้นโดยแสดงได้ดังรูปที่ 2 ในที่นี้กำหนดให้ A และ B เป็นลำดับสองลำดับที่มาเทียบเคียงกัน



#### 4. ระบบระบุผู้พูด

สถาปัตยกรรมของระบบระบุผู้พูดที่นำเสนอในบทความนี้แสดงได้ดังรูปที่ 3 โดยเสียงพูดจะถูกส่งเข้าในกระบวนการประมวลผลเบื้องต้นเพื่อจัดรูปแบบเสียงให้มีความเหมาะสมก่อนเข้าสู่ขั้นตอนการสกัดลักษณะสำคัญ จากนั้นก็จะนำเสียงไปเปรียบเทียบกับเสียงอ้างอิงโดยใช้เทคนิค DTW แล้วนำค่าระยะห่างส่งต่อไปให้ส่วนตัดสินใจระบุผู้พูดทำการตัดสินใจว่าเป็นเสียงของบุคคลใด และส่งผลการตัดสินใจออกมา



รูปที่ 3 สถาปัตยกรรมของระบบระบุผู้พูด

#### 4.1 ส่วนการประมวลผลขั้นต้น (Preprocessing)

เป็นการเตรียมสัญญาณเสียงก่อนสกัดค่าลักษณะสำคัญของเสียง เสียงพูดจะถูกบันทึกด้วยเครื่องคอมพิวเตอร์ผ่านการ์ดเสียงที่มีอัตราการสุ่มตัวอย่าง (Sampling rate) 11.025 กิโลเฮิร์ต โดยเก็บข้อมูลขนาด 16 บิต เนื่องจากการบันทึกเสียงผ่านเครื่องคอมพิวเตอร์จะมีสัญญาณรบกวนจากกระแสไฟฟ้าในเครื่องเกิดขึ้นประมาณ 50 กิโลเฮิร์ต จึงต้องนำสัญญาณที่ได้มาผ่านการกรองด้วยวงจรกรองสัญญาณแบบผ่านความถี่สูงเกิน (High Pass Filter) 200 เฮิร์ต ต่อจากนั้นจึงผ่านการตัดหัวท้ายค่า (Endpoint detection) แบบอัตโนมัติโดยใช้พลังงานเสียงเป็นหลัก [14] เพื่อให้ไม่มีการนำสัญญาณที่ไม่ใช่เสียงพูดจริงเข้ามาในระบบ

#### 4.2 การสกัดลักษณะสำคัญ (Feature Extraction)

เป็นการสกัดลักษณะสำคัญของเสียงของผู้พูดแต่ละคนจากสัญญาณเสียงที่บันทึกไว้ ค่าลักษณะสำคัญชนิด LPC, LPCC, PFL และ MFCC ถูกนำมาใช้ในการทดลอง การสกัดค่าลักษณะสำคัญของเสียงมีขั้นตอนดังนี้

- การเน้นสัญญาณขั้นต้น (Preemphasis) ด้วยการนำสัญญาณเสียงผ่านวงจรกรองอันดับที่ 1 เพื่อให้สัญญาณเสียงชัดเจนขึ้น สัญญาณรบกวนน้อยลง
- การแบ่งสัญญาณเป็นส่วนย่อย (Block to Frame) ขนาด 20 มิลลิวินาที โดยแต่ละส่วนย่อยมีส่วนซ้อนทับกัน 5 มิลลิวินาที



- การวางกรอบแบบแฮมมิง (Hamming Window) หลังจากแบ่งสัญญาณเป็นส่วนย่อยแล้ว ในแต่ละส่วนย่อยจะวางกรอบแฮมมิง เพื่อให้สัญญาณต่อเนื่อง และชัดเจนมากขึ้น
- การสกัดค่าลักษณะสำคัญของเสียงโดยวิธีต่าง ๆ ได้แก่ LPC, LPCC, PFL และ MFCC ดังได้กล่าวแล้วในส่วนที่ 2 ค่าสัมประสิทธิ์ของวิธีต่าง ๆ จะใช้จำนวนเท่ากันหมดคือ 15 อันดับสำหรับแต่ละส่วนย่อยของเสียง โดยค่าอันดับที่ 15 นี้เป็นค่าที่ให้ผลการระบุผู้พูดเบื้องต้นที่เหมาะสม [12]

#### 4.3 การเทียบเคียงรูปแบบสัญญาณอ้างอิงกับ สัญญาณทดสอบ (Pattern Matching)

ในบทความนี้ใช้เทคนิคไดนามิกสไลม์วอร์ปิง (DTW) ซึ่งเป็นเทคนิคที่นิยมนำมาใช้ในการเทียบเคียงรูปแบบ โดยการวัดระยะห่างของสัญญาณเสียงที่เข้ามาทดสอบกับสัญญาณเสียงอ้างอิงในแกนเวลา เทคนิค DTW นี้สามารถจับคู่ส่วนย่อยของสัญญาณเสียงที่เหมาะสมที่สุด และมีระยะห่างน้อยที่สุด รายละเอียดของวิธีนี้ได้กล่าวไว้ในส่วนที่ 3

#### 4.4 การตัดสินใจ (Decision Making) ของการระบุผู้พูด

หลังจากที่ DTW ได้วัดระยะห่างของสัญญาณเสียงที่ไม่ทราบว่าเป็นใครกับสัญญาณเสียงอ้างอิงแล้ว จะใช้วิธีการตัดสินใจเพื่อรู้จำ 2 วิธี วิธีแรกใช้สำหรับกรณีที่มีเสียงอ้างอิงเสียงเดียวต่อหนึ่งผู้พูด วิธีการตัดสินใจคือพิจารณาค่าระยะห่างต่ำที่สุด หรือกล่าวอีกนัยหนึ่งคือมีความคล้ายมากที่สุด (Maximum Likelihood : ML) วิธีที่สองใช้ในกรณีที่มีเสียงอ้างอิงของแต่ละผู้พูดจำนวนมาก วิธีการตัดสินใจใช้วิธีพิจารณา K ลำดับที่ใกล้ที่สุด (K-Nearest Neighbor : KNN) โดยจะพิจารณาสัญญาณเสียงอ้างอิงที่ให้ระยะห่างที่จับคู่ที่น้อยที่สุด K อันดับแรก ในสัญญาณเสียงอ้างอิง K สัญญาณนั้น หากเสียงอ้างอิงของบุคคลใดมีจำนวนมากที่สุด ระบบก็จะตัดสินใจว่าเสียงที่เข้ามาทดสอบเป็นเสียงของบุคคลนั้น และถ้าหากว่ามีเสียงอ้างอิงจาก 2 บุคคลจำนวนเท่ากัน ทำให้ไม่สามารถตัดสินใจได้ ก็จะพิจารณาจาก 1-NN ว่าเสียงอ้างอิงของบุคคลใดมีระยะห่างที่น้อยที่สุดอันดับแรก ก็จะตัดสินใจระบุว่าเป็นผู้พูดเป็นของบุคคลนั้นทันที

### 5. การทดลองและผลการทดลอง

เสียงที่ใช้ในการทดลองบันทึกโดยผู้พูด 50 คน (ชาย 30 คน หญิง 20 คน) โดยบันทึกเป็นเวลา 5 สัปดาห์ต่อเนื่อง ในแต่ละสัปดาห์ผู้พูดทุกคนต้องพูดเสียงตัวเลขโดด 0-9 ตัวเลขละ 10 ครั้ง โดยใช้โปรแกรมอัด

เสียงที่สร้างขึ้นโดยเฉพาะ การพูดเพื่อบันทึกเสียงตัวเลขจะเป็นแบบสุ่มไม่เรียงตามลำดับ 0-9 เพื่อป้องกันความคุ้นเคยในการพูดเลขต่อเนื่องกัน เสียงที่บันทึกได้จากสัปดาห์ที่ 1-3 จะใช้เป็นเสียงอ้างอิง ส่วนสัปดาห์ที่ 4-5 จะใช้เป็นเสียงสำหรับการทดสอบ ดังนั้นสำหรับผู้พูดแต่ละคน จะมีจำนวนเสียงอ้างอิงทั้งหมด 30 เสียง และเสียงเพื่อทดสอบ 20 เสียงต่อหนึ่งตัวเลข

ในการทดลองได้แบ่งการทดลอง 3 แบบด้วยกัน การทดลองแรกคือการทดลองกับเสียงตัวเลขโดด 0-9 โดยจะมีการทดลองหาจำนวนเสียงอ้างอิงที่เหมาะสมจุดมุ่งหมายเพื่อให้ระบบมีความถูกต้องของการระบุผู้พูดสูงและใช้เวลาไม่นานเกินไปนัก การทดลองเทคนิคการประยุกต์ใช้ DTW เพื่อหาวิธีวัดระยะห่างของเสียงอ้างอิงกับเสียงทดสอบที่เหมาะสมที่สุด การทดลองที่สองคือการทดลองกับเสียงตัวเลขคู่กัน เพื่อพิสูจน์ว่าเสียงที่ยาวขึ้นจะทำให้ผลการระบุผู้พูดถูกต้องมากขึ้น โดยได้ทดสอบเปรียบเทียบวิธีการตัดสินใจในการระบุผู้พูดด้วย การทดลองสุดท้ายคือการทดลองเพื่อเปรียบเทียบการใช้ค่าลักษณะสำคัญแบบต่างๆ เพื่อให้รู้ว่าการสกัดค่าแบบใดจะให้รายละเอียดของลักษณะเสียงได้ดี และเหมาะสมกับวิธี DTW

#### 5.1 การทดลองกับเสียงตัวเลขโดด 0-9

##### ก. การทดลองเพื่อหาจำนวนเสียงอ้างอิงที่เหมาะสม

การทดลองนี้มีจุดประสงค์เพื่อหาจำนวนเสียงอ้างอิงที่เหมาะสม เป็นที่ทราบกันดีว่าหากเสียงอ้างอิงมีจำนวนมาก ก็จะทำให้ความถูกต้องของการระบุผู้พูดมีมาก แต่อย่างไรก็ตามมีจำนวนเสียงอ้างอิงมากเกินไปก็เสียเวลาในการประมวลผลมากขึ้น ดังนั้นการทดลองนี้จึงนำเสนอเพื่อหาจำนวนเสียงอ้างอิงที่เหมาะสม 3 ระบบด้วยกันคือ

ระบบที่ 1 ทดลองใช้เสียงอ้างอิงจำนวน 30, 20 และ 10 เสียง ในการทดลองนี้จะนำเสียงที่บันทึก 3 สัปดาห์แรกมาเป็นเสียงอ้างอิง

ระบบที่ 2 ทดลองโดยการหาเสียงอ้างอิงที่ดีที่สุดจำนวน 5 เสียง โดยการใช้เสียงที่บันทึกจากสัปดาห์ที่ 1-3 จำนวน 1,500 เสียงของ 50 คนเป็นเสียงอ้างอิง แล้วใช้เสียงชุดเดียวกันนี้ในการทดสอบ หลังจากนั้นสำหรับผู้พูดแต่ละคน จะเลือกเสียงอ้างอิงที่ดีที่สุดมา 5 เสียง (เสียงอ้างอิงที่อยู่ในกลุ่มที่ให้คำตอบถูกต้องบ่อยที่สุด 5 อันดับแรก) และจะใช้เสียงทั้ง 5 เสียงนี้เป็นเสียงอ้างอิงในการทดสอบกับเสียงในสัปดาห์ที่ 4 และ 5

ระบบที่ 3 ทดลองหาเสียงอ้างอิงจำนวน 1 เสียงจากเสียงอ้างอิงในสัปดาห์ที่ 1-3 โดยสุ่มมา 10 ตัว และหาค่าเฉลี่ยรวมของเสียงอ้างอิงแต่ละตัว โดยอาศัย DTW ในการหาเส้นทางในการจับคู่จุดที่จะเฉลี่ยเข้า

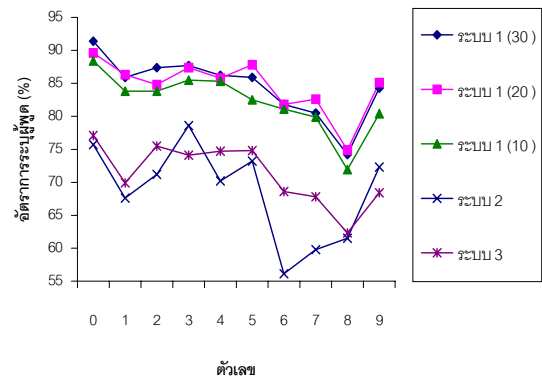
หากัน เพื่อนำมาเป็นเสียงอ้างอิงตัวเดียวในการทดสอบกับเสียงในสัปดาห์ที่ 4 และ 5

ระบบทั้ง 3 นี้ใช้ค่าลักษณะสำคัญแบบ LPCC ในการทดลอง ผลการทดลองแสดงในตารางที่ 1 และกราฟในรูปที่ 4 จะเห็นว่าอัตราการระบุผู้พูดของตัวเลขแต่ละตัวของทั้ง 3 ระบบเป็นไปในทิศทางเดียวกัน คือ เลข “0” มีอัตราการระบุผู้พูดสูงที่สุด ส่วนเลข “8” มีอัตราการระบุผู้พูดต่ำที่สุด ปัจจัยที่มีผลต่ออัตราการระบุผู้พูดคือ วรรณยุกต์ สระ และตัวอักษรที่ประกอบรวมขึ้นมาเป็นคำส่งผลให้อัตราการระบุผู้พูดของแต่ละตัวเลขแตกต่างกัน

ตารางที่ 1 : เปรียบเทียบอัตราการระบุผู้พูดตามจำนวนเสียงอ้างอิง

| เสียง<br>ตัวเลข | อัตราการระบุผู้พูด (%) |               |               |               |               |
|-----------------|------------------------|---------------|---------------|---------------|---------------|
|                 | ระบบที่ 1              |               |               | ระบบ<br>ที่ 2 | ระบบ<br>ที่ 3 |
|                 | อ้างอิง<br>30          | อ้างอิง<br>20 | อ้างอิง<br>10 |               |               |
| 0               | 91.40                  | 89.60         | 88.40         | 75.70         | 77.10         |
| 1               | 85.90                  | 86.30         | 83.80         | 67.60         | 69.90         |
| 2               | 87.40                  | 84.80         | 83.80         | 71.20         | 75.50         |
| 3               | 87.70                  | 87.40         | 85.50         | 78.60         | 74.10         |
| 4               | 86.20                  | 85.80         | 85.30         | 70.20         | 74.70         |
| 5               | 85.90                  | 87.80         | 82.50         | 73.20         | 74.80         |
| 6               | 81.80                  | 81.80         | 81.10         | 56.10         | 68.60         |
| 7               | 80.50                  | 82.60         | 79.90         | 59.80         | 67.80         |
| 8               | 74.20                  | 74.90         | 71.90         | 61.50         | 62.30         |
| 9               | 84.30                  | 85.10         | 80.40         | 72.30         | 68.40         |
| เฉลี่ย          | 84.53                  | 84.61         | 82.26         | 68.62         | 71.32         |

ในระบบที่ 1 อัตราการระบุผู้พูดเฉลี่ยของเสียงอ้างอิงจำนวน 20 เสียงมีค่ามากที่สุด แสดงให้เห็นว่าบางครั้งเสียงอ้างอิงจำนวนมากเกินไป อาจไม่ให้ผลลัพธ์ที่ดีที่สุดก็ได้ และยังทำให้เสียเวลามากขึ้นในการประมวลผลเพื่อระบุผู้พูด ระบบที่ 2 มีอัตราเฉลี่ยการระบุผู้พูดต่ำที่สุด แม้ว่าจะมีจำนวนเสียงอ้างอิงมากกว่าระบบที่ 3 แสดงให้เห็นว่าการเลือกเสียงอ้างอิงที่ดี โดยการใช้ชุดทดสอบใดๆ จะได้เสียงอ้างอิงที่ดีสำหรับชุดทดสอบนั้นเท่านั้น เมื่อเปลี่ยนชุดทดสอบไป ผลที่ได้จะต่ำมาก



รูปที่ 4 กราฟแสดงอัตราการระบุผู้พูดของเสียงอ้างอิงที่ต่างกัน

## ข. การทดลองเทคนิคการประยุกต์ใช้ DTW

ได้ทำการทดลองกับการประยุกต์ใช้เทคนิค DTW 3 แบบที่กล่าวในหัวข้อ 3.2 โดยสร้างระบบทดลอง 3 ระบบดังนี้

ระบบที่ 1 กำหนดค่า  $r = 5$  (เงื่อนไขข้อที่ 3 ในหัวข้อ 3.1) ใช้เสียงสัปดาห์ที่ 1-3 จำนวน 30 เสียงของแต่ละผู้พูดเป็นเสียงอ้างอิง หากเสียงที่นำมาทดสอบหรือเสียงอ้างอิงใด ที่สั้นกว่าก็จะเพิ่มค่า 0 ต่อท้ายเสียงสั้นจนมีความยาวเท่ากับเสียงที่ยาวกว่า

ระบบที่ 2 ให้ค่า  $r$  แปรผันตามความแตกต่างระหว่างความยาวของเสียงอ้างอิงและเสียงที่เข้ามาทำการทดสอบ และใช้เสียงอ้างอิง 1 เสียงโดยหาค่าเฉลี่ยของเสียง 10 เสียงที่สุ่มมาจากเสียง 30 เสียงในสัปดาห์ที่ 1-3

ระบบที่ 3 กำหนดค่า  $r = 5$  และใช้จำนวนเสียงอ้างอิง 30 เสียงเหมือนระบบที่ 1 แต่จะมีการจับคู่ส่วนย่อยที่เหมาะสมเข้าหากันก่อนตามสัดส่วนของความยาวของเสียงที่มาจับคู่กัน

ผลการทดลองแสดงในตารางที่ 2 ระบบที่ 1 กับระบบที่ 3 มีอัตราการระบุผู้พูดเฉลี่ยพอๆ กัน แต่สูงกว่าระบบที่ 2 มาก ทั้งนี้เพราะระบบที่ 2 มีจำนวนเสียงอ้างอิงเพียงเสียงเดียว จากผลการทดลองเราไม่สามารถสรุปได้ทันทีว่าระบบที่ 1 ดีกว่าระบบที่ 3 เพราะบางตัวเลขในระบบที่ 3 ให้อัตราการระบุผู้พูดสูงกว่าระบบที่ 1 ถึงแม้ว่าทั้งสองระบบจะให้อัตราเฉลี่ยการระบุผู้พูดพอกัน แต่ในรายละเอียดของการประมวลผลก็มีข้อแตกต่างกันอยู่ ประการแรกคือระบบที่ 3 จะมีการคำนวณที่ซับซ้อนกว่าในการจัดการจับคู่ส่วนย่อยของเสียง ประการที่สองนอกจากนี้ในทางทฤษฎีแล้ว การเทียบเคียงแบบระบบที่ 3 ถือว่าสมเหตุสมผลมากกว่าเพราะเป็นวิธีที่จับคู่ตามสัดส่วนของเสียงจึงทำให้การวัดค่าระยะห่างมีความถูกต้องมากกว่าโดยเฉพาะเสียงที่มีช่วงยาว

ตารางที่ 2 เปรียบเทียบอัตราการระบุผู้พูดของ DTW ทั้ง 3 เทคนิค

| เสียง<br>ตัวเลข | อัตราการระบุผู้พูด (%) |           |           |
|-----------------|------------------------|-----------|-----------|
|                 | ระบบที่ 1              | ระบบที่ 2 | ระบบที่ 3 |
| 0               | 91.40                  | 77.10     | 92.30     |
| 1               | 85.90                  | 69.90     | 90.60     |
| 2               | 87.40                  | 75.50     | 87.80     |
| 3               | 87.70                  | 74.10     | 85.40     |
| 4               | 86.20                  | 74.70     | 86.20     |
| 5               | 85.90                  | 74.80     | 85.60     |
| 6               | 81.80                  | 68.60     | 78.70     |
| 7               | 80.50                  | 67.80     | 78.20     |
| 8               | 74.20                  | 62.30     | 73.20     |
| 9               | 84.30                  | 68.40     | 84.90     |
| เฉลี่ย          | 84.53                  | 71.32     | 84.29     |

## 5.2 การทดลองกับเสียงตัวเลขต่อกัน

การทดลองที่ผ่านมาเป็นการทดสอบกับตัวเลขโดดเพียงตัวเดียว ซึ่งงานวิจัยที่ผ่านมาจะระบุว่าอัตราการรู้จำจะสูงขึ้นเมื่อเสียงมีความยาวมากขึ้น ดังนั้นในการทดลองนี้จะนำตัวเลขโดดมาต่อกัน และใช้ในการทดสอบเพื่อศึกษาอัตราการรู้จำเปรียบเทียบกับเสียงตัวเลขโดด

### ก. การทดสอบเพื่อหาความยาวของตัวเลขต่อกันและค่า $r$ ที่เหมาะสม

กลุ่มผู้วิจัยได้ทดสอบระบบระบุผู้พูดสำหรับผู้พูด 20 คน โดยใช้เสียงตัวเลขโดด 0-9 และใช้ค่าลักษณะสำคัญแบบ LPC [15] แล้วนำตัวเลขโดดที่ให้ผลอัตราการระบุผู้พูดสูงสุด 3, 5 และ 7 อันดับแรกมาต่อกัน ตัวเลขที่ได้คือ “145”, “14539” และ “1453902” จุดประสงค์ของการทดลองนี้ เพื่อพิสูจน์ว่าหากเสียงผู้พูดมีความยาวมากขึ้นจะทำให้อัตราการระบุผู้พูดเพิ่มมากขึ้น และเมื่อขนาดของเสียงยาวขึ้น ปัจจัยที่เกี่ยวข้องก็คือค่า  $r$  โดยจะมีการปรับค่า  $r$  ให้มีขนาดต่างกันออกไปตามความยาวของเสียง ในที่นี้ใช้ค่า  $r = 5, 10$  และ  $20$  ผลการทดลองแสดงในตารางที่ 3

เสียงที่มีความยาวขึ้นให้อัตราการระบุผู้พูดสูงกว่าค่าที่มีพยางค์เดียว อีกเหตุผลหนึ่งน่าจะมาจากผลของส่วนประกอบวรรณยุกต์ในเสียง ซึ่งเคยแสดงให้เห็นใน [16] นอกจากนี้ส่วนประกอบอื่นๆ ของเสียง เช่น พยัญชนะ และสระ ก็น่าจะส่งผลให้อัตราการระบุผู้พูดเปลี่ยนแปลงไปได้ จากตารางที่ 3 ค่า  $r$  มีผลต่ออัตราการระบุผู้พูดอย่างชัดเจน ผลการทดลองที่ได้สูงสุดคือ 98.75% ( $r=20$ ), 100% ( $r=20$ ) และ 99.20%

( $r=10$ ) โดยเป็นผลของตัวเลขที่มีอัตราการระบุผู้พูดสูงสุด 3 ตัว 5 ตัว และ 7 ตัวต่อกันตามลำดับ จากผลการทดลองแสดงให้เห็นว่าจะต้องกำหนดค่า  $r$  ให้เหมาะสมกับขนาดความยาวของเสียงด้วย เพราะถ้าเสียงมีความยาวมากเรากำหนดค่า  $r$  น้อยก็จะทำให้ได้ผลการระบุผู้พูดต่ำ แต่มีข้อที่น่าสังเกตจุดหนึ่งคือ สำหรับตัวเลขต่อกัน 7 ตัว อัตราการระบุผู้พูดสูงสุดที่ค่า  $r = 10$  ในขณะที่ เลขสูงสุดต่อกัน 3 ตัว และ 5 ตัว ให้อัตราการระบุผู้พูดสูงสุดที่  $r = 20$  เหตุที่ได้ผลเช่นนี้อาจเป็นเพราะว่าถ้าตัวเลขต่อกันมีความยาวมากเกินไปจะทำให้ระบบสับสนในการคำนวณการจับคู่ระหว่างช่วงของเสียงอ้างอิงและเสียงทดสอบเพื่อหาระยะห่างที่น้อยที่สุด อย่างไรก็ตาม การใช้ค่า  $r$  มากกว่า 20 อาจจะทำให้ผลต่ออัตราการระบุผู้พูดสูงเกินไปก็ได้ แต่เนื่องจากการเพิ่มค่า  $r$  มากเกินไป จะทำให้เสียเวลาในการประมวลผลมากเกินไปจนความจำเป็น นอกจากนี้การใช้ตัวเลขสูงสุดต่อกัน 5 ตัวก็เพียงพอที่จะให้ผลอัตราการระบุผู้พูดสูงถึง 100% แล้ว เมื่อพิจารณาปัจจัยทางด้านเวลาในการประมวลผล และอัตราการระบุผู้พูดที่ได้แล้ว จะพบว่าการใช้ตัวเลขต่อกันยาว 3 ตัวเหมาะสมที่สุดสำหรับระบบระบุผู้พูดนี้

ตารางที่ 3 อัตราการระบุผู้พูดของเสียงที่ต่อกัน

| ตัวเลข                       | $r = 5$ | $r = 10$ | $r = 20$ |
|------------------------------|---------|----------|----------|
| สูงสุด 1 อันดับ<br>“1”       | 95.83   | 94.58    | -        |
| สูงสุด 3 อันดับ<br>“145”     | 96.67   | 98.330   | 98.75    |
| สูงสุด 5 อันดับ<br>“14539”   | 96.83   | 97.50    | 100.00   |
| สูงสุด 7 อันดับ<br>“1453902” | 92.92   | 99.20    | 98.75    |

### ข. การทดลองเปรียบเทียบวิธีการตัดสินใจสำหรับตัวเลขต่อกัน

การทดลองนี้ได้นำผลการทดลองที่ได้จากหัวข้อที่ 5.1(ก) โดยนำตัวเลขที่ให้อัตราการระบุผู้พูดสูงสุด 3 อันดับแรกมาต่อกันคือ “053” สำหรับเสียงอ้างอิงจำนวน 20 เสียง เลข “032” สำหรับเสียงอ้างอิงจำนวน 30 เสียง มาใช้เป็นเสียงในการทดสอบ จากการทดลองที่ 5.2(ก) พบว่ากำหนด  $r = 20$  ให้อัตราการระบุผู้พูดสูงสุดสำหรับเลข 3 ตัวต่อกัน แต่พบว่าหากเพิ่มค่า  $r$  จาก 5 ไปเป็น 20 จะทำให้เสียเวลาในการประมวลผลเพิ่มมากขึ้นเป็นทวีคูณ ดังนั้นจึงทำการทดลองเปรียบเทียบ 2 ระบบที่จะเสนอต่อไปนี้

ระบบที่ 1 วิธีนี้เป็นวิธีปกติเหมือนในการทดลองที่ 5.2(ก) คือใช้เทคนิค DTW กำหนดค่า  $r = 20$  และใช้ 5-NN สำหรับการตัดสินใจระบบผู้พูด

ระบบที่ 2 แทนการใช้ DTW เพื่อเปรียบเทียบระยะทางของเสียงที่ต่อกันแบบระบบที่ 1 แล้วระบบผู้พูด ในระบบนี้จะใช้ DTW เปรียบเทียบตัวเลขโดดแต่ละตัวก่อน แล้วนำผล 5-NN ของตัวเลขทั้งสามมารวมกันเพื่อใช้ 15-NN ในการตัดสินใจระบบผู้พูด

ผลการทดลองแสดงในตารางที่ 4 จะเห็นว่าระบบที่ 2 มีอัตราการระบุผู้พูดสูงกว่าระบบที่ 1 อีกทั้งยังใช้เวลาในการประมวลผลน้อยกว่าระบบที่ 1 ด้วย จึงสรุปได้ว่าสำหรับเลขที่ต่อกันการรวมค่า KNN ที่คำนวณจากตัวเลขแต่ละตัวจะให้อัตราการระบุผู้พูดที่ดีกว่าวิธีธรรมดา

ตารางที่ 4 อัตราการระบุผู้พูดของเสียงต่อกัน  
โดยใช้การตัดสินใจระบบผู้พูดต่างกัน

| ระบบ      | อัตราการระบุผู้พูด (%) |                  |
|-----------|------------------------|------------------|
|           | อ้างอิง 20 เสียง       | อ้างอิง 30 เสียง |
| ระบบที่ 1 | 95.40                  | 95.10            |
| ระบบที่ 2 | 96.10                  | 96.50            |

### 5.3 การทดลองเพื่อเปรียบเทียบการใช้ค่าลักษณะสำคัญ

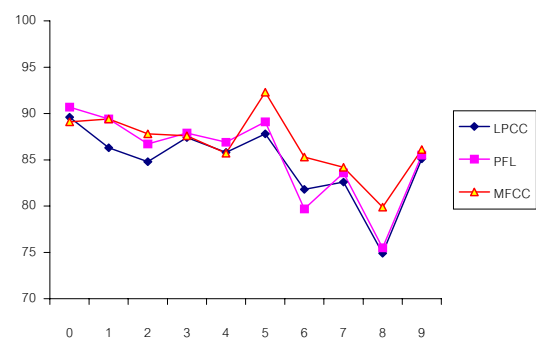
การทดลองนี้ได้เปรียบเทียบการสกัดลักษณะสำคัญ 3 แบบด้วยกันคือ LPCC, PFL และ MFCC โดยใช้ตัวเลขโดด 0-9 ของผู้พูด 50 คน และเพื่อให้การทดลองมีตัวคงที่เหมือนกัน จึงกำหนดให้ใช้จำนวนสัมประสิทธิ์ของค่าลักษณะสำคัญเท่ากับ 15 ลำดับ เท่ากันหมด ผลการทดลองแสดงในตารางที่ 5 และกราฟรูปที่ 5 จากผลการทดลองแสดงให้เห็นว่า MFCC ให้ผลการระบุผู้พูดที่ดีที่สุด และ LPCC ให้ผลที่ต่ำสุด ผลเป็นไปตามบทความ [17] MFCC ให้ผลการระบุผู้พูดสูงสุด 92.30% กับเลข “5” PFL ให้ผลดีกว่า LPCC เหมือนที่พิสูจน์ในบทความ [9] แต่อย่างไรก็ตามอัตราการระบุผู้พูดขึ้นกับเสียงของตัวเลขแต่ละตัวด้วย

ถึงแม้ว่า MFCC จะเป็นตัวแทนที่ดีที่สุด แต่ข้อด้อยก็มีอยู่บ้างคือ MFCC จะต้องมีการตั้งค่าตัวแปรจำนวนมากกว่า LPCC ซึ่งจะต้องเลือกให้เหมาะสมกับเสียง เพื่อให้ค่าลักษณะสำคัญที่สกัดออกมาใช้ได้ ซึ่งลักษณะของแต่ละบุคคลมากที่สุด ตัวอย่างเช่น การเลือกจำนวนของตัวกรองในชุดตัวกรอง รูปทรงของตัวกรอง การตั้งค่าช่วงของความถี่และความกว้างของช่องสัญญาณตัวกรอง นอกจากนั้นเรายังไม่สามารถสรุปได้ว่าจะต้องตั้งค่าตัวแปรอย่างไรจึงจะเหมาะสมที่สุดกับเสียงที่เก็บในฐานข้อมูลเว้นแต่ทำการทดลองเท่านั้น

ตารางที่ 5 อัตราการระบุผู้พูดของการสกัดคุณลักษณะ

แบบ LPCC PFL และ MFCC

| เสียง<br>ตัวเลข | อัตราการระบุผู้พูด (%) |       |       |
|-----------------|------------------------|-------|-------|
|                 | LPCC                   | PFL   | MFCC  |
| 0               | 89.60                  | 90.70 | 89.10 |
| 1               | 86.30                  | 89.40 | 89.40 |
| 2               | 84.80                  | 86.70 | 87.80 |
| 3               | 87.40                  | 87.90 | 87.60 |
| 4               | 85.80                  | 86.90 | 85.70 |
| 5               | 87.80                  | 89.10 | 92.30 |
| 6               | 81.80                  | 79.70 | 85.30 |
| 7               | 82.60                  | 83.60 | 84.20 |
| 8               | 74.90                  | 75.50 | 79.90 |
| 9               | 85.10                  | 85.50 | 86.10 |
| เฉลี่ย          | 84.61                  | 85.50 | 86.74 |



รูปที่ 5 กราฟรูปแสดงการเปรียบเทียบ

LPCC, PFL และ MFCC

การทดลองสุดท้ายของการเปรียบเทียบการสกัดค่าลักษณะของเสียงทั้ง 3 แบบ คือการนำเสียงตัวเลขโดดที่ให้อัตราการระบุผู้พูดสูงสุด 3 อันดับแรก จากตารางที่ 2 มาต่อกัน และใช้วิธีการตัดสินใจในการระบุผู้พูดเหมือนกับระบบที่ 2 ในหัวข้อที่ 5.2(ข) คือการรวมค่า 5-NN จากสามตัวเลข แล้วใช้ 15-NN ในการตัดสินใจ ผลการทดลองแสดงในตารางที่ 6

ตารางที่ 6 อัตราการระบุผู้พูดของการสกัดคุณลักษณะแบบ LPCC, PFL, และ MFCC โดยนำตัวเลขมาต่อกัน 3 อันดับแรก

| การสกัดคุณ<br>ลักษณะ | ตัวเลขสูงสุด<br>3 ตัว | อัตราการระบุผู้พูด |
|----------------------|-----------------------|--------------------|
| LPCC                 | “053”                 | 96.50              |
| PFL                  | “015”                 | 98.80              |

|      |       |       |
|------|-------|-------|
| MFCC | "510" | 98.70 |
|------|-------|-------|

จากตารางที่ 6 จะเห็นได้ว่าผลอัตราการระบุผู้พูดสูงขึ้นเป็นอย่างมากตามที่คาดไว้และ LPCC ยังคงให้อัตราการระบุผู้พูดต่ำที่สุดเช่นเดิม ในขณะที่ PFL และ MFCC ยังคงให้อัตราการระบุผู้พูดในระดับใกล้เคียงกัน

## 6. บทสรุป

บทความนี้เปรียบเทียบการใช้เทคนิค DTW กับเสียงตัวเลขโดด 0-9 ของภาษาไทยในหลาย ๆ แ่งมุม โดยเปรียบเทียบจำนวนเสียงอ้างอิงที่ใช้ ซึ่งสรุปได้ว่าไม่ควรใช้เสียงอ้างอิงจำนวนมากเกินไป เพราะจะทำให้เสียเวลาในการประมวลผลมากขึ้น อีกทั้งบางครั้งอาจทำให้อัตราการระบุผู้พูดลดลงได้ การเปรียบเทียบเทคนิคย่อยของการใช้ DTW โดยแยกย่อยได้ 3 แบบ แบบแรกคือการกำหนดให้ช่วงกรอบการจับคู่ ( $r$ ) คงที่แล้วเติมสัญญาณว่าง (ค่า 0) ต่อท้ายเสียง แบบที่ 2 กำหนดให้ค่า  $r$  แปรผันตามความยาวของเสียง และแบบที่ 3 คือ กำหนดให้ค่า  $r$  คงที่ แต่มีการจับคู่จุดที่เหมาะสมตามสัดส่วนความยาวของเสียงที่มาเทียบเคียงกัน พบว่าแบบที่ 3 มีความเหมาะสมผลที่สุดในการหาระยะห่างที่ต้องระหว่างเสียงอ้างอิงกับเสียงทดสอบ อย่างไรก็ตามระบบแบบที่ 3 นี้สิ้นเปลืองเวลาในการคำนวณมาก อีกทั้งให้ผลการทดสอบใกล้เคียงกับแบบที่ 1 และได้ทำการเปรียบเทียบระหว่างเสียงตัวเลขโดดและเลขต่อกัน เสียงที่มีขนาดยาวกว่ายอมให้อัตราการระบุผู้พูดที่สูงกว่า สุดท้ายคือเปรียบเทียบการสกัดลักษณะสำคัญของเสียง 3 แบบคือ LPCC, PFL, และ MFCC สรุปว่าการใช้ MFCC สกัดลักษณะสำคัญของเสียงได้ผลเฉลี่ยรวมดีที่สุด สำหรับเทคนิคที่ค้นพบใหม่ในบทความนี้ คือการใช้ผลรวมค่า KNN ของตัวเลขโดดแต่ละตัวของเลขที่นำมาต่อกัน พบว่าให้อัตราการระบุผู้พูดสูงกว่าวิธีปกติ และยังใช้เวลาประมวลผลในการระบุผู้พูดน้อยกว่าด้วย

งานวิจัยในอนาคตสำหรับการใช้เทคนิค DTW คือจะศึกษาตัวแปรแต่ละตัวให้เหมาะสม เช่นการจำกัดระยะเวลาในการจับคู่เสียง การหาความยาวเสียงที่ทำให้ DTW เปรียบเทียบแล้วมีประสิทธิภาพมากที่สุด และการลดเวลาในการเปรียบเทียบ ผลการวิจัยในส่วนต่าง ๆ เหล่านี้จะส่งผลให้การพัฒนาระบบระบุผู้พูดสำหรับภาษาไทยสามารถนำไปใช้งานได้จริงในอนาคตอันใกล้

## เอกสารอ้างอิง

[1] J. P. Campbell, Jr., "Prolog to Speaker Recognition: A Tutorial", *Proceedings of IEEE*, Vol. 85, No. 9, p. 1436-1462, September 1997.

[2] S. Furui, "Digital Speech Processing, Synthesis, and Recognition", New York and Basel: Marcel Dekker, Inc, 1989.

[3] G. R. Doddington, "Speaker Recognition-Identifying People by their Voices", *Proceedings of IEEE*, Vol. 73, No. 11, p.1651-1664, November 1985.

[4] J. M. Naik, "Speaker Verification: A Tutorial", *IEEE Communications Magazine*, pp. 42-48, January 1990.

[5] D. O'Shaughnessy, "Linear Predictive Coding", *IEEE Potentials*, p. 29-32, February 1988.

[6] L. R. Rabiner and B. -H. Juang, "Fundamentals of Speech Recognition", A. Oppenheim, Series Editor, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[7] S. Furui, "Cepstral Analysis Technique for Automatic Speaker Verification", *IEEE Transaction on Acoustic, Speech Signal Processing*, Vol. ASSP-29, pp.254-272, April 1981.

[8] R. J. Mammone, X. Zhang, and R. P. Ramachandran, "Robust Speaker Recognition, A Feature-based Approach", *IEEE Signal Processing Magazine*, p. 58-71, September 1996.

[9] M. S. Zilovic, R. P. Ramachandran, and R. J. Mammone, "Speaker Identification Based on the Use of Robust Cepstral Features Obtained from Pole-Zero Transfer Functions", *IEEE Transactions on Speech and Audio Processing*, Vol.6, No.3, pp.260-267, May 1998.

[10] K. Yu, J. Mason, and J. Oglesby, "Speaker Recognition using Hidden Markov Models, Dynamic Time Warping and Vector Quantisation", *IEE Proc.-Vis. Image Signal Process*, Vol. 142, No. 5, October 1995.

[11] R. A. Finan, A.T. Sapeluk, R. I. Damper, "Comparison of Multilayer and Radial Basis Function Neural Networks for Text-dependent Speaker Recognition", *1996 IEEE International Conference on Neural Networks (IJCNN'96)*, Vol. 4, pp. 1992-1997, 1996.

[12] วสิน สันธิบุญญ, เปรมนาถ คูเบ, สุทัศน์ แซ่ตั้ง, วารินทร์ อัจฉริยะกุลพร ชัย วุฒิวิวัฒน์ชัย และจุฬารัตน์ ต้นประเสริฐ, "การระบุผู้พูดด้วย LPC และ DTW สำหรับภาษาไทย", *เอกสารประกอบการประชุมวิชาการ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีประจำปีงบประมาณ 2542 ณ ศูนย์ประชุมสหประชาชาติ 30 มีนาคม-เมษายน 2542*

[13] M. S. Zilovic, R. P. Ramachandran, and R. J. Mammone, "Speaker Identification Based on the Use of Robust Cepstral Features Obtained from Pole-Zero Transfer Functions", *IEEE Transactions on Speech and Audio Processing*, Vol.6, No.3, pp.260-267, May 1998.

[14] ชัย วุฒิวิวัฒน์ชัย, "การรู้จำเสียงคำหลายพยางค์แบบไม่ขึ้นกับผู้พูด โดยใช้เทคนิคแบบฟัซซีและนิวรอลเน็ตเวิร์ก", *วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย*, 2540



- [15] C. Wutiwiwatchai, V. Achariyakulporn, and C. Tanprasert, "Text-dependent Speaker Identification using LPC and DTW for Thai Language", *1999 IEEE 10<sup>th</sup> Region Conference (TENCON'99)*, Vol. 1, September 1999.
- [16] C. Wutiwiwatchai, S. Sae-tang, and C. Tanprasert, "Text-dependent Speaker Identification Using Neural Network on Distinctive Thai Tone Marks", *Proceedings of International Joint Conference on Neural Networks*, July 1999.
- [17] S. Kasuriya, S. Jitapunkul, V. Ahkuputra, E. Maneenoi, and P. Tasakoo, "Text-Dependent Speaker Identification for Thai Language using Hidden Markov Model", *1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 781-784, December 1999.



**วรินทร์ อัจฉริยะกุลพร** สำเร็จการศึกษาระดับปริญญาตรีสาขาวิทยาการคอมพิวเตอร์ เกียรตินิยมอันดับ 2 จากมหาวิทยาลัยเกษตรศาสตร์ ปี พ.ศ. 2538 และการศึกษาในระดับปริญญาโทสาขาวิทยาการคอมพิวเตอร์ จากมหาวิทยาลัยมหิดล ปี พ.ศ. 2541 ปัจจุบันได้ร่วมงานในหน่วยปฏิบัติการวิจัยและพัฒนาวิศวกรรมภาษาและซอฟต์แวร์ โดยรับผิดชอบในหน้าที่การพัฒนาโปรแกรมพัฒนาเว็บเพจ และฐานข้อมูล นอกจากนี้ยังมีส่วนร่วมรับผิดชอบในโครงการแก้ปัญหาคอมพิวเตอร์ปี ค.ศ. 2000 โครงการ Asean-India Digital Archive (AIDA) และโครงการระบบระบุผู้พูดภายในเนตเทค มีความสนใจในด้านการออกแบบระบบงาน การออกแบบโปรแกรมเชิงวัตถุ และงานทางด้านปัญญาประดิษฐ์



**ชัย วุฒิวิวัฒน์ชัย** เกิดเมื่อวันที่ 6 ตุลาคม 2516 สำเร็จการศึกษาวิศวกรรมศาสตรบัณฑิต เกียรตินิยมอันดับ 1 ภาควิชาวิศวกรรมไฟฟ้า จากมหาวิทยาลัยธรรมศาสตร์เมื่อปี พ.ศ. 2537 และวิศวกรรมศาสตรมหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า

สาขาวิทยาการประมวลผลสัญญาณเชิงเลข จากจุฬาลงกรณ์มหาวิทยาลัยเมื่อปี พ.ศ. 2540 หลังจากนั้นได้เข้าทำงานในศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ โดยเริ่มทำงานวิจัย

และพัฒนาในโครงการซอฟต์แวร์ระบุผู้พูดจากเสียง มีความเชี่ยวชาญในงานวิจัยทางด้านการ



ประมวลผลสัญญาณเชิงเลข โดยเฉพาะการประมวลผลเพื่อรู้จำเสียงพูดและรู้จำผู้พูด

**จุฬารัตน์ ดันประเสริฐ** สำเร็จการศึกษาวิทยาสตรบัณฑิต (คณิตศาสตร์) เกียรตินิยมอันดับ 1 จากจุฬาลงกรณ์มหาวิทยาลัย ในปี 2532 และจบปริญญาโทและเอกในสาขา Computer Science จาก University of Louisiana at Lafayette ณ ประเทศสหรัฐอเมริกา ในปี 2534 และ ปี 2537 ตามลำดับ หลังจากสำเร็จการศึกษาได้เข้าทำงาน ณ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ในหน่วยปฏิบัติการวิจัยและพัฒนาวิศวกรรมภาษาและซอฟต์แวร์ โดยเป็นหัวหน้าโครงการพัฒนาโปรแกรมแปลงภาพเอกสารตัวพิมพ์ภาษาไทยและอังกฤษให้เป็นแฟ้มข้อความ (Thai OCR) และโครงการพัฒนาโปรแกรมระบุผู้พูดด้วยเสียงภาษาไทย (Speaker Identification) นอกจากนี้ยังรับผิดชอบหน้าที่ในการประสานงานโครงการอีกหลายโครงการ อาทิเช่น โครงการพัฒนาซอฟต์แวร์ประมวลผลคำไทย โครงการสอบถามเลขหมายโทรศัพท์ผ่านทางอินเทอร์เน็ต เป็นต้น มีความสนใจในงานวิจัยที่เกี่ยวข้องกับเรื่องโครงข่ายประสาทเทียม การรู้จำรูปแบบ ระบบฐานข้อมูล ปัญญาประดิษฐ์ และการประมวลผลภาพลักษณ์

## แนะนำ Bio-Mirror ของฐานข้อมูลชีวภาพของ เครือข่ายนักวิจัยภาคพื้นเอเชีย-แปซิฟิก

สมนึก คีรีโต และ จิตินันต์ เขียนดวงจันทร์

ฝ่ายสารสนเทศงานวิจัย สถาบันวิจัยและพัฒนาแห่งมหาวิทยาลัยเกษตรศาสตร์

พันธุ์ปิติ เปี่ยมสง่า

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

อีเมล {sk,g4265081,pp}@ku.ac.th

**ABSTRACT** – Kasetsart University Research and Development Institute (KURDI) by cooperation with Asia-Pacific Advanced Network (APAN) establishes and services biological sequence databanks. We provide free of charge for researchers and general person at web site: <http://bio-mirror.ku.ac.th> and ftp site: <ftp://bio-mirror.ku.ac.th/biomirror>. The currently total databanks are about 28 Gigabyte in compressed format and include many subdata that separate 3 widely groups. 1) DNA sequence database include GENBANK, BLAST, DDBJ, EMBL and RICE; 2) Protein sequence database include PROSITE, BLOCKS, PFAM, PIR, SWISSPROT, and TREMBL and 3) Other sub database include ENZYME, REBASE, TAXONOMY, and UNIGENE.

Beside, there are many software engines used for analyze and manage biological sequence database in Bio-Mirror that is useful to researcher and general person. Bio-Mirror helps users in Thailand for high-speed access to increase performance in using network that connects other countries. Moreover Bio-Mirror is useful to biological technology researcher and information technology researcher for research and development in the future.

**KEY WORDS** – Biological Database, DNA and Protein Sequences

**บทคัดย่อ** – สถาบันวิจัยและพัฒนาแห่งมหาวิทยาลัยเกษตรศาสตร์ โดยความร่วมมือกับเครือข่ายนักวิจัยภาคพื้นเอเชีย-แปซิฟิก (Asia-Pacific Advanced Network: APAN) ได้ดำเนินการติดตั้งและให้บริการระบบข้อมูลทางชีวภาพให้กับนักวิจัยและผู้สนใจทั่วไปโดยไม่คิดค่าบริการใดๆ ในลักษณะของ WWW ที่ <http://bio-mirror.ku.ac.th> และ FTP ที่ <ftp://bio-mirror.ku.ac.th/biomirror/> ระบบดังกล่าวเก็บรวบรวมข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพ ไว้จำนวนมากกว่า 19 กิกะไบต์ (ขนาดของข้อมูลเมื่อถูกบีบอัด) โดยประกอบด้วยข้อมูลย่อยต่างๆ มากมาย โดยแบ่งกว้างๆ ได้สามกลุ่มคือ 1) ฐานข้อมูลลำดับนิวคลีโอไทด์และกรดอะมิโนประกอบด้วย GENBANK, BLAST, DDBJ, EMBL และฐานข้อมูลลำดับนิวคลีโอไทด์ของข้าว (RICE); 2) ฐานข้อมูลลำดับกรดอะมิโนและโปรตีนประกอบด้วย PROSITE, BLOCKS, PFAM, PIR, SWISSPROT, และ TREMBL และ 3) ฐานข้อมูลย่อยอื่นๆ ประกอบด้วย ENZYME, REBASE, TAXONOMY, และ UNIGENE

นอกจากนี้ยังมีโปรแกรมที่ช่วยในการวิเคราะห์และจัดการฐานข้อมูลในระบบข้อมูล Bio-Mirror ในลักษณะต่างๆ ซึ่งจะเป็นประโยชน์ต่อนักวิจัยและผู้สนใจทั่วไปด้วย ระบบข้อมูล Bio-Mirror นี้จะช่วยทำให้ผู้ใช้งานในประเทศไทยสามารถเข้าถึงข้อมูลได้สะดวกและรวดเร็วขึ้น ช่วยเพิ่มประสิทธิภาพในการใช้สายเครือข่ายข้ามประเทศให้มี ประสิทธิภาพสูงขึ้น และเพื่อให้ นักวิจัยทั้งทาง ด้านเทคโนโลยีชีวภาพ และนักวิจัยทางด้านเทคโนโลยีสารสนเทศได้ใช้ประโยชน์ระบบข้อมูลนี้ เพื่อการค้นคว้า วิจัยและพัฒนาให้เกิด ประโยชน์มากขึ้นต่อไป

**คำสำคัญ** – ระบบข้อมูลทางชีวภาพ, ลำดับเบสของดีเอ็นเอและโปรตีน

## 1. บทนำ

ในปัจจุบัน การจัดทําระบบฐานข้อมูลที่เป็นศูนย์กลางของข้อมูลและสารสนเทศเพื่อการวิจัยเฉพาะเรื่องหรือเฉพาะสาขานับเป็นโครงสร้างพื้นฐานที่สำคัญมากในการดำเนินงานวิจัย และถือได้ว่าเป็นข้อมูลพื้นฐานที่นักวิชาการและนักวิจัยต้องการเพื่อทำการศึกษาดังข้อมูลเบื้องต้นหรือความเป็นไปได้ของการที่จะทำการวิจัยต่อไปในอนาคต รวมทั้งเป็นแหล่งที่สามารถใช้ในการศึกษาถึงความก้าวหน้าของศาสตร์นั้น ซึ่งจะเห็นได้ว่าในประเทศไทยจำนวนฐานข้อมูลสารสนเทศมิให้บริการไม่เพียงพอและยังขาดข้อมูลเฉพาะเรื่องอีกมาก ดังนั้นจึงได้มีการศึกษาและคิดดัดแปลงระบบข้อมูลทางด้านชีววิทยาขึ้นมา โดยระบบข้อมูลทางด้านชีววิทยานี้ได้เก็บรวบรวมฐานข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพไว้ในเซิร์ฟเวอร์ ที่เรียกว่า Bio-Mirror เซิร์ฟเวอร์นี้ได้ติดตั้งเว็บไซต์สำหรับเข้าถึงข้อมูลได้อย่างรวดเร็วและมีข้อมูลที่เปลี่ยนแปลงใหม่อยู่ตลอดเวลา ซึ่งระบบดังกล่าวนี้ได้รับความร่วมมือในการพัฒนาจากเครือข่ายนักวิจัยภาคพื้นเอเชียและแปซิฟิก หรือ Asia-Pacific Advanced Network (APAN) [1] การพัฒนาข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพนี้ได้มาจากความร่วมมือระหว่างนักวิจัยและพัฒนาจากสหรัฐอเมริกา ญี่ปุ่น และยุโรป โดยมีการแลกเปลี่ยนข้อมูลระหว่างเว็บไซต์กันตลอดเวลา การแลกเปลี่ยนข้อมูลกันได้อย่างรวดเร็วนี้ มีองค์กรและกลุ่มความร่วมมือต่างๆ ที่มีส่วนช่วยในการติดต่อสื่อสารคือ Very High Speed Backbone Service (vBNS) [26], TransPAC (Trans-Pacific Network) [22] และ Asia-Pacific Advanced Network (APAN)

## 2. จุดมุ่งหมายของระบบข้อมูล Bio-Mirror

การคิดดัดแปลงระบบข้อมูล Bio-Mirror มีจุดมุ่งหมายสำคัญคือ การให้บริการข้อมูลแก่นักวิจัยและผู้สนใจทั่วไปทั้งภายในและภายนอกประเทศ โดยผ่านทางเครือข่ายอินเทอร์เน็ตได้อย่างสะดวกและรวดเร็ว และเพื่อให้ นักวิจัยและผู้สนใจทั่วไปสามารถเข้ามาศึกษาหรือนำข้อมูลไปใช้โดยไม่คิดค่าบริการใดๆ สถาบันวิจัยและพัฒนาแห่งมหาวิทยาลัยเกษตรศาสตร์ ได้ให้ความร่วมมือกับเครือข่ายนักวิจัยระหว่างประเทศ (APAN) และได้จัดตั้ง Mirror เว็บไซต์เพื่อให้บริการสำหรับผู้ใช้ใน ประเทศไทยและประเทศใกล้เคียงได้ที่เว็บไซต์ <http://bio-mirror.ku.ac.th/> และที่เอฟทีพีไซต์ <ftp://bio-mirror.ku.ac.th/biomirror/> เว็บไซต์ที่จัดตั้งขึ้นนี้จะช่วยลดการใช้ bandwidth หรือช่วยเพิ่มอรรถประโยชน์จากสายเครือข่ายข้ามประเทศให้มีประสิทธิภาพสูงขึ้นด้วย

## 3. ฐานข้อมูลในระบบข้อมูล Bio-Mirror

ฐานข้อมูลลำดับเบสของดีเอ็นเอ หรือบางครั้งเรียกว่าลำดับนิวคลีโอไทด์ และลำดับเบสของโปรตีนหรือที่เรียกว่าลำดับกรดอะมิโนนั้น ในปัจจุบันมีอยู่มากมายมหาศาล นักวิจัยและพัฒนาระหว่างประเทศได้ร่วมกัน จัดเก็บรวบรวมข้อมูลที่เกี่ยวข้องกับข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพไว้อย่างมีระบบ ซึ่งฐานข้อมูลหลักในโลกจะมีอยู่ 3 ฐานคือ Genbank [18] ในสหรัฐอเมริกา, EMBL [7] ในทวีปยุโรป และ DDBJ [5] ในประเทศญี่ปุ่น ซึ่งฐานข้อมูลทั้งสามนี้จะมีการแลกเปลี่ยนข้อมูลกันอยู่ตลอดเวลาทำให้มีข้อมูลที่เรียกว่าเกือบจะเหมือนกัน ฐานข้อมูลเหล่านี้จะมีทั้งฐานข้อมูลทั่วไปและฐานข้อมูลเฉพาะ เช่น ฐานข้อมูลของลำดับนิวคลีโอไทด์ ฐานข้อมูลของลำดับกรดอะมิโนและเว็บไซต์ที่เกี่ยวข้องกับโปรตีน ฐานข้อมูลของโครงสร้างตติยภูมิ เป็นต้น นอกจากนี้ยังได้จัดตั้ง Mirror เว็บไซต์ของระบบข้อมูลทางชีวภาพที่เรียกว่า Bio-Mirror ขึ้นเพื่อให้บริการแก่นักวิจัยและผู้สนใจทั่วไป ระบบดังกล่าวเก็บรวบรวมข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพไว้จำนวนมากกว่า 19 กิกะไบต์ ในลักษณะข้อมูลบีบอัด ถ้านำมาขยายแล้วจะมีข้อมูลมากถึงประมาณ 80 กิกะไบต์ โดยประกอบด้วยข้อมูลย่อยต่างๆ มากมาย แยกได้ตามหมวดหมู่ดังนี้

### 3.1 ฐานข้อมูลของลำดับนิวคลีโอไทด์

#### 3.1.1 DDBJ

DDBJ [5] เป็นฐานข้อมูลของลำดับนิวคลีโอไทด์ในประเทศญี่ปุ่นซึ่งสนับสนุนโดยรัฐบาลญี่ปุ่น ฐานข้อมูลนี้เริ่มให้บริการในปี ค.ศ. 1986 โดยความร่วมมือของกลุ่มผู้เชี่ยวชาญทางด้านอนุชีววิทยาของประเทศญี่ปุ่น และการสนับสนุนของกระทรวงศึกษา วิทยาศาสตร์ กีฬา และวัฒนธรรม ฐานข้อมูลนี้ได้รับการดูแลโดย DDBJ ร่วมกับองค์กรหลักอื่นๆที่มีฐานข้อมูลลำดับนิวคลีโอไทด์อยู่

#### 3.1.2 EMBL

EMBL [7] เป็นฐานข้อมูลของลำดับนิวคลีโอไทด์ในทวีปยุโรป ซึ่งแหล่งข้อมูลหลักสำหรับข้อมูลลำดับเบสของดีเอ็นเอ และอาร์เอ็นเอ (RNA) ได้มาจากนักวิจัยที่ค้นคว้าด้านนี้โดยตรง โครงการหาลำดับของจีโนม (Genome) และงานอื่นๆที่จัดลิขสิทธิ์แล้ว ฐานข้อมูล EMBL สร้างขึ้นโดยความร่วมมือระหว่างประเทศคือ Genbank ของสหรัฐอเมริกา และ DDBJ ในประเทศญี่ปุ่น แต่ละกลุ่มจะเก็บสะสมข้อมูลลำดับเบสบางส่วนไว้และเผยแพร่ต่อทั่วโลก อนุญาตให้นำมาใช้ และฐานข้อมูลมีการเปลี่ยนแปลงใหม่ทุกวันระหว่างกลุ่มความร่วมมือของฐานข้อมูลนี้

### 3.1.3 GENBANK

GENBANK [18] เป็นฐานข้อมูลของลำดับนิวคลีโอไทด์ในประเทศสหรัฐอเมริกาซึ่งสนับสนุนโดย National Institute of Health (NIH) [11] ฐานข้อมูลนี้ให้บริการตั้งแต่ปี ค.ศ.1992 และบริหารโดย National Center for Biotechnology Information (NCBI) โดยได้รับความร่วมมือจากองค์กรหลักอื่นๆที่มีฐานข้อมูลลำดับนิวคลีโอไทด์และลำดับกรดอะมิโนอยู่ Genbank ได้จดทะเบียนการค้ากับ National Institute of Health (NIH) แต่ฐานข้อมูลที่มีอยู่เปิดให้บริการโดยไม่คิดค่าบริการใดๆ ทั้งสิ้น

### 3.1.4 RICE

RICE [19] เป็นฐานข้อมูลของลำดับนิวคลีโอไทด์ของข้าว

## 3.2 ฐานข้อมูลของโปรตีนและลำดับกรดอะมิโน

### 3.2.1 BLOCKS

BLOCKS [3] เป็นระบบสำหรับจำแนกโปรตีนซึ่งได้มาจากลำดับเบสอนุรักษ์ (Conserved Regions) ในกลุ่มของโปรตีนที่อยู่ในฐานข้อมูล PROSITE ฐานข้อมูลบล็อกนี้เป็นกลุ่มของลำดับเบสที่เหมือนกัน โดยได้รับจากลำดับเบสหลายสายที่นำมาเปรียบเทียบความคล้ายกัน ซึ่งการนำลำดับเบสหลายสายมาเปรียบเทียบกันนี้จะช่วยลดสิ่งรบกวน และเพิ่มประสิทธิภาพในการหาความสัมพันธ์ของลำดับเบส ฐานข้อมูลนี้มีจุดมุ่งหมายเพื่ออธิบายโครงสร้างและหน้าที่การทำงานของลำดับเบส

ขั้นตอนในการสร้างฐานข้อมูลบล็อกจะมีอยู่สามขั้นตอน โดยขั้นตอนแรกจะเป็นการหาลำดับเบสอนุรักษ์ในกลุ่มของโปรตีนที่อยู่ในฐานข้อมูล PROSITE และในขั้นตอนที่สองจะนำบล็อกของโปรตีนที่ได้มาจัดเรียงใหม่ให้ดีขึ้นเพื่อที่จะสร้างกลุ่มของบล็อกของโปรตีน ส่วนในขั้นตอนที่สามบล็อกของโปรตีนจะถูกวัดขนาดเทียบกับฐานข้อมูล SWISS-PROT เพื่อที่จะอธิบายระดับของการจับคู่ที่มีโอกาสเกิดขึ้นอีกครั้งหนึ่ง ฐานข้อมูลบล็อกจะถูกสร้างขึ้นเป็นกลุ่มของบล็อกของโปรตีนอย่างสมบูรณ์

### 3.2.2 PFAM

PFAM [16] เป็นฐานข้อมูลการเปรียบเทียบความคล้ายกันของลำดับเบสหลายสาย และโมเดล Hidden Markov (HMMs) ซึ่งครอบคลุมโดเมนของโปรตีน หรือลำดับเบสอนุรักษ์ของโปรตีนมากมาย ฐานข้อมูลนี้ได้จากความร่วมมือกันระหว่างห้องปฏิบัติการวิจัยของ Erik Sonnhammer (Sanger Centre, Hinxton UK) และ Sean Eddy (Dept. of Genetics, Washington University, St. Louis USA)

### 3.2.3 PIR

PIR [17] เป็นแหล่งข้อมูลของโปรตีนและฐานข้อมูลลำดับกรดอะมิโนของนานาชาติ ซึ่งได้รับการดูแลรักษาโดย PIR ที่ National Biomedical Research Foundation (NBRF) ในสหรัฐอเมริกา Martinsried Institute for Protein Sequence (MIPS) ในเยอรมัน และ Japan International Protein Information Database (JIPID)

ในญี่ปุ่น ฐานข้อมูลนี้เป็นที่เก็บรวบรวมข้อมูลของโปรตีนที่เกิดขึ้นตามธรรมชาติซึ่งลำดับกรดอะมิโนของมันจะถูกกำหนดขึ้น ข้อมูลในฐานข้อมูลนี้ประกอบด้วยข้อมูลอ้างอิงถึงชื่อ ประเภท และแหล่งที่ได้รับโปรตีน และลักษณะทางชีวภาพของโปรตีน เช่น เครื่องหมายแสดงและขั้นตอนการทำงาน ขอบเขตและความเปลี่ยนแปลงภายในโปรตีน ลักษณะหน้าที่การทำงาน และข้อมูลของลำดับเบสขั้นแรก นอกจากนี้ข้อมูลจำนวนมากในฐานข้อมูลนี้ยังถูกเก็บไว้ในฐานข้อมูลอื่นๆ ด้วยเช่น ฐานข้อมูล SWISS-PROT และฐานข้อมูลอื่นๆ ซึ่งดูแลรักษาโดย NCBI และ EBI

### 3.2.4 PROSITE

PROSITE [14] เป็นแหล่งข้อมูลซึ่งเก็บโครงร่างของโปรตีน และเว็บไซต์ที่เกี่ยวข้องกับโปรตีน ฐานข้อมูลนี้ได้รับการบริหารและจัดการโดย Dr. Amos Bairoch ที่มหาวิทยาลัยของเมืองเจนีวา ประเทศสวิสเซอร์แลนด์ ซึ่งสามารถได้รับข้อมูลของ PROSITE จากซีดีรอมของ EBI โดยตรง หรือโดยเอฟทีพีจากเซิร์ฟเวอร์ของ NCBI EBI หรือ DDBJ หรือส่งผ่านเซิร์ฟเวอร์ ExPASy [9] บนเว็บไซต์

### 3.2.5 SWISSPROT

SWISSPROT [15] เป็นฐานข้อมูลของลำดับกรดอะมิโนซึ่งได้จากแหล่งข้อมูลหลายแหล่ง ประกอบด้วยกรดอะมิโนที่แปลมาจากกรดนิวคลีอิก กรดอะมิโนจากฐานข้อมูลลำดับกรดอะมิโนหลักของนานาชาติ ฐานข้อมูลนี้เริ่มต้นใน ปี ค.ศ. 1985 และในปี ค.ศ. 1987 ได้รับการดูแลรักษาโดยมหาวิทยาลัยของเมืองเจนีวา ประเทศสวิสเซอร์แลนด์ และห้องสมุดข้อมูล EBI

ข้อมูลใน SWISS-PROT มาจากแหล่งข้อมูล 3 แหล่งคือ การแปลลำดับนิวคลีโอไทด์จากฐานข้อมูลลำดับนิวคลีโอไทด์ของ EMBL จากกลุ่มของนักวิจัยโดยตรง และจากเอกสารทางวิทยาศาสตร์ ฐานข้อมูลโปรตีนอื่นๆมากมายเช่น PROSITE, BLOCKS และ PRINTS ได้รับข้อมูลมาจากฐานข้อมูล SWISS-PROT

### 3.2.6 TREMBL

TREMBL [23] เป็นฐานข้อมูลลำดับกรดอะมิโนซึ่งแปลมาจากฐานข้อมูลลำดับนิวคลีโอไทด์ของ EMBL โดยตรง ประกอบด้วยส่วนเสริม

ให้กับฐานข้อมูล SWISS-PROT เป็นคำอธิบายประกอบของ คอมพิวเตอร์ ซึ่งรูปแบบจะเหมือนกับรูปแบบของ SWISS-PROT รูปแบบนี้ได้มาจากการแปลลำดับเบสที่รู้ลำดับแล้วทั้งหมดหรือ Coding Sequences (CDS) ในฐานข้อมูลลำดับนิวคลีโอไทด์ของ EMBL ซึ่งฐานข้อมูลของ EMBL ต้องการลำดับเบสที่รู้ลำดับเรียบร้อยแล้วรวมทั้งฐานข้อมูล SWISS-PROT ด้วย

### 3.3 ฐานข้อมูลอื่นๆ

#### 3.3.1 BLAST

BLAST [2] เป็นฐานข้อมูลลำดับนิวคลีโอไทด์และกรดอะมิโนของ Genbank ซึ่งเก็บอยู่ในรูปของ FASTA Format เป็นฐานข้อมูลหลักที่ใช้ในโปรแกรมสำหรับวิเคราะห์และจัดการข้อมูลในระบบข้อมูล Bio-Mirror

#### 3.3.2 ENZYME

ENZYME [8] เป็นฐานข้อมูลที่เกี่ยวข้องกับการตั้งชื่อวิทยาศาสตร์ของ เอนไซม์ ซึ่งได้รับการสนับสนุนและแนะนำเบื้องต้นจากคณะกรรมการการตั้งชื่อทางชีวเคมีและอนุชีววิทยาแห่งสหประชาชาติ หรือ Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) และอธิบายรายละเอียดลักษณะของ เอนไซม์แต่ละชนิดพร้อมรหัส EC (Enzyme Commission)

#### 3.3.3 REBASE

REBASE [24] เป็นฐานข้อมูลของเอนไซม์ตัดจำเพาะซึ่งเก็บรายละเอียดของข้อมูลเกี่ยวกับชนิดเอนไซม์ตัดจำเพาะ ชนิดของจุลินทรีย์ที่ผลิตเอนไซม์ ลำดับเบสที่เอนไซม์เลือกตัดจำเพาะ ตำแหน่งที่เอนไซม์ตัดจำเพาะ เอนไซม์ที่หาได้ในเชิงพาณิชย์ และเอกสารอ้างอิงทั้งที่ตีพิมพ์และยังไม่ได้ตีพิมพ์

#### 3.3.4 SRS-DATABANKS

SRS-DATABANKS [20] เป็นฐานข้อมูลที่ได้รับมาจาก EBI [6] เป็นฐานข้อมูลที่เก็บรวบรวมข้อมูลของระบบสืบค้นข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพ ซึ่งระบบสืบค้น SRS เขียนขึ้น เพื่อเลือกข้อมูลย่อยจากฐานข้อมูลลำดับเบสทั้งหมด

#### 3.3.5 TAXONOMY

TAXONOMY [21] เป็นฐานข้อมูลลำดับชั้นของสิ่งมีชีวิตซึ่งเก็บอยู่ในรูปแบบของโครงสร้างต้นไม้ โดยจะเก็บชื่อของสิ่งมีชีวิตทั้งหมดที่เป็นตัวแทนในฐานข้อมูลลำดับเบส และลำดับนิวคลีโอไทด์หรือกรดอะมิโนที่น้อยที่สุดหนึ่งสาย ฐานข้อมูลนี้ส่วนใหญ่ได้รับมาจาก NCBI และ EBI

#### 3.3.6 UNIGENE

UNIGENE [25] เป็นระบบการทดลองเพื่อแบ่งลำดับเบสของ Genbank ให้เป็นเซตที่ไม่ซ้ำซากหรือยาวเกินไปของกลุ่มยีนที่จัดตำแหน่งแล้วอย่างอัตโนมัติ แต่ละกลุ่มของ UniGene เก็บลำดับเบสซึ่งเป็นตัวแทนยีนที่มีลักษณะเฉพาะ และข้อมูลที่เกี่ยวข้องเช่น ชนิดของเนื้อเยื่อในยีนที่สามารถระบุตำแหน่งได้ ฐานข้อมูลนี้ได้รับจาก National Center for Biotechnology Information (NCBI) หรือ Genbank ฐานข้อมูลนี้จะเก็บแยกเป็นไฟล์ด้วยคำนำหน้าชื่อที่ต่างกันตามชนิดของยีน ไฟล์ในฐานข้อมูลที่น่าหน้าด้วย Hs ย่อมาจาก Homo sapiens ไฟล์ที่น่าหน้าด้วย Mm ย่อมาจาก Mus musculus ไฟล์ที่น่าหน้าด้วย Rn ย่อมาจาก Rattus norvegicus และไฟล์ที่น่าหน้าด้วย Dr ย่อมาจาก Danio rario

ตารางที่ 1. เซิร์ฟเวอร์ของ Bio-Mirror ที่มีอยู่ในปัจจุบัน

| ประเทศ        | เว็บเซิร์ฟเวอร์และเอฟทีพีเซิร์ฟเวอร์  |
|---------------|---|
| เกาหลีใต้     | http://bio-mirror.kr.apan.net<br>ftp://bio-mirror.kr.apan.net/pub/biomirror |
| จีน           | http://bio-mirror.cn.apan.net<br>ftp://bio-mirror.cn.apan.net               |
| ญี่ปุ่น       | http://bio-mirror.jp.apan.net<br>ftp://bio-mirror.jp.apan.net/pub/biomirror |
| ไทย           | http://bio-mirror.ku.ac.th<br>ftp://bio-mirror.ku.ac.th/biomirror           |
| สหรัฐ-อเมริกา | http://bio-mirror.us.apan.net<br>ftp://bio-mirror.us.apan.net/biomirror     |
| สิงคโปร์      | http://bio-mirror.sg.apan.net<br>ftp://bio-mirror.sg.apan.net/biomirrors    |
| ออสเตรเลีย    | http://bio-mirror.au.apan.net<br>ftp://bio-mirror.au.apan.net/biomirrors    |

## 4. ระบบวิเคราะห์และจัดการฐานข้อมูลชีวภาพ

ในทศวรรษที่ผ่านมาจนถึงปัจจุบันนับได้ว่าเป็นยุคของเทคโนโลยีทางวิทยาศาสตร์การแพทย์ ความรู้จากการค้นพบสิ่งใหม่ๆ ก่อให้เกิดประโยชน์นานัปการ การนำคอมพิวเตอร์เข้ามาช่วยในงานทางด้านวิทยาศาสตร์การแพทย์นับว่ามีบทบาทสำคัญมาก ช่วยทำให้เกิดความก้าวหน้าและการค้นพบสิ่งใหม่ๆ ขึ้นมา ในการศึกษาภาคการดำเนินงานของมนุษยศาสตร์ซึ่งมีปริมาณถึง 3 พันล้านนิวคลีโอไทด์ หรือประมาณ 1 แสนยีนซึ่งเป็นเสมือนหนึ่งพิมพ์เขียวของมนุษย์นั้น หากว่านักวิทยา

ศาสตร์มีความสามารถในการวิเคราะห์ชั้นได้วินาทีละ 1 นิวคลีโอไทด์ แม้จะทำงานวันละ 24 ชั่วโมงทุกวัน ก็ยังต้องใช้เวลานานนับ 100 ปีที่จะศึกษาได้ทั้งหมด ดังนั้นจึงได้มีการพัฒนาเครื่องมือและเทคนิคต่างๆ เข้ามาช่วยเพื่อวิเคราะห์และประมวลผลได้อย่างสะดวกและรวดเร็ว โปรแกรมประยุกต์มากมายได้นำมาใช้กับฐานข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพสามารถแยกแยะเป็นกลุ่มดังนี้

#### 1. โปรแกรมแปลลำดับเบส (Translation)

- โปรแกรมแปลลำดับเบสของดีเอ็นเอเป็นโปรตีน (DNA Translation)
- โปรแกรมแปลลำดับเบสของโปรตีนกลับเป็นดีเอ็นเอ (Back Translation)

#### 2. โปรแกรมค้นหาลำดับเบสของดีเอ็นเอและโปรตีนโดยเปรียบเทียบความคล้ายกันของลำดับเบส (Sequence Homology or Similarity Searching)

#### 3. โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีนที่กำลังศึกษา (Sequence Alignment)

- โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีนสองเส้น (Pairwise sequence alignment)
- โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีนมากกว่าสองเส้น (Multiple sequence alignment)

#### 4. โปรแกรมจำแนกชนิดและลักษณะของโปรตีน (Protein Identification and Characterization)

#### 5. โปรแกรมวิเคราะห์โครงสร้างปฐมภูมิ (Primary Structure Analysis)

- โปรแกรมคาดคะเนตำแหน่งของสายโปรตีนที่เกิดการขดเป็นวง
- โปรแกรมคาดคะเนตำแหน่งที่เอนไซม์ตัดจำเพาะ

#### 5. โปรแกรมการหาตำแหน่งของ Transmembrane (Transmembrane regions detection)

ตารางที่ 2. โปรแกรมประยุกต์ที่ใช้กับฐานข้อมูลลำดับเบสของดีเอ็นเอและโปรตีน

| 1. โปรแกรมแปลลำดับเบส                       |   |
|---|---|
| โปรแกรมแปลลำดับเบสของดีเอ็นเอเป็นโปรตีน     |   |
| DNA Sequence Translation                    | <a href="http://alces.med.umn.edu/cuse.html">http://alces.med.umn.edu/cuse.html</a> |
| Protein Machine                             | <a href="http://www2.ebi.ac.uk/translate/">http://www2.ebi.ac.uk/translate/</a>     |
| โปรแกรมแปลลำดับเบสของโปรตีนกลับเป็นดีเอ็นเอ |   |

| Back Translation  | <a href="http://www.entelechon.com/english/backtranslation.html">http://www.entelechon.com/english/backtranslation.html</a>                           |
|---|---|
| 2. โปรแกรมค้นหาลำดับเบสของดีเอ็นเอและโปรตีนโดยเปรียบเทียบความคล้ายของลำดับเบส |   |
| FASTA   | <a href="http://crick.genes.nig.ac.jp/homology/fasta-e.html">http://crick.genes.nig.ac.jp/homology/fasta-e.html</a>                                   |
| BLAST   | <a href="http://www.ncbi.nlm.nih.gov/blast/blast.cgi">http://www.ncbi.nlm.nih.gov/blast/blast.cgi</a>   |
| SSEARCH   | <a href="http://watson.genes.nig.ac.jp/homology/ssearch-e.html">http://watson.genes.nig.ac.jp/homology/ssearch-e.html</a>                             |
| 3. โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีนที่กำลังศึกษา    |   |
| โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีน 2 เส้น             |   |
| ALIGN   | <a href="http://vega.igh.cnvs.fr/bin/align-guess.cgi">http://vega.igh.cnvs.fr/bin/align-guess.cgi</a>   |
| LALIGN  | <a href="http://vega.igh.cnvs.fr/bin/lalign-guess.cgi">http://vega.igh.cnvs.fr/bin/lalign-guess.cgi</a>   |
| LFasta  | <a href="http://www2.igh.cnrs.fr/fasta/lfasta-query.html">http://www2.igh.cnrs.fr/fasta/lfasta-query.html</a>   |
| โปรแกรมเปรียบเทียบความคล้ายกันระหว่างสายดีเอ็นเอหรือโปรตีนมากกว่า 2 เส้น      |   |
| CLUSTALW  | <a href="http://dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html">http://dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html</a> |
| MAP   | <a href="http://dot.imgen.bcm.tmc.edu:9331/seq-seqrch/protein-search.html">http://dot.imgen.bcm.tmc.edu:9331/seq-seqrch/protein-search.html</a>       |
| MSA   | <a href="http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html">http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html</a>       |
| MultAlign   | <a href="http://www.toulouse.inra.fr/multalin.html">http://www.toulouse.inra.fr/multalin.html</a>   |
| Molsoft   | <a href="http://molsoft.com/serv/walign.htm">http://molsoft.com/serv/walign.htm</a>   |
| 4. โปรแกรมจำแนกชนิดและลักษณะของโปรตีน   |   |
| ACompIdent  | <a href="http://expasy.nhri.org.tw/tools/aacomp/">http://expasy.nhri.org.tw/tools/aacomp/</a>   |
| MultiIdent  | <a href="http://expasy.nhri.org.tw/tools/multiident/">http://expasy.nhri.org.tw/tools/multiident/</a>   |
| PeptideMiss   | <a href="http://expasy.nhri.org.tw/tools/peptide-mass.html">http://expasy.nhri.org.tw/tools/peptide-mass.html</a>                                     |
| 5. โปรแกรมวิเคราะห์โครงสร้างปฐมภูมิ   |   |
| โปรแกรมคาดคะเนตำแหน่งของสายโปรตีนที่เกิดการขดเป็นวง                           |   |
| Coils   | <a href="http://www.ch.embnet.org">http://www.ch.embnet.org</a>   |
| PairCoils   | <a href="http://nightingale.lcs.mit.edu">http://nightingale.lcs.mit.edu</a>   |
| MultiCoils  | <a href="http://nightingale.lcs.mit.edu">http://nightingale.lcs.mit.edu</a>   |



| โปรแกรมคาดคะเนตำแหน่งที่เอนไซม์ตัดจำเพาะ                                  |   |
|---|---|
| DGPI  | <a href="http://www.bigfoot.com">http://www.bigfoot.com</a>                             |
| NetPicoRNA  | <a href="http://www.cbs.dtu.dk">http://www.cbs.dtu.dk</a>                               |
| 6. โปรแกรมการหาตำแหน่งของ Transmembrane (Transmembrane regions detection) |   |
| DAS   | <a href="http://www.biokemi.su.se">http://www.biokemi.su.se</a>                         |
| HMMTOP  | <a href="http://www.enzim.hu">http://www.enzim.hu</a>                                   |
| SOSUI   | <a href="http://azusa.proteome.bio.tuat.ac.jp">http://azusa.proteome.bio.tuat.ac.jp</a> |
| Tmpred  | <a href="http://www.ch.embnet.org">http://www.ch.embnet.org</a>                         |

## 5. ตัวอย่างการประยุกต์ Bio-Mirror

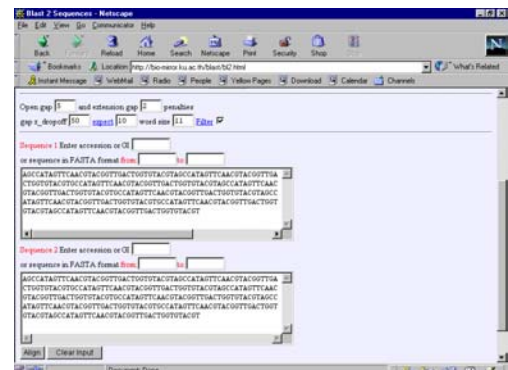
ในทศวรรษที่ผ่านมาจนถึงปัจจุบันนับได้ว่าเทคโนโลยีทางชีวภาพได้ก้าวหน้าไปอย่างรวดเร็ว การค้นพบสิ่งใหม่ๆ ก่อให้เกิดประโยชน์นานัปการ เช่น การวินิจฉัยโรคโดยอาศัยการตรวจความผิดปกติทางโครโมโซม การพัฒนาผลิตภัณฑ์ยาและวัคซีนรักษาโรค ความเข้าใจกลไกของโรคภัยไข้เจ็บ ฯลฯ หนึ่งในบรรดาผลสำเร็จเหล่านั้นที่ได้รับความสนใจจากสังคมโลกคือ เทคโนโลยีเกี่ยวกับดีเอ็นเอที่สามารถประยุกต์ใช้กับงานต่างๆ ได้มากมาย เช่น

- งานทางด้านการพิสูจน์หลักฐาน
  - การพิสูจน์ความเป็นพ่อแม่ ลูก
  - การพิสูจน์หลักฐานในคดีฆาตกรรม
  - การพิสูจน์หลักฐานในโบราณคดี
- งานทางด้านเกษตรกรรม
  - การปรับปรุงพันธุ์พืช พันธุ์สัตว์
  - การตรวจความผิดปกติในพืช และสัตว์
- งานทางด้านวิทยาศาสตร์การแพทย์
  - การตรวจสายพันธุ์ของไวรัส
  - การวินิจฉัยโรคโดยอาศัยการตรวจความผิดปกติของลำดับเบสของดีเอ็นเอ

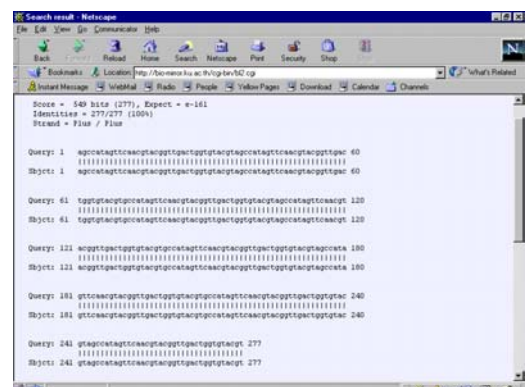
เพื่อให้เห็นถึงการนำข้อมูลของ Bio-Mirror มาประยุกต์ได้อย่างชัดเจน เราจะแสดงตัวอย่างการประยุกต์ดังรูปที่ 1 และรูปที่ 2 ซึ่งเป็นรูปการเปรียบเทียบลำดับเบสของดีเอ็นเอของไวรัสสองสาย โดยสมมติให้ลำดับเบสของดีเอ็นเอสายหนึ่งได้มาจากเชื้อไวรัสที่ได้มาจากคนที่มีความผิดปกติของโรค คล้ายกับคนเป็นโรคไวรัสตับอักเสบ B คือ Sequence 1 ในรูปที่ 1 และให้ลำดับเบสอีกสายหนึ่งคือลำดับเบสของดีเอ็นเอของไวรัสตับอักเสบ B ที่มีอยู่ในฐานข้อมูล Bio-Mirror คือ

Sequence 2 ในรูปที่ 1 เมื่อกดปุ่ม Align เพื่อเปรียบเทียบความคล้ายกันของลำดับเบสทั้ง 2 สายจะได้ผลลัพธ์ดังรูปที่ 2

ในการวินิจฉัยโรคโดยอาศัยการตรวจลำดับเบสของดีเอ็นเอ เราจะนำลำดับเบสสองสายมาเปรียบเทียบความคล้ายกัน หากลำดับเบสทั้งสองสายมีความคล้ายกัน เราก็สามารถวินิจฉัยได้ว่าคนๆนั้นติดเชื้อไวรัสตับอักเสบ B จริง ซึ่งมีความถูกต้องเกือบร้อยเปอร์เซ็นต์ในการตรวจวินิจฉัยโรคด้วยวิธีนี้



รูปที่ 1 การเปรียบเทียบความคล้ายกันระหว่างลำดับเบสของเชื้อไวรัส 2 สายเพื่อตรวจวินิจฉัยโรค



รูปที่ 2 ผลลัพธ์ของการเปรียบเทียบความคล้ายกันระหว่างลำดับเบสของเชื้อไวรัสสองสาย

จากตัวอย่างจะเห็นว่ากรณีนำข้อมูลลำดับเบสในระบบข้อมูล Bio-Mirror มาประยุกต์ใช้ในงานทางด้านวิทยาศาสตร์การแพทย์นั้นนับว่ามีประโยชน์มาก ยังมีงานวิจัยที่สามารถนำข้อมูลเหล่านี้ไปใช้ประโยชน์ได้อีกมากมาย งานทางด้านเกษตรกรรม การปรับปรุงพันธุ์พืช พันธุ์สัตว์ก็เป็นอีกงานหนึ่งที่ในปัจจุบันมีการศึกษาและวิจัยกันอยู่มาก ข้าวซึ่งเป็นพืชทางเศรษฐกิจที่สำคัญของประเทศนั้น ในปัจจุบันก็ได้มีองค์กรต่างๆ มากมายทั่วโลกได้ร่วมมือกันหาลำดับทางพันธุกรรมของข้าว เพื่อประโยชน์ในการปรับปรุงพันธุ์ สามารถต้านทานศัตรูพืช ทนต่อสภาพแวดล้อม และแพร่พันธุ์ได้มากขึ้น [13] เป็นต้น นอกจากนี้การทำโคลนนิ่งวัวที่สำเร็จไปไม่นานมานี้ ก็ได้มา

จากการวิเคราะห์และศึกษาข้อมูลลำดับเบสของดีเอ็นเอของมันนั่นเอง ซึ่งนับได้ว่าข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพในระบบข้อมูล Bio-Mirror นั้นมีประโยชน์สูงมาก หากมีผู้สนใจและนำข้อมูลเหล่านี้ไปประยุกต์ใช้หรือศึกษาเพื่อนำมาทำการวิจัยต่อไปในอนาคต

## 6. งานวิจัยที่สามารถทำต่อไปในอนาคตกับระบบข้อมูล

### Bio-Mirror

งานวิจัยที่สามารถทำต่อไปกับฐานข้อมูลลำดับเบสของดีเอ็นเอและโปรตีนทางชีวภาพนั้น มีมากมายหลายด้าน แต่ที่จำเป็นมากที่สุดคือ งานทางด้านชีววิทยาซึ่งใช้ทั้งในงานวิจัยทั่วไปและการเรียนระดับสูง งานวิจัยที่จำเป็นอีกงานหนึ่งก็คือ งานทางด้านวิทยาการคอมพิวเตอร์ เพราะฐานข้อมูลนี้มีขนาดใหญ่และเป็นข้อมูลเฉพาะด้าน ในการวิเคราะห์และจัดการกับฐานข้อมูลนี้เพื่อให้ได้ผลลัพธ์ที่ถูกต้องและรวดเร็วจำเป็นต้องใช้โปรแกรมคอมพิวเตอร์เข้ามาช่วย ซึ่งในปัจจุบันมีโปรแกรมคอมพิวเตอร์สำหรับวิเคราะห์และจัดการกับฐานข้อมูลนี้มากมายดังที่กล่าวมาแล้ว แต่ขั้นตอนวิธีและเทคนิคต่างๆ ที่ใช้ในการวิเคราะห์และจัดการข้อมูลให้ได้อย่างรวดเร็ว และมีความถูกต้องแม่นยำตามความต้องการของผู้ใช้นั้นยังไม่มีดีเพียงพอ

ดังนั้นการออกแบบขั้นตอนวิธีและหาเทคนิคใหม่ๆ มาใช้ในการวิเคราะห์และจัดการข้อมูลจึงเป็นสิ่งจำเป็น นอกจากนั้นฐานข้อมูลที่เพิ่มขึ้นอยู่ทุกวันซึ่งมีทั้งข้อมูลใหม่ที่เป็นข้อมูลชนิดเดิมและชนิดใหม่รวมอยู่ ยังต้องการโปรแกรมใหม่ที่จะรองรับสำหรับวิเคราะห์และจัดการกับข้อมูลนั้นตามความต้องการของผู้ใช้งานอีกด้วย

### เอกสารอ้างอิง

- [1] Asia-Pacific Advanced Network (APAN), Available URL: <http://www.apan.net/>
- [2] BLAST, Available URL: <http://www.ncbi.nlm.nih.gov/BLAST/>
- [3] Blocks WWW Server, Available URL: <http://blocks.fhrc.org/>
- [4] Dan Gusfield, "Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology", Cambridge University Press, 1997.
- [5] DDBJ Dna Data Bank of Japan, Available URL: <http://ddbj.nig.ac.jp/>
- [6] EBI, Available URL: <http://www.ebi.ac.uk/>
- [7] EMBL Outstation European Bioinformatics Institute, Available URL: <http://www.ebi.ac.uk/embl/>
- [8] Enzyme Nomenclature Database, Available URL: <http://www.expasy.ch/enzyme/>
- [9] Expasy Database, Available URL: <http://www.expasy.ch/>
- [10] FASTA, Available URL: <http://www.ncbi.nlm.nih.gov/BLAST/fasta.html>
- [11] National Institute of Health (NIH), Available URL: <http://www.nih.gov>
- [12] Pieve Baldi and Soren Brunak, "Bioinformatics: The Machine Learning Approach", The MIT Press Cambridge, 1999.
- [13] Prapa Sripichit, Pradit Pongtongkam and Apichart Vanavichit, Breeding of rice (oryza sativa.) variety KDML 105 for photoperiod intensivity through anther culture of F1 Hybrids, Kasetsart Journal, 1994, 28:499-511.
- [14] Prosite Database of protein families and domains, Available URL: <http://www.expasy.ch/prosite/>
- [15] SWISS-PROT annotated protein sequence database, Available URL: [http://expasy.nhri.org.tw/sprot/sprot\\_details.html](http://expasy.nhri.org.tw/sprot/sprot_details.html)
- [16] The PFAM Database of protein domains and HMMS, Available URL: <http://pfam.wustl.edu/>
- [17] Protein Information Resource, Available URL: <http://www-nbrf.georgetown.edu/pir/>
- [18] Genbank, Available URL: <http://www.ncbi.nlm.nih.gov/Database/>
- [19] Rice Genome Project, Available URL: <http://www.dna.affrc.go.jp/>
- [20] SRS, Available URL: <http://srs.ebi.ac.uk/>
- [21] Taxonomy, Available URL: <http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html>
- [22] Trans-Pacific Network (TransPAC), Available URL: <http://www.transpac.org/>
- [23] TREMBL, Available URL: <http://srs.ebi.ac.uk/>
- [24] The Restriction Enzyme Database, Available URL: <http://www.neb.com/rebase/>
- [25] UniGene Resources, Available URL: <http://www.ncbi.nlm.nih.gov/UniGene/>
- [26] Very High Speed Backbone Service (vBNS), Available URL: <http://www.vbns.net/>

## Somnuk Keretho

Assistant Professor



Somnuk Keretho received bachelor degree and master degree in electrical engineering from Kasetsart University in 1970 and 1976 respectively and received master degree in computer application and doctoral degree in computer science from Asian Institute of Technology in 1995 and 1982 respectively. He has research experiences in the fields of Artificial Intelligence, Neural Computation, Object-Oriented and Java Technology, Design and Analysis of Algorithms and Management Information Systems. At the present, he is an assistant professor at the Department of Computer Engineering and a deputy director of Kasetsart University and Research Development Institute (KURDI). His current research interests are Management Information Systems and Bioinformatics.

## Punpiti Piamsa-nga

Lecturer

Punpiti Piamsa-nga received his B.Eng. and M.Eng. (Electrical Engineering) from Kasetsart University (KU,) Thailand in 1988 and 1991, respectively, and received his D.Sc. (Computer Engineering) from George Washington University in 1999. He has research experiences in the fields of Natural Language Processing, Parallel Computing, and Multimedia. He has been a faculty at the Department of Computer Engineering, KU since 1991. His current research interests are Bioinformatics and Multimedia.

## New VLSI circuits for fast, low-cost binary adders

Vitit Kantabutra<sup>(1)</sup>, Pasquale Corsonello<sup>(2)</sup>, Stefania Perri<sup>(3)</sup>

(1) College of Engineering

Idaho State University, Pocatello, Idaho 83209, USA

e.mail: [vkantabu@computer.org](mailto:vkantabu@computer.org)

(2) Department of Electronics Engineering and Applied Mathematics

University of Reggio Calabria, Loc. Vito de Feo,

89060 Reggio Calabria, ITALY

e.mail: [pascor@deis.unical.it](mailto:pascor@deis.unical.it)

(3) Department of Electronics, Computer Science and Systems

University of Calabria, Arcavacata di Rende

87036 - Rende (CS), ITALY

e.mail: [perri@deis.unical.it](mailto:perri@deis.unical.it)

**Abstract** - Adders are of fundamental importance in a wide variety of digital systems. Many fast adders exist, but adding fast using low area and power is still challenging. This paper presents a new bit block structure that computes propagate signals called “carry strength” in a ripple fashion. Several new adders based on the new bit block structure are proposed. Comparison with well-known conventional adders demonstrates that the usage of carry-strength signals allows high-speed adders to be realised at significantly lower cost and consuming lower power than previously possible.

**Key Words** – Adder, carry-skip, VLSI

**บทคัดย่อ:** วงจรบวกเลข เป็นวงจรที่มีความสำคัญเป็นอย่างยิ่งในอุปกรณ์ดิจิทัลหลายชนิด รวมทั้งเครื่องคอมพิวเตอร์ โทรศัพท์มือถือ และอุปกรณ์สื่อสารอื่นๆ แม้ว่าจะมีวงจรบวกเลข ที่เร็วมากอยู่แล้วหลายชนิดในโลกนี้ ก็ยังมีปัญหาในเรื่องการบวกที่รวดเร็ว และใช้พลังงานต่ำ คือกินไฟฟ้าน้อยและไม่ร้อน บทความนี้ กล่าวถึงวงจรบวกเลขใหม่หลายชนิด ซึ่งบวกได้เร็ว ในขณะที่ใช้พลังงานต่ำ และใช้เนื้อที่บนชิปน้อยอีกด้วย วงจรใหม่นี้ ใช้สัญญาณที่เรียกว่า ความแข็งแรงของตัวทด

**คำสำคัญ:** วงจรบวกเลข อุปกรณ์โทรคมนาคม วงจรคอมพิวเตอร์

### 1. Introduction

The importance of a fast, low-cost binary adder in a digital system is difficult to overestimate. Not only are adders used in every arithmetic operation, they are also needed for computing the physical address in virtually every memory fetch operation in most modern CPUs. Adders are also used in many other digital systems including telecommunications systems in places where a full-fledged CPU would be superfluous. Since delay is one of the most important properties of any addition circuit the design of high-speed adders is one of the main goals of the digital system designers. Nevertheless, in the last few years many

researchers have redirected their efforts to the realization of low-power adders that also achieve high speed. Due to this many styles of adders exist in literature, but it is quite difficult identifying among them an adder having a satisfactory trade-off between high-speed and low-power. This is true either for traditional adders, such as ripple-carry, carry-skip, carry-select and carry look-ahead, or for the more technology-specific addition circuits proposed in the last few years [5, 6, 7, 8, 9].

The goal of the research presented in this paper is to consider adders that can be implemented in a wide variety of technologies. Among such adders *ripple adders* are the smallest but also the slowest. More recently, *carry-skip*

Dr. Vitit Kantabutra was supported in part by an Idaho State University Faculty Research Committee grant.

adders [1, 2, 3] are gaining popularity due to their high speed and relatively small size. Normally, in an  $N$ -bit carry-skip adder divided into a proper number of  $M$ -bit blocks [1, 4], a long-range carry signal starts at a generic block  $B_i$ , rippling through some bits in that block, then skips some blocks, and ends in a block  $B_j$ . If the carry does not end at the LSB of  $B_j$  then rippling occurs in that block and an additional delay is needed to compute the valid sum bits. Carry-look-ahead and carry-select adders [1] are very fast but far larger and consume much more power than ripple or carry-skip adders. Two of the fastest known addition circuits are the Lynch-Swartzlander's [5] and Kantabutra's [6] hybrid carry-look-ahead adders. They are based on the usage of a carry tree that produces carries into appropriate bit positions without back propagation. In order to obtain the valid sum bits as soon as possible, in both Lynch-Swartzlander's and Kantabutra's adders the sum bits are computed by means of carry-select blocks, which are able to perform their operations in parallel with the carry-tree. The adders proposed in [7] and [8] are also examples of very high-speed circuits, but they are not cut out for low-power systems. On the contrary the addition circuits proposed in [9] could satisfy low-power requirement, but they are very slow.

This paper presents two new families of adders which allow a reasonable time power trade-off to be obtained. Both the proposed families are based on a new bit block structure that computes propagate signals called "carry-strength" in a ripple fashion. The first family of adders is a family of new carry-skip adders that are significantly faster than traditional carry-skip adders while not much larger. The second family of adders is a family of hybrid lookahead adders similar to those presented in [5, 6] but significantly smaller and still comparable in speed.

In our new type of carry-skip adder, the new block structure eliminates the delay due to the rippling at the end of the life of a long-range carry signal. The main idea is, that for each bit position  $k$  in a block  $B_j$  we compute whether the carry-in to position  $k$  comes from the carry-in to block  $B_j$ , or whether this carry is internally generated in block  $B_j$ . To this purpose we will use a new type of bit block, in which we will compute propagate signals that start at the LSB of the block and end at every bit position. We find it helpful to call the complements of these "carry-strength" signals, because they indicate for each bit position whether the carry-in to that position originates within the same bit block.

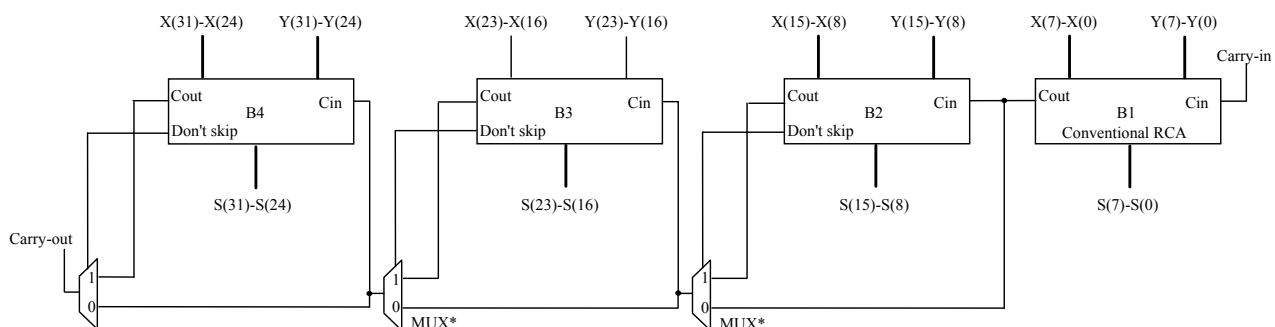


Figure 1.  $N$ -bit one-level carry-skip adder.

As mentioned above we will also show that carry-strength signals can also be successfully used in hybrid carry-look-ahead adders. The same principle described above for the carry-skip addition mechanism is applied to bit blocks to replace the larger blocks designed for Lynch-Swartzlander's and Kantabutra's adders. These new bit blocks allow us to avoid carry-select stages, saving significant area and power with little speed loss.

In Section 2, the meaning of the carry-strength signals and their computation are detailed. Then the new bit blocks for carry-skip and hybrid carry-look-ahead adders, as well as complete adders of these types, are presented in Sections 3 and 4, respectively. Section 5 discusses an optimisation procedure for bit block sizes and compares an optimised carry-skip adder of our type with a related design called the *Carry-Increment Adder (CIA)* (arrived at independently using a very different approach) by Zimmermann and Kaeslin [10], extending the work of Tyagi [11]. In these papers the authors found a way to reduce the redundancy in carry-select

adders, and came up with adders that are minimally slower than regular carry-select adders, requiring significantly less space. Interestingly enough, our blocks do have similarity to theirs, even though the two kinds of blocks were designed from very different points of view. Our adders are faster and require less silicon area.

Section 6 contains all the simulation results. The new bit blocks have been used in a 32-bit carry-skip adder and a 32-bit hybrid carry-look-ahead adder realised in AMS 0.6 $\mu$ m CMOS standard cells. In order to compare the new addition

circuits to existing ones, several conventional adders have been simulated. The new carry-skip adder is 57% of the layout area and consuming only 58% of the power. Surprisingly the new hybrid carry-look-ahead adder shows a slight speed advantage with respect to the Lynch-Swartzlander's scheme (also realised using the AMS 0.6 $\mu$ m

been also realised using the same technology. The new carry-skip adder has shown a speed of only 5% lower than a traditional carry-look-ahead adder, taking only 59% of the layout area and consuming only 58% of the power. Surprisingly the new hybrid carry-look-ahead adder *shows a slight speed advantage* with respect to the Lynch-Swartzlander's scheme (also realised using the AMS 0.6 $\mu$ m CMOS standard cells library), while taking only 76% of the layout area and consuming only 67% of the power.

## 2. Carry-skip addition and carry strength signals.

The main observation that led to the carry-skip adder is that, any bit position where the two operand bits differ will propagate its carry in. That is, if  $x_i$  and  $y_i$  are the two operand bits,  $c_i$  the carry in and  $c_{i+1}$  the carry out, then  $x_i \neq y_i$  implies  $c_{i+1} = c_i$ . For the sake of simplicity, let us assume that as shown in Figure 1 an  $N$ -bit one-level carry-skip adder is divided into  $N/M$  equal-length blocks ( $B_1, \dots, B_{N/M}$ ), each of which contains  $M$  bits [1,4]. (Our ideas are used even when the block sizes are unequal in order to optimise speed, as shown in Section V). Moreover, let's  $X$  and  $Y$  be the two  $N$ -bit operands of the adder. Any block in which all the positions  $i$  have unequal operands ( $x_i \neq y_i$ ) will propagate the carry into the block. That is, all the carries inside the block as

well as the carry out of the block are going to be the same as the carry into the block. We will call a block with this property a *skip block*.

Normally a long-range carry signal starts at a block  $B_i$ , rippling through some bits in that block, then skips some blocks, and ends in a block  $B_j$ . If the carry does not end at the LSB of  $B_j$  then rippling occurs in that block. The worst case delay occurs when  $i=1, j=N/M$ , the carry signal starts at the LSB of  $B_i$  and ends at the MSB of  $B_j$ . In such a scenario rippling occurs through  $(M-1)$ -bit positions of  $B_j$ . In order to eliminate the delay due to this rippling, we define the carry-strength ( $CS$ ) signal for each bit position in an  $M$ -bit block as follows:

if  $k$  is the LSB of the block then  $CS_{k+1} = \overline{x_k \oplus y_k}$ ,  
 otherwise  $CS_{k+1} = CS_k + (x_k \oplus y_k)$ .

In other words, for bit position  $k$  that is not the LSB of a block of bits, the incoming carry-strength  $CS_k$  is high if and only if the carry into the same position ( $C_k$ ) is independent of the carry-in of the block containing that bit position. In case  $CS_k=1$ , we also say that the carry-in  $C_k$  is *strong*. Otherwise  $C_k$  is *weak*. Carry-strength signals are useful in a block in which a long-range carry signal ends. To demonstrate the usefulness of carry-strength signals, let us consider the following two complementary cases. If  $CS_k=0$ ,

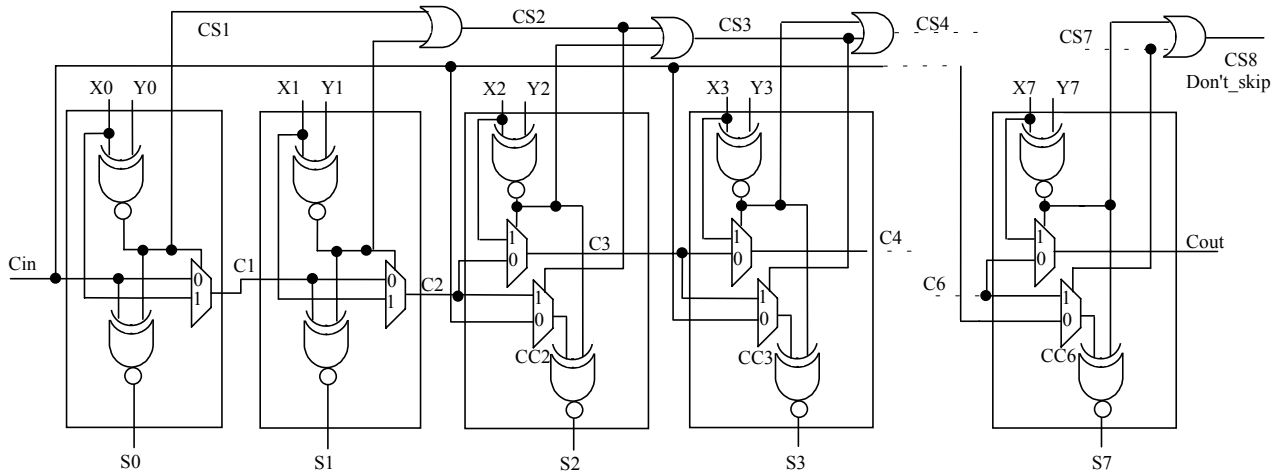


Figure 2. The new 8-bit block for carry-skip adders.

that is, the carry is weak, it is easy to verify that  $C_k$  corresponds to the block carry-in. Thus, we can just select  $C_k$  to be the same as the block carry-in, eliminating the delay due to rippling. On the other hand, if  $CS_k=1$ , then  $C_k$  is independent of the carry-in, and is therefore known quickly. (In other words the computation of  $C_k$  starts as soon as the adder's operands appear, without waiting for an incoming block carry-in). For these reasons, the existence of carry-

strength signals trivialises the delay in the ending block of a long-range carry signal, provided that the block carry-in is fed into a large enough buffer. Carry-strength signals can be easily computed in a ripple fashion implementing the above recursive definition. However, the latter rippling starts right when the operand are ready, not having to wait for the carry-in signal.



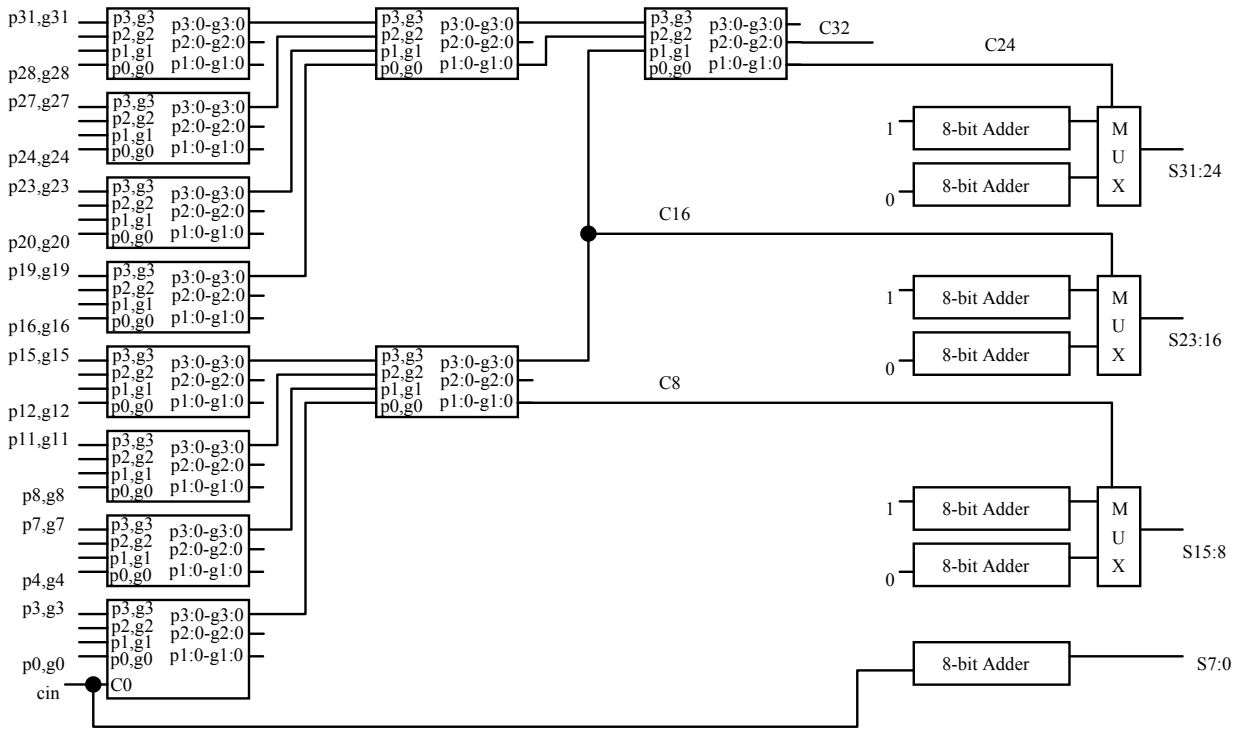


Figure 3. The carry tree of the 32-bit Lynch-Swartzlander's adder.

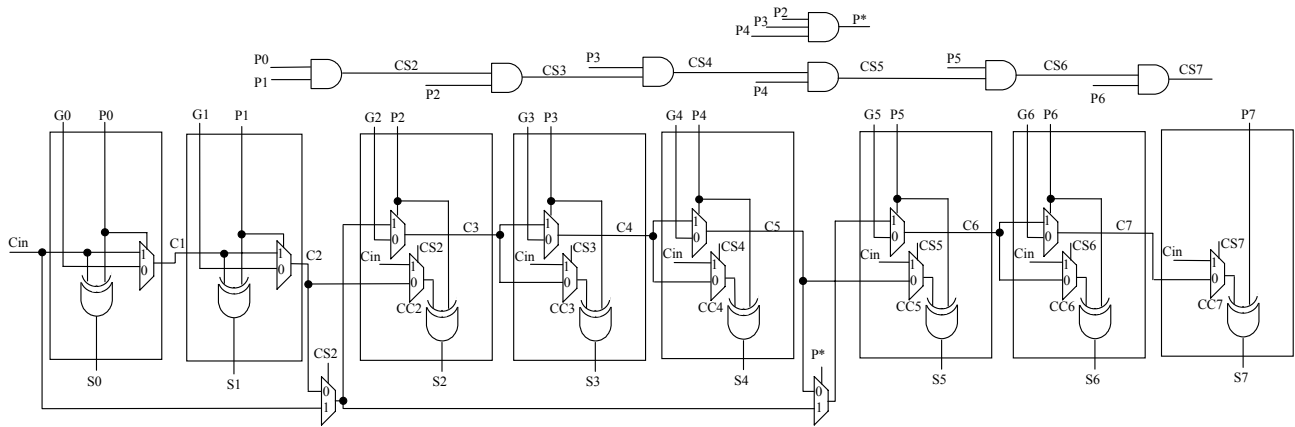


Figure 4. The 8-bit block for our version of Lynch-Swartzlander-style adder.

For this reason, their computation does not influence the critical path. It is worth pointing out that the carry-strength signal of the MSB position in a block ( $CS_M$ ) indicates whether that block can be skipped. Thus, no additional circuitry is required to compute that event.

### 3. A new eight-bit block for carry-skip adders.

The new 8-bit block for carry-skip adders is depicted in Figure 2. It can be seen that it has been designed to exploit the carry-strength signals without increasing the delay of the block when the carry is internally generated. In fact, the carry

propagation path is unchanged with respect to that of a conventional ripple-carry adder and totally new signals are formed to compute sum bits. These new signals (namely  $CC_i$ ) exploit in the best manner the parallelism allowed by the carry-strength signals ( $CS_2, \dots, CS_7$ ). In fact, even though a long-range weak carry will ripple through the carry propagation path (i.e.  $C_1, C_2, C_3, \dots, C_6$ ) the sum bits will always be valid after just  $\tau_{MUX} + \tau_{XNOR}$  from the time at which the carry arrives to the block. On the other hand, if the  $k$ -th FA receives a strong carry-in, produced in the LSB of the block calculates the carry-out after a delay  $\tau_{XNOR} + (k+1) \cdot \tau_{MUX}$  and the sum bit after a delay  $2 \cdot \tau_{XNOR} + (k+1) \cdot \tau_{MUX}$  from the time at which the operands arrive.

Conventional  
carry-skip

New carry-skip

### Carry-select

Carry  
look-ahead

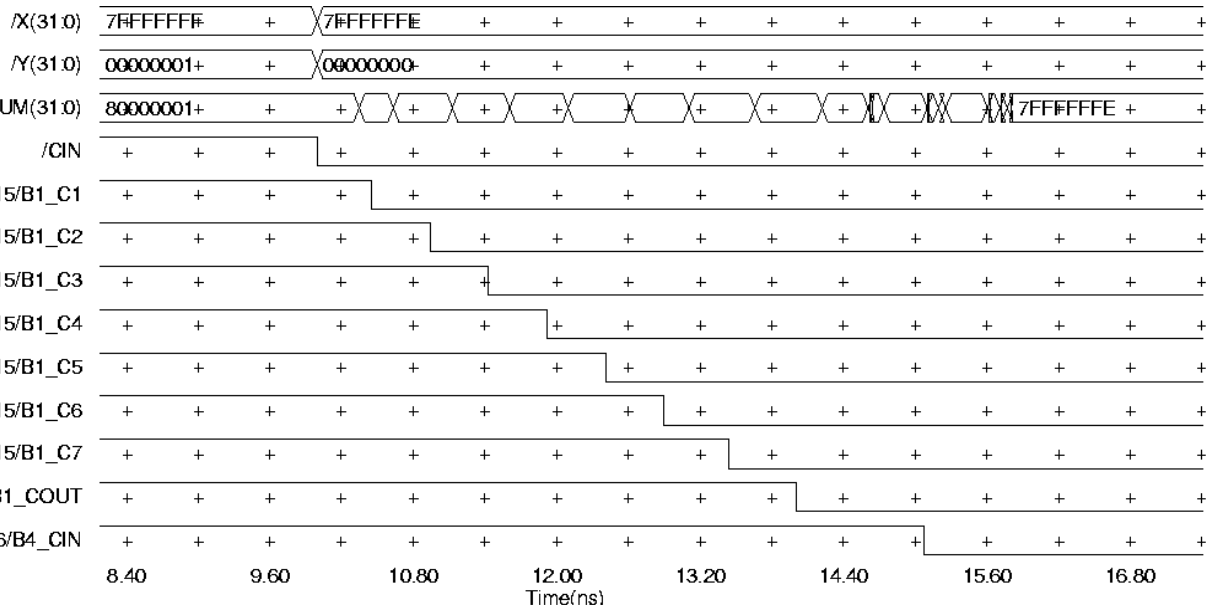


Figure 6. Gate level simulation results: worst operating transition for the new carry-skip adder.

#### 4. Introduction of modules based on carry strengths in the Lynch-Swartzlander adder.

Lynch-Swartzlander's [5] and Kantabutra's [6] hybrid carry-look-ahead adders are two of the fastest known adders whose area requirements are high because of the usage of carry select stages. Using carry select stages implies a duplication of the sum computation circuitry and

the use of a large number of multiplexers. As shown in Figure 3, the carry tree of the 32-bit version of the Lynch-Swartzlander's adder uses 4-bit lookahead generators (CLAG) to generate the carries into bit positions 8, 16, 24 and 32. The sum bits are obtained by means of 8-bit carry-select blocks, which perform their operations in parallel with the carry tree. The carries generated by the tree are then used to select the valid 8-bit sum words.

The same principle is applied to the Kantabutra's adder, which reaches higher speed performance due to a non-uniform carry-tree and recursive structure. Also the Kantabutra's adder generate the sum bits by means of carry-select stages. Thus in both these hybrid carry-look-ahead adders the delay introduced by the sum computation circuitry from the time in which the carry tree ends its computation is just  $\tau_{MUX}$ .

We investigate the possibility of using non-duplicate stages to obtain sum bits. These stages could be realised as carry-skip adders, which use the carry-strength logic. Since the blocks we are planning to use must work in parallel with the carry tree, we must make sure that when a carry is generated in such blocks, the blocks will complete their computations during the time in which the tree performs carries calculations.

The 8-bit block shown in Figure 4 has been designed specifically for our version of Lynch-Swartzlander-style adder. The signals  $P0, \dots, P7$  and  $G0, \dots, G6$ , correspond to the propagate and generate terms, respectively.

The new block is organised as a carry-skip adder using carry-strength signals ( $CS2, \dots, CS7$ ). It contributes to the

global delay with  $\tau_{MUX} + \tau_{XOR}$  when a weak carry-in (generated by the carry tree) dies in that block. On the other hand, when a carry is internally generated the worst case delay of such block is  $7 \cdot \tau_{MUX} + \tau_{XOR}$ . Thus its worst-case delay is not greater than that shown by a carry-select 8-bit block.

## 5. Optimisation of Bit Block Lengths

The bit block lengths can be optimised using a procedure adapted from that which was presented by Kantabutra [2]. We will only describe it roughly here. The procedure starts off by finding the largest MSB m-bit block such that its delay generated from the least significant bit of this block and terminated at the MSB of the same block is no more than some figure d, which we try to minimise. Then, the procedure must add less significant blocks to the left of the first one without making the worst case delay path longer than d. To this end, we must take into account the fact that a carry generated in such less significant blocks will terminate (in the worst case) in a more significant block, increasing its delay by just 1 MUX.

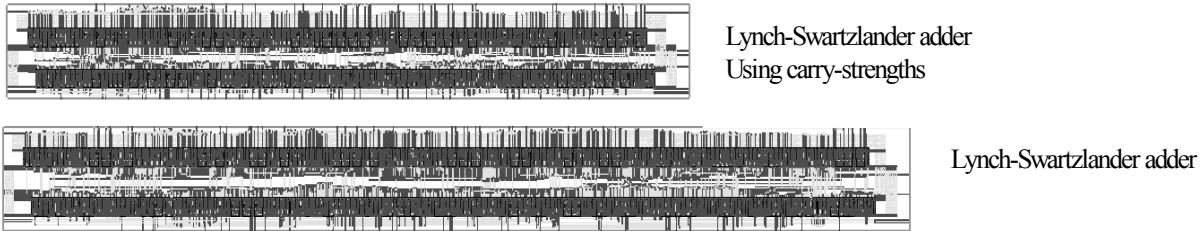


Figure 7. Layouts of the Lynch-Swartzlander style adders.

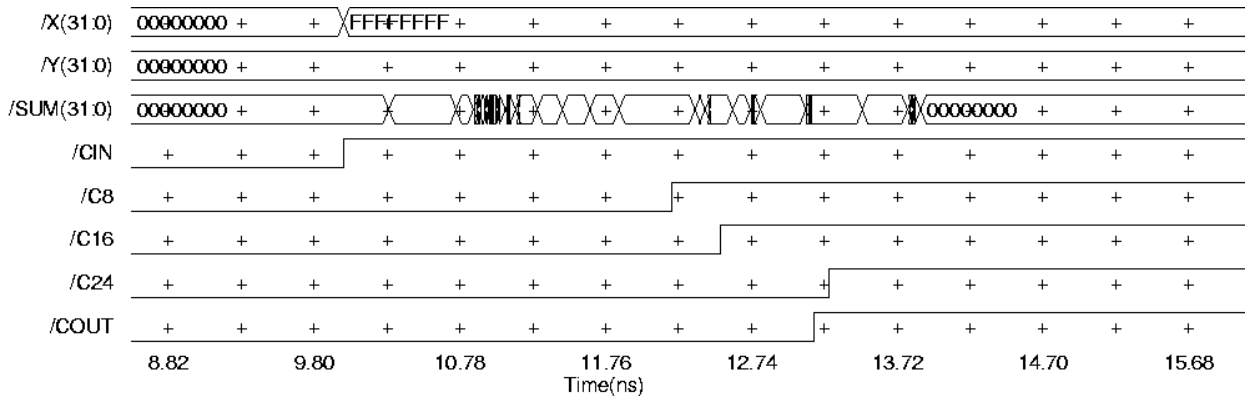


Figure 8. Gate level simulation results: worst operating transition for the new spanning tree adder.

Using this optimisation procedure gave rise to the following block sizes for a 32-bit adder: 2-4-5-6-7-8, with the largest block on the MSB side. The delay, area, and power are reported in the next section along with all the other simulation results.

## 6. Results

The new 8-bit block described in Section 3 has been used realising a new carry-skip adder. It has been compared to a conventional ripple-carry, a BCLA (implemented using 4-bit CLA blocks and two levels of look-ahead-carry generator [4]), a conventional carry-skip, and a carry-select adder



## New VLSI circuits for fast, low-cost binary adders

(carried out by means of 4 8-bit blocks). All the above adders have been realised using Austria Mikro Systeme p-sub, 2-metal, 1-poly, 5 V, 0.6 $\mu$ m CMOS process (CUB process) and have been implemented for 32-bit wide operands. Their layouts have been obtained using Mentor Graphics tools and the same level of optimisation has been used for each one. In Figure 5, the obtained layouts are depicted.

Gate-level and transistor-level (using BSIM3v3 device models at 27°C) simulations have been performed.

Worst-case delays have been identified considering the worst transition on operands inputs. This implies the asymmetric behaviour of logic gates being taken into account. As an example, note that the analysed adders for the operand transition 7FFFFFFF+00000001+1  $\rightarrow$  7FFFFFFE+00000000+0 shows a delay slightly greater than the transition 00000000+00000000+0  $\rightarrow$  7FFFFFFF+00000001+0, which corresponds to an opposite carry transition.

Power measurements have been carried out for the operand transition which appears to produce the maximum number of gates switching (FFFFFFF+ FFFFFFFF +0  $\rightarrow$  FFFFFFFF +00000000+0), assuming a 40-MHz repetitive frequency.

In Figure 6, gate level simulation results showing the low carry propagation through the critical path of the new carry-skip adder are reported. It can be seen that normal rippling occurs in block  $B_7$ . Then, the travelling carry reaches block  $B_4$ . There, each full-adder receives a weak carry-in. Thus,  $B_4$  computes sum bits in just 860ps, avoiding the need of rippling.

Post-layout simulation results summarised in Table 1 demonstrate that the new adder has a performance comparable to that of the BCLA, while its area requirement and power consumption are similar to that of a conventional carry-skip adder.

*Table 1. Performances and costs comparison.*

| 32-bit adders            | Area [ $\mu$ m <sup>2</sup> ] | Delay [ns] | Max Power [mW] |
|--------------------------|-------------------------------|------------|----------------|
| Ripple-Carry             | 137700                        | 15.8       | 15.8           |
| Carry-skip               | 160173                        | 9          | 17.1           |
| New e s t Implementation | 181944                        | 5.8        | 21             |
| Carry-Select             | 264966                        | 5.9        | 26             |
| BCLA                     | 310168                        | 5.5        | 36.4           |

The new 8-bit block described in Section 4 has been used to realise a 32-bit hybrid carry-look-ahead adder and it has been compared to the 32-bit version of the Lynch-Swartzlander's adder. Both circuits have been carried-out using the above-mentioned technology and their layouts are depicted in Figure 7.

In Figure 8, gate level simulation results related to the new hybrid adder are reported. Note that sum bits are computed

just 700ps later than  $C_{24}$ . Post-layout simulation results summarised in Table 2 show that the new implementation allows power dissipation and area to be significantly reduced without compromising speed.

It is worth pointing out that under the crude gate-counting delay model the new hybrid carry-look-ahead adder was expected to be slower than the conventional one by about  $\tau_{XOR}$ . However, post-layout simulations have shown that the new adder is in fact slightly *faster* than the original! This is due to the lower loads on the carry signals produced by the carry tree in the new implementation. In fact in the new adder the above lines are loaded by the input  $I1$  of eight multiplexers (25fF $\times$ 8), while in the Lynch-Swartzlander's implementation they drive the selection input of eight multiplexers (45fF $\times$ 8). Moreover, a reduction in net congestion leads more compact layout and shorter interconnection delays.

*Table 2. Performances and costs comparison.*

| 32-bit adders            | Area [ $\mu$ m <sup>2</sup> ] | Delay [ns] | Max Power [mW] |
|--------------------------|-------------------------------|------------|----------------|
| L y n c h - Swartzlander | 419244                        | 4.08       | 51             |
| N e w implementation     | 318550                        | 3.88       | 34             |

Finally we turn our attention to the carry-skip adder with optimised block sizes introduced in Section V. Using the same 0.6 micron, 5V AMS standard cell technology as before, we report our result in Table 3:

*Table 3. Characteristics of the optimised carry-skip.*

|          |                             |
|----------|-----------------------------|
| Delay    | 4.51ns                      |
| Area     | 194766 $\mu$ m <sup>2</sup> |
| MaxPower | 28 mW @ 40MHz               |

Table 4 shows a summary of Zimmermann's study, using a 0.8-micron standard-cell process. The figures are normalised so that they are 1 for the ripple adder. Table 5 shows a summary of our study, likewise normalised. All results shown in these tables are for 32-bit adders. Differences in the tables are due to several reasons. Firstly, the adders referred in [10] have been realised using a 0.8 $\mu$ m CMOS technology, whereas those summarised in Table 5 have been realised with a 0.6 $\mu$ m CMOS technology. However, the observable differences between Tables 4 and 5 are mainly due to the fact that different topologies have been used for the Carry-skip, Carry-Select and Carry Look-Ahead adders. For these addition circuits we used a uniform number of levels as well as uniform bit-block sizes (with the exception of our "new, optimised" adder presented in the last row of Table 5), whereas non-uniform number of levels and non-uniform bit-block sizes have been used in [10].

*Table 4. Relative comparison obtained in [10].*

| 32-Bit Adders | Area | Delay | Avg. Power |
|---------------|------|-------|------------|
| Ripple        | 1    | 1     | 1          |
| Carry-Skip    | 1.3  | 0.45  | 1.13       |
| Carry-Select  | 1.8  | 0.36  | 1.72       |

|            |      |      |      |
|------------|------|------|------|
| Lookahead  | 1.8  | 0.36 | 1.4  |
| CIA-1Level | 1.48 | 0.37 | 1.24 |
| CIA-2Level | 1.57 | 0.35 | 1.29 |

Table 5. Our relative comparison results.

| 32-Bit Adders  | Area | Delay | Max Power |
|----------------|------|-------|-----------|
| Ripple         | 1    | 1     | 1         |
| Carry-Skip     | 1.16 | 0.57  | 1.08      |
| Carry-Select   | 1.92 | 0.37  | 1.65      |
| Lookahead      | 2.25 | 0.35  | 2.3       |
| New, not opt.  | 1.32 | 0.37  | 1.33      |
| New, optimised | 1.41 | 0.28  | 1.78      |

It can be seen that our new optimised adder outperforms both CIAs saving silicon area.

## 7. Conclusions

In this paper adders based on the *carry-strength* signals have been presented. Carry strength signals are the negation of propagate signals, but represented as functions of one variable (bit position) instead of two. Such clearer representation gave rise to an entirely new bit block structure that has been applied to two families of adders. In these new results we have, apparently for the first time since the carry itself, utilised a ripple signal that uses regular gate logic. We demonstrated here that our new principle allows reaching very high speed at a low cost. The new carry-skip adder implementation is only ~5% slower than traditional BCLAs, while requiring much less silicon area and power dissipation. Significant reduction in silicon area and power has been also obtained in hybrid carry-look-ahead adders using the carry-strength signals without compromising speed. Finally, comparison with 1-level and 2-level carry-increment adders, though not direct, also gives favourable results.

## References.

- [1] Koren, I.: "Computer arithmetic algorithms", Prentice-Hall, 1993
- [2] Kantabutra, V.: "Designing optimum one-level carry-skip adders", *IEEE Trans. on Comp.*, 1993, Vol. 42, n°6, pp.759-764.
- [3] Chan, P.K., Schlag, M.D.F., Thomborson, C.D., Oklobdzija, V.G.: "Delay optimization of carry-skip adders and block carry-look-ahead adders", *Proc. of Int'l Symposium on Computer Arithmetic*, 1991, pp.154-164.
- [4] Nagendra, C., Irwin, M.J., Owens, R.M.: "Area-time-power tradeoffs in parallel adders", *IEEE Trans. CAS-II*, 43, (10), pp. 689-702.
- [5] T. Lynch, E.E. Swartzlander, "A spanning-tree carry-look-ahead adder", *IEEE Trans. on Comp.*, Vol. 41, n°8, Aug. 1992.
- [6] V. Kantabutra, "A recursive carry-look-ahead/carry-select hybrid adder", *IEEE Trans. on Comp.*, Vol. 42, n°12, Dec. 1993.
- [7] J. Park, H.C. Ngo, J.A. Silberman, S.H. Dhong, "470ps 64bit Parallel Binary Adder", *2000 IEEE Symposium on VLSI Circuits Digest of Technical Papers*.

[8] K. Raahemifar, M. Ahmadi, "Fast Carry-Look-Ahead Adder", *IEEE Canadian Conference on Electrical and Computer Engineering*, 1999.

[9] C.C. Yu, C.S. Lin, B.D. Liu, "A Generalized Block Distribution for Fast Carry-Skip Adder Design", *IEEE Tencon*, 1999.

[10] R. Zimmermann and H. Kaeslin, "Cell-Based multilevel Carry-Increment Adders with Minimal AT- and PT-Products, unpublished manuscript. <http://www.iis.ee.ethz.ch/~zimmi/>

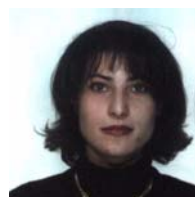
[11] A. Tyagi, "A reduced-area scheme for carry-select adders" *IEEE Trans. on Comp.*, Vol. 42, n°10, Oct. 1993.



**Vitit Kantabutra** was born in Bangkok, Thailand in 1958, and graduated from the Demonstration School of the Srinakharinwirot University, Patumwan in 1975. He graduated with Bachelor's Degree in Electrical Engineering, Honours Program at McGill University, Canada in 1979, and a Ph.D. in Electrical Engineering and Computer Science at The Johns Hopkins University, U.S.A., in 1985. He is currently an associate professor of computer science at Idaho State University, Pocatello, Idaho, U.S.A. He has numerous publications in international journals and conferences in the fields of computer algorithms, computational geometry, computer arithmetic circuits, and VLSI. He also holds U.S. patents on computer arithmetic circuits. Dr. Kantabutra is a senior member of the IEEE.



**Pasquale Corsonello** was born in 1964. He received his Masters degree in Electronics Engineering from the University of Naples, Italy, in 1988. He joined the Institute of Research on Parallel Computer System (IRSIP) of National Council of Research of Italy, working on the design and modeling of electronic transducers for high precision measurement. Since 1992 he joined the Department of Electronics, Computer Sciences and Systems of the University of Calabria, where he was involved in application specific IC design. Currently he is assistant professor at the Department of Electronics Engineering and Applied Mathematics of the University of Reggio Calabria. His main research interests are in arithmetic circuit, reconfigurable computing and VLSI design. Dr. Corsonello is a member of the IEEE.



**Stefania Perri** was born in 1971. In 1996 she received her degree in Computer Science Engineering from the University of Calabria, Italy. Since 1996 she joined the Department of Electronics, Computer Sciences and Systems of the University of Calabria, where she is working towards the Ph.D. degree in Electronics Engineering. Her current research interests include arithmetic circuit, reconfigurable computing and VLSI design.



# Programmers: A Critique on Communication and Software Engineering Issues

*Chatpong Tangmanee, Ph.D.*

*Faculty of Commerce and Accountancy, Chulalongkorn University*

**ABSTRACT** -- Viewing computer programmers from behavioral perspectives, this paper offers a critique on the essence of communication in how programmers do their work. The significance of communication is more notable when it is combined with rigorous software engineering practices and subsequently enables programmers to achieve and retain superb quality of their work. Toward the end, the paper identifies traditional communication practices as well as well-tested software engineering guidelines. Finally, it suggests three research opportunities, of which the findings would definitely fill gaps of knowledge about programmers, especially those concerning their behaviors.

**Key Words:**-- Computer Programmers, Interpersonal Communication, Software Engineering, Software Development

**บทคัดย่อ** -- บทความนี้นำเสนอโปรแกรมเมอร์ จากมุมมองของพฤติกรรมศาสตร์ โดยเฉพาะอย่างยิ่งจากมุมมองของทักษะการสื่อสาร คงไม่มีใครปฏิเสธถึงความสำคัญความสามารถในการติดต่อสื่อสารของโปรแกรมเมอร์ที่มีต่อการพัฒนาซอฟต์แวร์และระบบสารสนเทศ แต่ก็น่าแปลกที่มีนักวิจัยจำนวนไม่มากนักได้ทำการศึกษาเกี่ยวกับทักษะการติดต่อสื่อสารของโปรแกรมเมอร์โดยเฉพาะความสำคัญของการติดต่อสื่อสารของโปรแกรมเมอร์นี้ ไม่ได้หยุดอยู่ที่ว่าโปรแกรมเมอร์ควรเปิดกว้างและรับฟังความคิดเห็นของผู้ใช้ซอฟต์แวร์และผู้ร่วมงานเท่านั้น ทว่านักวิชาการทางวิศวกรรมซอฟต์แวร์ได้ตระหนักถึงความสำคัญของการติดต่อระหว่างบุคคลและนำมาปรับร่วมกับแนวคิดทางการพัฒนาระบบสารสนเทศ จนได้ออกเป็นแนวปฏิบัติหลักทางวิศวกรรมซอฟต์แวร์ ที่ทำให้โปรแกรมเมอร์สามารถสร้างและรักษาผลงานที่มีคุณภาพนั้นไว้ได้ด้วย ผู้เขียนได้สรุปในตอนท้ายถึงแนวทางของการวิจัยในอนาคตเพื่อที่ผลลัพธ์และข้อสรุปจะช่วยฉายภาพที่เฉพาะ สำคัญและชัดเจนขึ้นของโปรแกรมเมอร์ โดยเฉพาะอย่างยิ่งส่วนที่แสดงถึงทักษะและพฤติกรรมทางการติดต่อสื่อสาร

**คำสำคัญ** -- โปรแกรมเมอร์ การติดต่อสื่อสาร วิศวกรรมซอฟต์แวร์ การเขียนโปรแกรม การพัฒนาระบบสารสนเทศ

## 1. Introduction

The goal of this manuscript is to explicate significant implications of communication in the context of programmers and to call for more attention on research into the programmers' behavior. As one of the most critical components in the information superhighway (Tapscott, 1996), computer programmers must possess not only technical expertise but also communication skills so as to achieve programming excellency.

Next to this introduction is a description of computer programmers and their work. The following section will discuss the importance of communication to computer programming. It addresses basic communication practices and software engineering guidelines, both of which demand efficient interpersonal skills among programmers as well as those between them and users. In the final section is a concluding remark that suggests future research opportunity to explore programmers' communication behavior.

## 2. Computer Programmers and Their Work

Computer programmers design, write, test and maintain instructions--commonly known as a program or software--that describe logical steps for a computer to perform functions and produce final output (Brooks, 1995). Instructional steps correspond to (1) the input that a computer needs to perform the functions, (2) the process by which a computer could generate the results, and (3) the output and its appearance. These three major parts are known in general as program requirements. As programmers learn about the requirements, they may break them down into a logical series of instructions and write computer code representing the instructions for a computer to follow and generate an output. The code can be in any conventional programming language such as C or a more advanced one such as C++ or Java.

Programmers are not required to have formal education or training (i.e., they could be self-learning programmers or hold educational degrees in various fields) and may have different job titles (such as programmers, software

developers, or software engineers). Such a hybrid combination of background information makes the profiles of programmers relatively unique and different from those of other similar professionals and, thereby, implies a serious urgent need for specific research into programmers.

Researchers have found that programmers hold a set of basic personality traits (Brooks, 1995; Weinberg, 1982, 1998). They are frequently seen as having a relatively large ego. Perhaps, it is this large ego that makes it difficult for programmers to accept any criticisms or to admit their faults (Weinberg, 1971). Weinberg (1998) further notes that the frequent incidents of programmers' inability to admit their mistakes may be due to the feeling of being attached to their work. This type of feeling is generally shared among individuals whose routine work heavily involves creativity and abstract symbols. Such individuals include artists or poets (Brooks, 1995; Weinberg, 1982). Programmers do enjoy "making things" (Brooks, 1995, p. 7). They are often excited to see their own computer programs compiled and executed. The climax of writing code is to learn if the code could generate the desired output (Weinberg, 1971). Programmers are also highly tolerant to working under pressure. They can stay overnight to meet the due date (Ullman, 1995). Finally, Brooks (1995) claims that programmers have a passion for learning. They make every effort to stay up-to-date in technology and tend to apply what they learn to their work. This may be one of the reasons that programmers tend to "over-do" their work (i.e., try to make a program shorter or use fewer computer resources for the same output), which leads to the high cost of development (Weinberg, 1971; Yourdon, 1996). Neumann (1998) comments that "even the very best programmers may have a tendency toward total absorption in writing and debugging code" (p. 128).

Not only do programmers have certain personality traits, but their programming work also retains certain characteristics. Software researchers note that these characteristics of program development may explain a great demand for effective communication between programmers (Kiesler, et al., 1994; Kraut & Streeter, 1995). These characteristics include:

#### ◆ Scale

Kraut and Streeter (1995) note that one of the fundamental characteristics of software systems is that they are so large and complex that an individual or a small group is unable to handle the development. Programming is no longer the only essential skill for the success of software development (Kiesler, et al., 1994; Turley, 1991). The development of software containing millions of lines of code requires that the software be "compartmentalized" into small, manageable modules for which a group of programmers could be responsible (Kraut & Streeter, 1995, p. 70). Effective communication between group members and across the groups is thus a challenging task.

As the size of software becomes large, the task of software development also becomes tedious and expensive (Brooks, 1995). One approach used to minimize parts of the task involves communication-related activities between

programmers (Humphrey, 1994; Yourdon, 1996). These activities include peer reviews in which all parties are routinely informed of work progress. Thus a programmer leaving a group would not stall group progress since other members have learned the programmer's work from the reviews and they are able to carry on the development.

#### ◆ Uncertainty

A large scale per se may not be a unique characteristic of software development since other manufacturing processes are also large and complex. According to Kiesler and colleagues (1994), the production lines of automotive factories, textile mills, or tuna canneries are complex, yet many of them run smoothly. Manufacturing involves pre-determined steps of producing the same parts repeatedly. Developing software, however, is more uncertain because specifications of the software frequently change over time (Brooks, 1995; Kraut & Streeter, 1995). The change is due to two critical factors. One is the limited shared knowledge among involved parties such as users, stakeholders, programmers or system analysts. This limited knowledge includes the knowledge of software functionality identified by users and that of application domain in which the software will be exploited (Kiesler, et al., 1994). As Kraut and Streeter (1995) claim, system analysts handle user requirements and translate them into software requirements. While the analysts "adopt users' points of view," programmers frequently focus on technical aspects of the software such as ease of development or efficiency of computation (Kraut & Streeter, 1995, p. 70). Different perspectives from different groups of involved parties often lead to the constant change of software specifications.

The second group of factors that cause uncertainty in software development are those from external environments. The dynamics of financial conditions, regulations, company matters, or competition could invoke a change in software specifications (Kiesler, et al., 1994). New telecommunication regulations may force a development team to redesign software features, or users may demand, after the first weeks of working with a new software system, new capabilities that they had not envisioned earlier. Software scholars maintain that the changes in software specifications result in more uncertainty in the programming environment than do the external factors (Kraut & Streeter, 1995).

#### ◆ Interdependence.

According to Kiesler and her colleagues (1994), "... uncertainty in software development would be less of a problem if software did not require precise integration of its components at the end of the project" (p. 217). Because its gigantic scale (e.g., millions of lines of code) demands compartmenting a project into several modules, programmers must integrate all of them at the end with absolutely no defects (Brooks, 1995; Cusumano & Smith, 1997; Kraut & Streeter, 1995). A defect from improper integration, though each component performs properly, could result in fatal consequences (Kraut & Streeter, 1995). Hence, to succeed in building software, programmers must

take into account the interdependent nature of software development and consider an appropriate process to handle it.

Working in such a meticulous manner (i.e., making certain all modules and their integration function properly) places an enormous amount of pressure on programmers. As such, it should not be surprising that programmers often demand to work in isolation with, ideally, no interruption so that they can concentrate on their work and, perhaps, cope with stress caused by the pressure (Ullman, 1995). Besides the office condition, the pressure of building software stems from a rigorous routine of software testing (Cusumano & Smith, 1997). Programmers are engaged in different testing steps to assure high quality of the developed software (Kiesler, et al., 1994). Late delivery of many software products has happened because of a delay in the testing process (Yourdon, 1996).

In summary, a focus of this paper is on programmers who have main responsibilities for designing, writing, testing or maintaining computer programs. Coming from varied backgrounds, programmers are uniquely characterized as having (1) a large ego, (2) an ability to work under pressure, (3) a psychological attachment to their programming work and (4) a passion for constant learning and for working with abstract instructions. Further, the software development process holds a set of distinct characteristics in which the process is considered large in scale, uncertain and interdependent. It should be noted that the list of characteristics of programmers and that of program development are by no means exhaustive. They are instead better conceptualized as a set of major personality traits of programmers and critical attributes of software development. Learning about these personalities and attributes helps in understanding the implications of communication for programming work, discussed in the next section.

### 3. Critical Implications of Communication to Computer Programming

Programming is a collaborative effort (Brooks, 1995; Kraut & Streeter, 1995). According to McConnell (1993), a highly experienced programmer “recognizes that programming is only 15 percent communication with the computer, that it’s 85 percent communicating with people” (p. 764). Communication is thus one of the essential factors leading to programming success. This section presents in three parts a review that illustrates the importance of communication to programming work. The first part deals with general communication practices that have a positive impact on the ways in which programmers do their work. The second part is an overview of software engineering approaches that offer programmers rigorous communication-oriented guidelines to improve their programming performance. Finally, problems and possible solutions are included so as to emphasize on the significance of communication to programming work.

#### 3.1 Basic Communication Practices

Organizations often suggest that programmers have open communication with their colleagues or clients. The emphasis of open communication in the literature has been on the interaction between programmers and users during the stage of determining software requirements (or specifications) (Sawyer & Guinan, 1998).

Not only is communication between programmers and users important, but the interaction among programmers also plays a major role in the success of program development. In a study of coordination among software developers, Kraut and Streeter (1995) discovered that the participating developers placed high value on exchanging information and coordinating programming activities. They recommended that a large team of programmers whose work is to develop a technologically intensive program pay more attention to communication among team members as it could improve their performance.

Among several general communication practices, software researchers and practitioners agree on the critical implications of communication structures, feedback exchange and informal communication for programming work, each of which is discussed below.

##### ◆ Communication Structures

The communication in a programming team can be centralized or decentralized. Through the former, any interaction among team members involves the team leader, who acts as the center of communication. The latter, however, has no nucleus for interaction and the team members can communicate directly to each other (Shapiro, 1993). To compare productivity, satisfaction and leadership between the two communication structures in the programming environment, Shapiro (1993) conducted a controlled experiment on eighteen teams of four student programmers. While the members of the first nine teams had interaction with each other only through the appointed leader (centralized communication), those of the other nine teams did not have this constraint. Instead, they interacted freely with each other (decentralized communication). All the interaction had to be through computers; no other communication was reported to be allowed in Shapiro’s study. Her findings showed no differences in programming productivity, satisfaction or leadership between these two communication structures. However, there were several weaknesses in Shapiro’s (1993) study. Besides using students in substitution for actual programmers, her experiment design was flawed by the controlled condition that allowed only computer-mediated communication even though the subjects were in the same (or adjacent) room. Under different conditions, the effects of the two communication structures might have been distinct.

##### ◆ Feedback Exchange

Feedback exchange is a form of communication that has a positive impact on programming work. Programmers at Microsoft customarily subject their computer code for feedback not only to immediate colleagues but also to those

in different departments (e.g., Marketing or Pricing) who are stakeholders in the program (Cusumano & Smith, 1997; Yourdon, 1996). Bill Gates emphasizes the importance of becoming a highly competent programmer by learning how to give and receive feedback on programming products (Lammers, 1986). To offer feedback on any product, one has to read and thoroughly understand the program. This means that a feedback giver, after reading the code, should share an understanding of the program with its author. To receive feedback, however, may not be so smooth as it should be. Some programmers may have difficulty receiving and handling feedback (Brooks, 1995). Weinberg (1971) blames it on their large ego. Nonetheless, once the programmers agree to listen to feedback, they are likely to use it to improve the quality of their programs (Weinberg, 1982). Sawyer and Guinan (1994) found that using an electronic meeting system to project programming code on a large, wall-mounted screen seemed to draw attention of the meeting members to the screen, not to the code owner. This may make it easier for the code owner to receive feedback from other members.

The process of exchanging feedback has been incorporated in the hiring process at Microsoft (Cusumano & Selby, 1997). Not only will the new Microsoft programmers learn from the experienced colleagues about what has not been taught in college, the mentors also benefit from the younger programmers' fresh programming concepts and other innovative technological aspects (Yourdon, 1996).

#### ◆ Informal Communication

Informal communication (i.e., the interaction in a hallway or a request for a personal favor) has been shown to be at least as efficient as formal contacts (Fish, et al., 1993; Kraut, et al., 1993). The interaction among programmers at a water cooler or a vending machine during coffee breaks appears to have a positive impact on programming performance (Weinberg, 1998).

Regarding informal communication, Kraut and Streeter (1995) found that information obtained informally is considered more practical than that received via formal channels such as a departmental meeting or a session with clients. Additionally, subjects in Kraut and Streeter's (1995) study noted that the discussions with peers offered valuable feedback, but occurred less often. The similar incident of help or feedback offered through informal communication was also evident in software design teams (Curtis, et al., 1988).

This section (2.3.1 Basic Communication Practices) has presented important evidence that programmers "talk" to colleagues and clients as a means to enhance their productivity. Findings from the literature also exhibit how basic communication practices such as informal interaction, centralized or decentralized communication, and feedback exchange affect programmers and their performance. Organizations frequently encourage programmers to have open communication with other people. In addition to this basic communication, software engineers have offered rigorous approaches by which programmers who routinely practice them can substantially improve their performance.

The following section will discuss these software engineering approaches.

### 3.2 Software Engineering Practices

A set of software engineering practices provide instructions through which participating programmers are able to enhance their programming performance. Some techniques are communication-oriented while the others focus solely on innovative programming concepts. Since the major interest in this paper is on programmers' communication issues, this section is limited to the software engineering practices that combine communication effort with programming expertise in order to improve programming productivity. The practices reviewed in this section are chief and egoless programming techniques, technical reviews, and personal software process (PSP) strategy. Note that the paper does not offer complete instructions for these instances, or include research involving them. Moreover, it makes no attempt to imply that these three practices are the most effective ways of developing software or are the only set of the kind. Rather, the paper includes them for their critical focus on combining communication effort with programming expertise to improve the software development process. Details of these software engineering practices are included primarily to underscore the critical implications of communication for programming work.

#### ◆ Chief and egoless programming approaches

Chief and egoless programming approaches are two team-oriented software engineering practices that can be used in combination with several software development techniques. A team of programmers who practice chief programming has a centralized autocratic structure in which major decisions are made at the top (e.g., the chief) level and its interaction pattern is vertical along the structure (Mantei, 1981). The egoless approach holds a democratic structure in which team leadership is rotated and given to the individual whose expertise is needed at a given time. Communication among members in an egoless team is horizontal (Mantei, 1981). Shapiro (1993) claims that communication among chief team members centers on the leader and implies that the chief approach is comparable to centralized communication. Members of the egoless team, on the other hand, do not have this constraint as they can interact freely with each other. The communication in an egoless team is thus conceptually similar to a decentralized pattern (Shapiro, 1993).

Mantei (1981) recommends that the chief programming approach is appropriate for quick development of large-scale software that handles a routine, well-structured problem. This type of software could be a typical business application such as an inventory system. The egoless approach, in contrast, is recommended for a small-scale and technically difficult project where accurate and reliable results are expected. Because of the constraints of the difficult requirements and the result reliability, the project requires a flexible development schedule and a great deal of collaboration among developers. Examples of projects for which the egoless approach may be a candidate are those in

research and development (R&D) departments (Bishop, 1995; Mantei, 1981).

#### ◆ Technical reviews

Technical reviews are software engineering practices that help a group of programmers (1) detect any potential flaws in a programming product, (2) keep group members, colleagues or clients informed of work detail and progress, and (3) achieve more uniform, or at least more predictable, programming work which makes the subsequent routine more manageable (Freedman & Weinberg, 1990; Johnson, 1998). In general, technical reviews call for meetings of individuals who are involved in the work under review (Fagan, 1976). The meeting members are those who are responsible for creating the work and those whose expertise could help improve the work quality.

In a review session, a person or a team responsible for the work under review (e.g., a blueprint of a system design, computer programming code or a plan of system testing) presents it to review members. These members may then discern possible errors and offer different ways to correct them. At the end of the review, the person (or the team) would learn about the hidden mistakes and possible corrections. The other members could learn about the work as well, and share the principle of how the work is created. One critical aspect of performing a technical review is that the review is not intended to be a serious evaluation in which the individuals in charge of creating the work are criticized (or evaluated) for their productivity. Rather, the review emphasizes possible improvement of the work and shared understanding among the review members (Glass, 1999; Johnson, 1998).

Software engineering practitioners and researchers have suggested different approaches for performing technical reviews (Fagan, 1976, 1986; Freedman & Weinberg, 1990). Included in this literature review are two approaches, inspections and walkthroughs. These techniques have been selected because of their wide acceptance in software development firms and because of the important role that communication plays in practicing them (Lammers, 1986; McConnell, 1993).

Referred to as a “formal” process of locating errors (Fagan, 1976, p. 189), an inspection is a technical review in which the artifact under scrutiny is often software requirements, program design, or computer code. The formality comes from an inspection’s rigorous three-step process: preparation, collection and repair (Porter, et al., 1997). First, a team of reviewers meet with an individual (or a team) in charge of creating the artifact for its review. Next, each reviewer examines the work thoroughly (e.g., reading in a line-by-line fashion is suggested for computer code) and attempts to detect as many defects in the artifact as possible. All members then hold the final session together to discuss the defects and possible corrections. For a very difficult project demanding highly reliable outcomes, the inspection process may be repeated (Fagan, 1986; Porter, et al., 1997).

A walkthrough is a technical review conceptually similar to an inspection. The distinctions between these two review techniques are (1) a walkthrough is performed in an informal and less structured manner than is an inspection, and (2) a walkthrough meeting is adjustable to a broader type of work while the inspection is commonly done for review of a program design or computer code (Freedman & Weinberg, 1990). A basic walkthrough session consists of a few people, one of whom is most familiar with the work and will present it to other members. During the session, the members could ask for clarification or suggest an alternative to enhance the presented work. Freedman and Weinberg (1990) comment that due to the informal style of a walkthrough, the presenter may choose to present some parts with which he is familiar and skip others which may cause controversy. As a result, software practitioners have deemed walkthroughs less useful than inspections.

#### ◆ Personal Software Process (PSP) Strategy

PSP is software engineering methodology based on the premise that improved discipline of programmers can help enhance their programming productivity. Humphrey (1995) asserts that a programmer can strengthen his discipline of building software by “improving communication among team members” (p. 104). According to Humphrey’s (1995) notion of discipline, PSP is not a new concept of software development, but rather a set of well-assessed guidelines that can be used in conjunction with other software development techniques.

In general, the PSP strategy offers four-step instruction by which programmers could improve their performance (Humphrey, 1988). Each step, except the first, is built on the previous one. The following is a brief overview of each PSP step, synthesized from Humphrey (1995). The first step, namely “baseline,” is to familiarize programmers with the personal discipline approach by which they learn how to assess their routine programming work. Results of the first step of PSP are not enhanced productivity, but rather a rigorously standardized and well-documented procedure of routine programming work. Once the programming work is standardized and its output can be measured objectively, programmers are ready for the second step of PSP known as the “personal planning process.” Based on concepts of software planning tools (e.g., CASE tools), the second step provides a programmer instructions on how to organize a quality plan. Several measurements (e.g., the numbers of lines of code written per hour or per module) are incorporated to ensure that the quality plan is well implemented (Humphrey, 1995, p. 65). Implementation of the quality plan is the third step of PSP. This step, referred to as the “personal quality management,” focuses on feedback and technical reviews as critical means to detect any errors or mistakes as early as possible so that correction could be made at the early stage of program development. The sooner a mistake is detected, the greater the quality and the smaller the cost of the final product (Humphrey, 1995). The last step of PSP is to maintain the programming quality. Known as the “cyclic personal process,” this step provides instructions on how to adjust the first three steps to handle other projects with more requirements while

maintaining the program quality standard (Humphrey, 1995).

Based on the four-step instructions, the PSP strategy incorporates several communication practices (e.g., feedback exchange, technical reviews, and communication structures) into programming expertise in order to achieve and maintain a programmer's improved performance (Humphrey, 1994, 1995).

In summary, chief and egoless programming approaches, technical reviews and the PSP strategy are three significant instances of software engineering practices that emphasize the abstract yet useful application of communication to programming work. As stated previously, these three instances are included so as to elaborate the critical implication of communication to program development, but not to instruct for each instance. Complete descriptions can be found in Freedman and Weinberg (1990) for walkthroughs and inspections, Weinberg (1998) for the egoless programming approach, Baker (1972) and Fagan (1976) for chief programming teams and Humphrey (1995) for PSP strategy.

### 3.3 Problems and Possible Solution

Evidence from academic research and practitioner's report have confirmed numerous problems in software development (Humphrey, 1995; Weinberg, 1998; Yourdon, 1996, 1999). These problems include long delay in delivery of over-budgeted software products, unmet user requirements, hidden programming bugs, or insoluble network problems. Some problems yield constant serious defect to software while the others may still wait for disaster.

A close look at the problems in software development shows that they can be classified into two major categories: technical and management problems. The former address difficulty in developing technically feasible software while the latter raise concerns on managing a software project. A variety of recent development in computer engineering (such as distributed environment for vast data processing) could be a key that solves technical problems. However, despite numerous tools (such as visual programming languages, or web-based collaboration tools), managerial difficulty still depends heavily on communication behavior of programmers. Such a set of these tools aim primarily at alleviating interaction problems among developers (Sommerville, 1996) and subsequently enhance productivity in managing software projects. This thus again emphasizes the essential role of programmers' efficient communication skills.

## 4. Concluding Remark

A review of computer programmer profiles exhibits that they possess unique characteristics. These distinctive traits prompt a serious need for research that deals specifically and exclusively with programmers. Programming work is also multifaceted and its success depends greatly on effective communication of programmers. This critique has indicated a critical relation of basic communication

practices to programming work. These practices include communication structure, feedback exchange, and informal communication. In addition, software engineering researchers have developed intensive guidelines from which the programmers learn how to balance their communication effort with programming expertise in a way that maximizes their productivity. The guidelines covered in this paper are chief and egoless programming styles, technical reviews and personal software process strategy.

Given the unique characteristics of programmers and their work, generalizing knowledge about their communication behavior from what other researchers have previously learned in other similar professions (e.g., engineers or scientists) could be inconclusive. This therefore leads to at least three research opportunities. First, it will be much enlightening for communication scholars and software practitioners to explore what the communication behavior of computer programmers actually is. The research topics may include their interpersonal skills, preferred modes of communication, mass communication consumption (e.g., newspaper reading), or information seeking behavior. Findings from such research could portray a behavioral side of programmers and may further, perhaps, help software project leaders to properly handle communication issues among the project members. Moreover, knowing programmers' communication activities may confirm (or contradict) professional harmony among programmers, engineers and scientists, frequently asserted in previous studies (Brooks, 1995).

Second, in addition to programmers' communication competence, knowledge about a link between communication practices and programming excellency is still largely unknown. It would yield a remarkable benefit to the programming community if we realize, for example, what are the group communication structure and the mode of communication that may excel programming performance. In addition, given the disparity between two major types of software development (e.g., in-house software and off-the-shelf package developments), recommendations on how to arrange effective communication policies would draw a great amount of attention from software development firms such as Microsoft.

Finally, there is virtually no studies investigating the combined effect of (1) basic communication activities and (2) software engineering guidelines on programming performance, although each has been proved to enhance productivity among programmers (Kraut and Streeter, 1995; Weinberg, 1998). An example of such a study could be an experiment in which comparison of productivity in three groups of programmers is made. Those in the first group employ only basic communication practices (e.g., communication structures) ascertained in previous studies to have positive impact on programming work. Those in the second group practice only recommended software engineering guidelines (e.g., technical review or egoless programming). In the final group are programmers engaged in both basic communication and software engineering practices. Should the joint effect is confirmed, it would



emphasize one line of critical yet largely ignored research that addresses behavioral aspects of programmers. On a practical side, the joint impact may suggest that software development firms pay equal attention to rigorous innovative software practices as well as programmers' essential traditional communication skills.

In conclusion, programmers are one of the critical constituents in the information-based environment. Although they may share professional similarity with other walks of life, a close look at the programmers' profile and their communication issues make them relatively unique. This thus suggests an urgent need to investigate a behavioral side of the programmers, in addition to the plenty of results on their technical side. Until the findings from systematic research become available, we may just see only the "hard" side of programmers and miss their "soft" side.

## 5. References

- Baker, F. T. (1972). Chief programmer team management of production programming. *IBM system journal*, 56-73.
- Bishop, A. P. (1995). *The use of computer networks in aerospace engineering*. Unpublished doctoral dissertation, NY: Syracuse university.
- Brooks, F. P. (1995). *The mythical man-month: Essays on software engineering*. MA: Addison-Wesley.
- Curtis, W., Krasner, H., & Iscoe, N. (1988). A field study of the software design process for large systems. *Communications of the ACM*, 31(11), 1268-1287.
- Cusumano, M. A., & Selby, R. W. (1997). How Microsoft builds software. *Communications of the ACM*, 40(6), 53-61.
- Cusumano, M. A., & Smith, S. A. (1997). Beyond the waterfall: Software development at Microsoft. In D. B. Yoffie (Ed.), *Competing in the age of digital convergence* (pp. 371-412). MA: Harvard business school press.
- Fagan, M. E. (1976). Design and code inspections to reduce errors in program development. *IBM system journal*, 3, 182-211.
- Fagan, M. E. (1986). Advances in software inspections. *IEEE transactions on software engineering*, 12(7), 744-751.
- Fish, R. S., Kraut, R. E., Root, R. W., & Rice, R. E. (1993). Video as a technology for informal communication. *Communications of the ACM*, 36(1), 48-61.
- Freedman, D. P., & Weinberg, G. M. (1990). *Walkthroughs, inspections, and technical reviews: Evaluating programs, projects, and products*. NY: Doset Book.
- Humphrey, W. S. (1988). Characteristics of the software process: A maturity framework. *IEEE software*, 5(2), 73-79.
- Humphrey, W. S. (1994). The personal process in software engineering. In *Proceedings of the third international conference on the software process: Applying the software process* (pp. 69-77). Los Alamitos, CA: IEEE computer society press.
- Humphrey, W. S. (1995). *A discipline for software engineering*. MA: Addison-Wesley.
- Kiesler, S., Wholey, D., & Carley, K. M. (1994). Coordination as linkages: The case of software development teams. In D. Harris (Ed.), *Organizational linkages: Understanding the productivity paradox* (pp. 214-239). Washington, D.C: National Academy Press.
- Kraut, R., Fish, R. S., Root, R. W., & Chalfonte, B. L. (1993). Informal communication in organizations: Form, function, and technology. In R. M. Baecker (Ed.), *Readings in groupware and computer-support cooperative work* (pp. 287-314). San Mateo, CA: Morgan Kaufmann Publications.
- Kraut, R., & Streeter, L. (1995). Coordination in software development. *Communications of the ACM*, 38(3), 69-81.
- Lammers, S. (1986). *Programmers at work*. Redmond, Washington: Microsoft Press.
- Mantei, M. (1981). The effects of programming team structures on programming tasks. *Communications of the ACM*, 24(3), 106-113.
- McConnell, S. (1993). *Code complete*. Washington: Microsoft Press.
- Neumann, P. G. (1998). Are computer addictive? *Communication of the ACM*, 40(3), 128.
- Porter, A. A., Siy, H. P., Toman, C. A., & Votta, L. (1997). An experiment to assess the cost-benefits of code inspections in large scale software development. *IEEE transactions on software engineering*, 23(6), 329-345.
- Sawyer, S., & Guinan, P. J. (1994). Team-based software development using an electronic meeting system: The quality pay-off. In M. Lee, B. Barta & P. Jaliff (Eds.), *Software quality and productivity: Theory, practice, education and training* (pp. 78-83). New York, NY: Chapman Hall.

- Sawyer, S., & Guinan, P. J. (1995). Applications development as ceremony: A cultural perspective. In M. K. Ahuja, D. F. Galletta & H. J. Watson (Eds.), *Proceedings of the first America conference of information systems* (pp. 54-56). NY: AIS.
- Shapiro, E. J. (1993). *Productivity, satisfaction, leadership and interaction: A comparison of two communication networks in software development teams*. Unpublished doctoral dissertation, MN: University of Minnesota.
- Sommerville, I. (1996). *Software engineering*. Fifth edition. NY: Addison-Wesley.
- Tapscott, D. (1996). *The digital economy: Promise and peril in the age of networked intelligence*. NY: McGraw-Hill.
- Ullman, E. (1995). Out of time: Reflections on the programming life. In J. Brook & I. Boal (Eds.), *Resisting the virtual life: The culture and politics of information* (pp. 131-143). San Francisco, CA: City Lights.
- Weinberg, G. M. (1971). *The psychology of computer programming*. NY: Von nostrand reinhold.
- Weinberg, G. M. (1982). *Understanding the professional programmer*. NY: Dorset House.
- Weinberg, G. M. (1998). *The psychology of computer programming (the silver anniversary edition)*. NY: Dorset House Inc.
- Yourdon, E. (1996). *Rise and resurrection of the American programmer*. NJ: Yourdon Press.
- Yourdon, E. (1999). *Death march*. NY: Prentice Hall, Inc.

# Optical Soliton Equation and Its Application in Optical Devices Design For High Speed Communication

C. Taepanich and P.P. Yupapin

Lightwave Technology Research Center, Department of Applied Physics, Faculty of Science

King Mongkut 's Institute of Technology Ladkrabang, Bangkok 10520

Tel: 7373000 ext. 6271/6213, Fax: 3269981

E-mail: s0065005@kmitl.ac.th or Yupapin.Preecha@kmitl.ac.th

**ABSTRACT:** The study of nonlinear Schrodinger equation is used to describe the propagation of optical signals in single mode optical fiber that is represented a model of transmission light pulse signals in an optical fiber. Basic soliton equation is briefly reviewed for the explanation of signals group velocity dispersion (GVD) and self-phase modulation (SPM) which effect the required optical transmission signals. The modeling of soliton pulses travelling into two or three coupling fibers are modeled then input into the soliton wave equations. The input optical signal parameters such as light pulse width, input signal bit rate, light source wavelength and bandwidth, and coupling power are employed in the soliton equation. The coupling output signals obtained are described for the use in optical transmission devices such as optical filter, switching and multiplexer. The physical behaviors of the optical output signals obtained by numerical technique are associated with the optical device parameters.

**Keywords:** Optical Communication , Optical Devices.

**บทคัดย่อ:** การศึกษานี้พิจารณาสมการชเรอดิงเงอร์แบบไม่เชิงเส้นในการใช้อธิบายการแพร่ของสัญญาณแสงในใยแก้วนำแสงชนิดโหมดเดี่ยว ซึ่งเป็นระบบหรืออุปกรณ์ที่แทนที่ด้วยระบบสมการที่สามารถอธิบายการส่งผ่านสัญญาณพัลส์แสงได้ โดยสมการนี้เรียกว่าสมการโซลิตอนได้ถูกนำมาใช้อธิบายปรากฏการณ์ของการกระจายความเร็วกลุ่ม และปรากฏการณ์เซลเฟสมอดูเลชัน ซึ่งเป็นปรากฏการณ์ที่มีความสำคัญของพัลส์โซลิตอนในใยแก้วนำแสง การศึกษานี้จะพิจารณาลักษณะของพัลส์โซลิตอนที่เดินทางไปในคัปเปิลอร์ใยแก้วนำแสงชนิด 2 และ 3 แกน และผลของสัญญาณเอดด์พุดจากการคัปปลิงระหว่างกัน นอกจากนั้นยังได้พิจารณาค่าพารามิเตอร์ของสัญญาณแสง ดังเช่น ความกว้างของพัลส์แสง อัตราการส่งข้อมูล หรือ ความเร็ว ความยาวคลื่นของแหล่งกำเนิดแสง แบนด์วิดท์ และกำลังจากการคัปปลิงแสง สัญญาณเอดด์พุดที่ได้จากการคัปปลิงนี้สามารถนำไปใช้พิจารณาในการออกแบบอุปกรณ์ทางแสง เช่น อุปกรณ์กรองสัญญาณ สวิตช์ ออสซิลเลเตอร์ และ มัลติเพลกเซอร์ โดยลักษณะของพารามิเตอร์และสัญญาณเอดด์พุดหาได้โดยวิธีการทางตัวเลข

## 1. บทนำ

ในระบบการสื่อสารทางแสงนั้นค่าของอัตราการส่งข้อมูล (bite rate ) จะถูกจำกัดโดยคุณสมบัติของการกระจาย (dispersion ) ของใยแก้วนำแสง ดังนั้นเพื่อเป็นการชดเชยผลที่วุ่นนี้ได้มีการใช้ลักษณะความไม่เป็นเชิงเส้นของตัวกลาง (nonlinearity ) ที่มีผลที่ทำให้พัลส์ที่ไม่เป็นเชิงเส้น (nonlinear pulse) มีความเสถียรและเคลื่อนที่ได้อย่างไม่มีการกระจายหรือเปลี่ยนรูปร่างซึ่งถูกเรียกว่า “โซลิตอน” (solitons) [1-6] ในการแพร่ของเอนเวลอปของโซลิตอน (envelope soliton) ในใยแก้วนำแสงจะอธิบายได้ด้วยสมการชเรอดิงเงอร์แบบไม่เชิงเส้น (Nonlinear Schrodinger Equation, NLSE ) สมการนี้สามารถหาคำตอบ

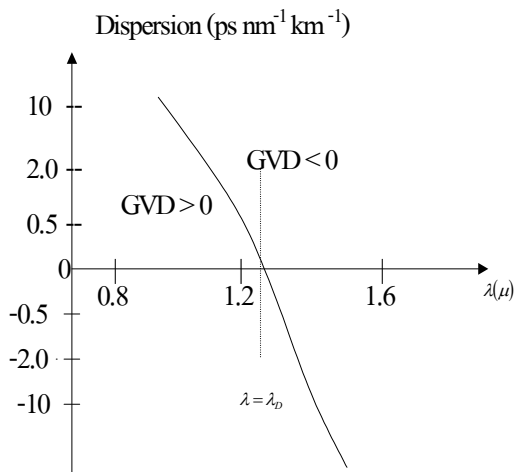
ได้ทั้งโดยวิธีการกระเจิงกลับ ( inverse scattering technique, IST) และโดยวิธีทางตัวเลข (numerical method) ปรากฏการณ์ที่สำคัญของพัลส์ที่เคลื่อนที่ในใยแก้วนำแสงชนิดโหมดเดี่ยวที่ทำให้เกิดโซลิตอนก็คือการกระจายความเร็วกลุ่ม (group velocity dispersion, GVD) และเซลเฟสมอดูเลชัน (self phase modulation, SPM) เนื่องด้วยข้อดีของคุณสมบัติของโซลิตอนดังกล่าวจึงทำให้มีการนำไปประยุกต์ใช้งานอย่างกว้างขวางโดยเฉพาะในอุปกรณ์สวิตช์สำหรับการสื่อสารความเร็วสูง เป็นต้น

## 2. การกระจายความเร็วกลุ่ม

ในใยแก้วนำแสงชนิดโหมดเดี่ยวนั้นการกระจายที่สำคัญที่สุดคือการกระจายความเร็วกลุ่มซึ่งเป็นผลจากความเร็วในการเดินทางของแสงมีความสัมพันธ์กับความยาวคลื่นแสงดังแสดงในรูปที่ 1

โดยเริ่มพิจารณาจากเลขคลื่น(k) ดังนี้

$$k(\omega) = n(\omega) \frac{\omega}{c} = k_0 + \left( \frac{\partial k}{\partial \omega} \right) (\omega - \omega_0) + \frac{1}{2} \left( \frac{\partial^2 k}{\partial \omega^2} \right) (\omega - \omega_0)^2 + \dots$$



รูปที่ 1 แสดงความสัมพันธ์ของการแพร่กระจาย

โดยที่  $k(\omega)$  เป็นเลขคลื่น  $n(\omega)$  เป็นดัชนีหักเห และ  $\omega$  เป็นความถี่เชิงมุม และจาก

$$\frac{\partial k}{\partial \omega} = \frac{1}{V_g} = k', \quad \frac{\partial^2 k}{\partial \omega^2} = k'' = \lambda^2 \frac{D}{2\pi c}$$

เมื่อ  $V_g$  เป็นความเร็วกลุ่ม และ  $k''$  สอดคล้องกับการกระจายความเร็วกลุ่ม (GVD) เมื่อ  $D$  คือค่าพารามิเตอร์การกระจาย จากกราฟในรูปที่ 1 แสดงการกระจาย (D) จะมีค่าต่ำเมื่อ  $\lambda = \lambda_0 \approx 1.3 \mu\text{m}$  และเมื่อ  $\lambda < \lambda_0$  จะได้การกระจายปกติ (normal GVD) ซึ่งมีค่า  $k''$  เป็นบวก ในทางตรงกันข้ามเมื่อ  $\lambda > \lambda_0$  จะได้การกระจายไม่ปกติ (anomalous GVD) ซึ่งมีค่า  $k''$  เป็นลบ ผลจากการกระจายนี้ทำให้พัลส์ขยายกว้างออก ด้วยเหตุนี้จึงจำเป็นต้องใช้ปรากฏการณ์ของความไม่เป็นเชิงเส้นที่จะทำการชดเชยการกระจายออกของพัลส์ได้

## 3. ลักษณะไม่เป็นเชิงเส้นของตัวกลาง

ในการตอบสนองของตัวกลางสารไดอิเล็กตริกที่มีต่อสนามไฟฟ้าความเข้มสูงนั้น จะเริ่มแสดงคุณสมบัติไม่เป็นเชิงเส้นเมื่อเกิดจากการเปลี่ยนแปลงดัชนีหักเหของแสง มีความสัมพันธ์กับความเข้มของแสงดังนี้

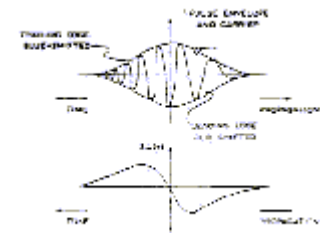
$$n = n_0 + n_2 I$$

โดยที่  $n_0$  เป็นดัชนีหักเหเชิงเส้น  $I$  แทนความเข้มแสง  $n_2$  เป็นสัมประสิทธิ์ของเคอร์ (Kerr coefficient)

จากสมการนี้ทำให้ทราบว่าค่าดัชนีหักเหของแสงขึ้นอยู่กับค่าความเข้มแสง ซึ่งเป็นผลให้เกิดการเลื่อนเฟส (phase shift) เป็นดังนี้

$$\Delta \phi = \frac{2\pi}{\lambda} L n_2 I$$

โดยลักษณะการเลื่อนเฟสที่เกิดขึ้นในทิศทางกันข้ามกับการกระจายแบบไม่ปกติซึ่งเรียกลักษณะการเลื่อนเฟสกรณีนี้ว่าเซลเฟสมอดูเลชัน และจากรูปที่ 2 เป็นการแสดงการเชิฟ (chirp) ของความถี่ซึ่งเนื่องจาก SPM ผลก็คือ มีความถี่สูงขึ้นที่บริเวณขอบของพัลส์ ซึ่งลักษณะของการเชิฟนี้จะแตกต่างกันกับการเชิฟที่เกิดจากการกระจายความเร็วกลุ่ม ดังนั้น SPM จะช่วยทำให้ที่ความยาวคลื่นยาวถูกหน่วงโดยผลการกระจายความเร็วกลุ่มแบบไม่ปกตินั้นสามารถทำให้สัญญาณกลับสู่สภาพที่ไม่มีการกระจายได้เพราะว่าปรากฏการณ์ทั้งสองเป็นปฏิกิริยาที่หักล้างกัน ส่วนที่ความยาวคลื่นสั้นนั้นก็เกิดขึ้นในลักษณะคล้ายกันซึ่งผลลัพธ์ก็คือได้พัลส์โซลิตอนที่ไม่มีการเปลี่ยนแปลงรูปร่าง



รูปที่ 2 แสดงเชิฟ(chirp) ความถี่เนื่องจาก SPM

## 4. สมการชเรอดิงเงอร์ไม่เชิงเส้น

พิจารณาสมการที่ใช้อธิบายการแพร่ของเอนเวลลอปของพัลส์โซลิตอน โดยคิดมาจากผล GVD และ ความไม่เป็นเชิงเส้นคือ

$$j \frac{\partial U}{\partial \xi} - \text{sgn}(k'') \frac{1}{2} \frac{\partial^2 U}{\partial \tau^2} + N^2 |U|^2 U = 0 \quad (1)$$

เมื่อ  $U$  เป็นแอมพลิจูดของพัลส์ที่มีการเปลี่ยนแปลงอย่างช้า ๆ  $\xi$  เป็นโคออร์ดิเนต  $\tau$  เป็นเวลาของพัลส์สัญญาณที่เคลื่อนที่ในกรอบอ้างอิงของความเร็วกลุ่ม และ  $j$  เป็นตัวแปรเชิงซ้อน และโดยที่  $\text{sgn}(k'')$  ให้ค่า 1 เมื่อ  $k''$  เป็นบวก (normal GVD) และให้ค่าเป็น -1 เมื่อ  $k''$  เป็นลบ (anomalous GVD) และพารามิเตอร์  $N$  คืออันดับ (order) ของโซลิตอน

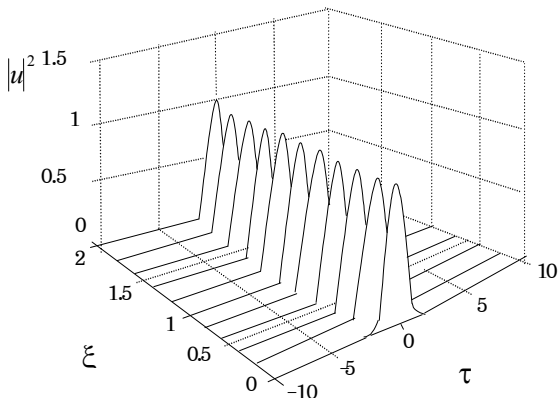
## 5. โซลิตอนพื้นฐาน และในอันดับสูงขึ้น

สมการ NLS สามารถแก้หาคำตอบได้อย่างถูกต้องด้วยเทคนิคทางคณิตศาสตร์ด้วยวิธีวิเคราะห์ที่เรียกว่าวิธีการการกระเจิงกลับ (IST) ทั้งสำหรับกรณีที่เป็นการกระจายแบบปกติ หรือไม่ปกติก็ตาม โดยที่ค่า  $k'' < 0$  ซึ่งเป็นการกระจายแบบไม่ปกติเรียกว่า “โซลิตอนด้านสว่าง” (bright soliton) ส่วนกรณี  $k'' > 0$  หรือการกระจายแบบปกติเรียกว่า “โซลิตอนด้านมืด” (dark soliton) ในทางปฏิบัติมักสนใจแต่ bright soliton เท่านั้นและมักสนใจในกรณีที่  $\lambda = 1.5 \mu\text{m}$  ซึ่งเป็นย่านที่ตรงกับการสูญเสียที่ต่ำที่สุดและการขยายสัญญาณด้วยเส้นใยแก้วนำแสงได้เป็นอย่างดี

### 5.1. โซลิตอนพื้นฐาน

ในการหาคำตอบโซลิตอนพื้นฐานนั้นได้จากการแก้สมการที่ (1) โดยวิธีวิเคราะห์ (analytical method) ซึ่งมีคำตอบเป็นฟังก์ชันลักษณะดังแสดงในรูปที่ 3

$$u(\xi, \tau) = \text{sech}(\tau) \exp(j\xi/2)$$



รูปที่ 3 แสดงการเคลื่อนตัวของโซลิตอนพื้นฐาน

จากรูป 3 จะเห็นว่าในกรณี  $N = 1$  นี้ พัลส์จะไม่เปลี่ยนรูปร่างตลอดการแพร่ของโซลิตอน

### 5.2. โซลิตอนอันดับสูง

กรณีนี้ค่า  $N$  เป็นจำนวนเต็มที่มีค่ามากกว่า 1 ขณะที่พัลส์เดินทางไปบนเส้นใยแก้วนำแสงนั้นพัลส์จะหดตัวและแตกออก และก็จะกลับมารวมกันอีกครั้งเหมือนกับพัลส์เริ่มต้นที่มีลักษณะเป็นคาบ  $2\pi$  ตำแหน่งซึ่งเหมือนเดิมนี้คือ  $\xi = m\pi/2$  เมื่อ  $m$  เป็นจำนวนเต็ม ลักษณะรูปคลื่นที่เปลี่ยนแปลงไปดังกล่าวเป็นผลมาจากความต่างเฟสระหว่างโซลิตอน

## 6. สมการการคับปลิง

พิจารณาระบบสมการ NLS ที่มีการคับปลิงด้วยการพิจารณาสมการที่ได้จากพัลส์แสงในใยแก้วนำแสงโหมดเดียว 3 แกน โดยของแต่ละแกนที่มีการกระจายความเร็วกลุ่มแบบไม่ปกติที่มีความไม่เป็นเชิงเส้นเป็นบวก คือ [1]

$$j \frac{\partial u_1}{\partial \xi} = -\frac{1}{2} \frac{\partial^2 u_1}{\partial \tau^2} - K(u_2 + u_3) - |u_1|^2 u_1$$

$$j \frac{\partial u_2}{\partial \xi} = -\frac{1}{2} \frac{\partial^2 u_2}{\partial \tau^2} - K(u_1 + u_3) - |u_2|^2 u_2$$

$$j \frac{\partial u_3}{\partial \xi} = -\frac{1}{2} \frac{\partial^2 u_3}{\partial \tau^2} - K(u_1 + u_2) - |u_3|^2 u_3$$

เมื่อ  $K$  คือสัมประสิทธิ์ในการคับปลิงเชิงเส้น  $u_1$ ,  $u_2$  และ  $u_3$  คือแอมพลิจูดของพัลส์แสง

## 7. วิธีการทางตัวเลข

สมการที่ 1 เป็นสมการ NLS ซึ่งเป็นสมการที่ไม่คิดผลของการรบกวน (Perturbation) ซึ่งสามารถแก้ไขได้โดยใช้วิธีการการกระเจิงกลับ แต่ถ้าเราต้องการคิดผลจากการรบกวนด้วยแล้วจะไม่ใช่การง่ายที่จะใช้วิธีการวิเคราะห์ ดังนั้นวิธีที่ดีและสะดวกกว่าก็คือวิธีการทางตัวเลข โดยเฉพาะกับการแก้ระบบสมการ NLS ซึ่งมีวิธีทางตัวเลขมีอยู่หลายวิธี เช่น วิธีการไฟไนต์ดิฟเฟอเรนซ์ (Finite Difference Method, FDM) วิธีการบีมพรอพเกชัน (Beam Propagation Method, BPM) และวิธีไฟไนต์เอลิเมนต์ (Finite Element Method, FEM) ซึ่งในที่นี้จะขอกกล่าวแต่เพียงวิธีไฟไนต์เอลิเมนต์เท่านั้น พิจารณาสมการ NLS ซึ่งคิดผลของการสูญเสียในใยแก้วนำแสง และ SSFS เข้าไปด้วยมีสมการเป็น [3]

$$-j \frac{\partial q}{\partial \xi} + \frac{1}{2} \frac{\partial^2 q}{\partial \tau^2} + |q|^2 q - j\Gamma q - R \frac{\partial |q|^2}{\partial \tau} q = 0 \quad (2)$$

จากสมการที่ (2) ถ้าเราตัดผลของการสูญเสียของใยแก้วนำแสง และ SSFS ( $\Gamma = R = 0$ ) ออกแล้วสมการที่ 2 จะลดรูปเป็นสมการ NLS ดังสมการ (1) นั่นเอง เมื่อ  $u$  และ  $q$  เป็นค่าเดียวกัน และในการคำนวณโดยมีเงื่อนไขเริ่มต้นคือ

$$q(\tau) = A \text{sech}(\tau) \quad (3)$$

สมการไฟไนต์เอลิเมนต์เริ่มต้นด้วยการแบ่ง  $\tau$  ออกเป็นคอร์ดควอดราติกไลน์เอลิเมนต์ (quadratic line elements) จำนวนมาก และเขียน  $q$  ในแต่ละเอลิเมนต์เป็นดังนี้

$$q(\xi, \tau) = \{N(\tau)\}^T \{q(\xi)\}_e \quad (4)$$

โดยการใช้วิธีการทาง FEM กับสมการ (2) จะได้

$$-j[M]\frac{d\{q\}}{d\xi} + [K]\{q\} = \{0\} \quad (5)$$

เมื่อ

$$[K] = \sum_c \int_c \left[ -\frac{1}{2} \{N_c\} \{N_c\}^T + |q|^2 \{N_c\} \{N_c\}^T - j\Gamma \{N_c\} \{N_c\}^T - R \frac{\partial |q|^2}{\partial \tau} \{N_c\} \{N_c\}^T \right] d\tau$$

$$[M] = \sum_c \int_c \{N_c\} \{N_c\}^T d\tau$$

โดยใช้ FDM กับสมการ 5 จะได้สมการเมตริกซ์ ภายในช่วง  $i \Delta \xi \leq \xi < (i+1)\Delta \xi$  คือ

$$-j[M]\{\{q\}_{i+1} - \{q\}_i\} / \Delta \xi + [K]_i [\theta \{q\}_{i+1} + (1-\theta)\{q\}_i] = \{0\}$$

ซึ่งสามารถสรุปได้เป็น

$$[L(\theta)]_i \{q\}_{i+1} = [L(\theta-1)]_i \{q\}_i$$

เมื่อ

$$[L(\theta)]_i = -j[M] + \theta \Delta \xi [K]_i$$

โดยที่  $\theta$  คือตัวควบคุมความเสถียรในการคำนวณโดย  $0 \leq \theta \leq 1$  โดยที่  $\theta = 1/2$  สำหรับวิธีการ แครง์-นิโคลสัน (Crank-Nicolson, CN) ที่ใช้ขั้นตอนสปริต (split-step procedure, SSP) ต่อไปนี้จะเป็นการเพิ่มประสิทธิภาพในการคำนวณได้

**Step 1 (dispersion effect) :**

$$-j[M]\frac{d\{q^{(1)}\}}{d\xi} + [K_0]\{q^{(1)}\} = 0 \quad (6)$$

เมื่อ

$$[K_0] = \sum_c \int_c \left[ -\frac{1}{2} \{N_c\} \{N_c\}^T \right] d\tau$$

**Step 2 (nonlinear effect) :**

$$\{q^{(2)}\} = \{q^{(1)}\} \exp(-j \int k d\xi) \quad (7)$$

เมื่อ

$$k = |q|^2 - j\Gamma - R \partial |q|^2 / \partial \tau$$

โดยที่ตัวเลข 1 และ 2 แทนขั้นตอน step ที่ 1 และ 2 ตามลำดับ และโดยใช้ FDM กับสมการ 6 จะได้

$$[L(\theta)] \{q^{(1)}\}_{i+1} = [L(\theta-1)] \{q^{(1)}\}_i \quad (8)$$

โดยที่

$$[L(\theta)] = -j[M] + \theta \Delta \xi [K_0]$$

จากสมการ 7 เขียนอีกรูปได้เป็น

$$\{q^{(2)}\}_{i+1} = \{q^{(1)}\}_{i+1} \exp(-jk_i \Delta \xi) \quad (9)$$

โดยการแทนสมการ 8 ลงในสมการ 9 สุดท้ายเราจะได้

$$\{q\}_{i+1} = [L(\theta)]^{-1} [L(\theta-1)] \{q\}_i \times \exp(-jk_i \Delta \xi) \quad (10)$$

สมการเหล่านี้เป็นสมการที่จะถูกนำไปประดิษฐ์โปรแกรมคอมพิวเตอร์โดยวิธีการทาง FEM ต่อไป ขั้นตอนดังกล่าวมาทั้งหมดนี้เป็นเพียงแค่สมการที่พร้อมนำไปประดิษฐ์โปรแกรมคอมพิวเตอร์เท่านั้น และขั้นตอนต่อไปจะต้องนำไปเขียนผังงานและในท้ายสุดก็ทำการเขียนโปรแกรม โดยในการทำจะเริ่มจากการใช้แบบจำลองของโซลิตอนพื้นฐานก่อน และหลังจากนั้นก็ทำการเพิ่มเป็นระบบสมการในการคำนวณทั้งในแบบ 2 และ 3 แกนเพื่อนำไปพิจารณาออกแบบอุปกรณ์ใยแก้วนำแสงเป็นสวิตช์ มัลติเพลกเซอร์และฟิลเตอร์ เป็นต้น

## 8. สรุป

คำตอบที่ได้จากระบบสมการการคำนวณนั้นสามารถนำไปใช้ในการพิจารณาออกแบบอุปกรณ์สื่อสารความเร็วสูงทางแสงได้มากมายเช่น อุปกรณ์สวิตช์ กรองสัญญาณ และมัลติเพลกซ์ เป็นต้น ตัวอย่างนี้เป็นเพียงบางตัวอย่างของการประยุกต์ใช้โซลิตอนเท่านั้น ซึ่งคาดว่าในอนาคตอันใกล้โซลิตอนคงเข้ามามีบทบาทอย่างมากแน่นอน และในส่วนของข้อมูลนั้นผู้เขียนยังอยู่ในขั้นตอนการทำโปรแกรมคอมพิวเตอร์ สำหรับการคำนวณของใยแก้วนำแสงโดยใช้ NLS ซึ่งจะได้นำผลของข้อมูลสำหรับออกแบบอุปกรณ์เสนอต่อไป

## เอกสารอ้างอิง

- [1] ปรีชา ยุพาพิน, พรทิพย์ อนุรัตน์, ลลิตา มั่นยานนท์ และ ภักกนิชิตสกุล, "การสื่อสารโซลิตอน", วารสารวิทยาศาสตร์ มข., ปีที่ 24 ฉบับที่ 2, 2539, หน้า 91-99.
- [2] Hasagawa, A. and Kodama, Y., Soliton in Optical Communications, 1995, Clarendon Press, Oxford, pp.7-42.
- [3] Masanori, K., Optical Waveguide Theory by the Finite Element Method, Tokyo, KTK Scientific Publishers, 1992.
- [4] Mayteevarunyoo, T., "Soliton in Optical Fibers", *J. Engineering Transaction*, Mahanakorn University of Technology, vol. 3, 2000, pp. 45-52.
- [5] Newell, A.C. and Moloney, J.V., Nonlinear Optics. ATIMS, California, 1992.
- [6] Scheuer J. and Orenstein, M., "Interaction and Switching of Spatial Soliton Pairs in the Vicinity of a Nonlinear Interface", *Opt. Lett.*, vol.24, no. 23, pp.1735-1737.





**Mr. Chainarong Taepanich** He was born in 1973 in Chonburi, Thailand. He received B.Sc. in Physics, from Burapha University, Chonburi, in 1996. He is a lecturer in the Department of Physics, Faculty of Science, Burapha University, Chonburi, in 1999. He is a M.Sc.

student in the Department of Applied Physics, King Mongkut's Institute of Technology Ladkrabang. His research interests are soliton communication and optical signal processing. He is a member of Thai Institute of Physics (**TIP**).



**Assoc. Professor Dr. Preecha P. Yupapin**

He was born in 1958 in Bangplama District, Suphanburi Province, Thailand. He received B.Ed. in Physics, M.Sc. in Mathematical Physics, and Ph.D. in Electrical Engineering from Srinakharinwirote University, Bangkok, Mahidol University, Bangkok and City

University, London in 1981, 1983 and 1993 respectively. He was a post doctoral research fellowship at the City University in 1995, researching in fiber optic devices for sensor and communication use. He is an associate professor in applied physics in the Light-wave Technology Research Center (**LTRC**), Department of Applied Physics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. He has authored/co-authored either research journals of 100 papers which have been published worldwide, he has also written 5 textbooks. His research interests are either experimental or theoretical works in fiber optic devices for sensors and communication, wavelength division multiplexing (WDM) and signal processing, optical soliton and communication, and optical metrology. Dr. Yupapin is a member of Thai Institute of Physics (**TIP**) and Optical Society of America (**OSA**).

# Fast Fourier Transform Photoreflectance

Jiti Nukeaw

Department of Applied Physics, King Mongkut's Institute of Technology Ladkrabang,  
Bangkok 10520, Thailand

E-mail: jiti@physics02.sci.kmitl.ac.th

**Abstract:** The electric fields in semiconductor were investigated by using fast Fourier transform photoreflectance (FFT PR) techniques. The Aspnes derivative function was introduced to analyze the PR spectra as a function of photon energy. Using the Airy function at several different electric fields in GaAs, the Franz-Keldysh Oscillations (FKO) above the band-gap energy ( $E_g$ ) of PR spectra was generated. It was confirmed that FKO period increased with increasing electric field. The application of FFT to the PR spectra was demonstrated to be able to calculate the electric field from FKO period under uniform electric field.

**บทคัดย่อ:** สนามไฟฟ้าที่เกิดขึ้นภายในสารกึ่งตัวนำ สามารถตรวจสอบได้โดยเทคนิคฟูริเยร์ทรานส์ฟอร์มโฟโตรีเฟลกแตนซ์ ฟังก์ชันของAspnes derivative ถูกนำมาใช้ในการวิเคราะห์สเปกตรัมของโฟโตรีเฟลกแตนซ์กับพลังงานโฟตอน การแกว่งของ Franz-Keldysh ปรากฏในสเปกตรัมของโฟโตรีเฟลกแตนซ์เหนือค่าของแถบพลังงานต้องห้ามของสารกึ่งตัวนำแกลเลียมอาร์เซไนด์ได้ถูกสร้างจำลองขึ้นโดยใช้ฟังก์ชัน Airy ที่ค่าสนามไฟฟ้าต่างๆ คาบการแกว่งของสเปกตรัมมีค่าเพิ่มขึ้นเมื่อสนามไฟฟ้ามีค่าเพิ่มขึ้น การประยุกต์ใช้เทคนิคฟูริเยร์ทรานส์ฟอร์มกับสเปกตรัมของโฟโตรีเฟลกแตนซ์ถูกแสดงให้เห็นว่า สามารถหาค่าสนามไฟฟ้าได้จากคาบการแกว่งของ Franz-Keldysh ภายใต้สนามไฟฟ้าที่เกิดขึ้น

## 1. Introduction

With the development and application of high-quality thin-film growth techniques such as molecular-beam epitaxy (MBE) and organometallic vapor phase epitaxy (OMVPE), groups of new materials such as semiconductor heterostructures including superlattices (SLs), single quantum wells (SQWs), and multiple quantum wells (MQWs) have been produced because of their new physical properties and device applications. A variety of characterization methods including photoluminescence (PL), photoluminescence excitation (PLE) spectroscopy, modulation spectroscopy, Raman and resonant Raman spectroscopy, absorption spectroscopy, transmission electron microscopy (TEM), Hall measurements, etc., have been applied to exploit various physical information from those structures, but most of the methods mentioned above must work under specific conditions such as low temperature, e.g., PL, PLE, or special sample preparation, e.g., TEM, etc. In experiments, the most popular methods should be simple as well as informative. This aim is satisfied in photoreflectance (PR), one of the powerful contactless modulation spectroscopy suggested by Wang, Albers, and Bleil in 1967 [1]. It was widely used in extraction of semiconductor band-structure parameters [2] in the 1970s.

Since the first report on the application of the PR technique to semiconductor heterostructure studies by Glembocki et al [3] in 1985, hundreds of papers have been published on

this subject. PR is contactless, requires no special mounting of the sample, can be performed in a variety of transparent ambients and is sensitive to surface and interface electric fields. It is well known that at high electric fields Franz-Keldysh oscillation (FKO) appears in the PR spectra above the band-gap energy. The period of oscillations is directly related to the electric field at the surface or interface. Another more important feature of this technique is that even at room temperature, it can still provide as much information as other methods, i.e., PL, PLE, etc., at very low temperature [3]. For example, by using PR at 300 K, it is possible to determine the interband transition energies in semiconductor heterostructures (SLs, SQWs and MQWs) with an accuracy within a few meV. By analyzing the PR spectra, one can obtain even more information about any interested transition and can obtain even more information about the electric fields at surface or interface of semiconductor heterostructures.

In this paper, the application of fast Fourier transform (FFT) to the PR spectra is demonstrated to be able to calculate the electric field from FKO.

## 2. Fundamental PR Line shapes

The representation of the complex dielectric function  $\epsilon$  of semiconductors and semiconductor heterostructures in a uniform electric field as a low-field perturbation expansion has proved to be particularly useful in the modulation

spectroscopy such as photoreflectance and electoreflectance [4]. The low-field expression for  $\Delta\epsilon$ , the field-induced change in  $\epsilon$ , is related to the third derivative of the unperturbed dielectric function and is sufficiently simple to allow certain critical-point characterization to be obtained directly from low-field PR line shapes [4-7]. To consider the perturbation on the reflectivity  $R$ , when the electric field is applied to a sample, the reflectivity is changed through the change of the dielectric constants  $\epsilon_1$  and  $\epsilon_2$ . The relative change of the reflectivity  $dR/R$  can be written as [6-7]

$$\frac{dR}{R} = \alpha(\epsilon_1, \epsilon_2) \Delta\epsilon_1 + \beta(\epsilon_1, \epsilon_2) \Delta\epsilon_2, \quad (1)$$

where  $\alpha$  and  $\beta$  are the Seraphin's coefficients;  $\epsilon_1$ ,  $\epsilon_2$  and  $\Delta\epsilon_1$ ,  $\Delta\epsilon_2$  are the real and imaginary parts of the dielectric constant, respectively, and its variation, i.e.,  $\epsilon = \epsilon_1 + i\epsilon_2$  and  $\Delta\epsilon = \Delta\epsilon_1 + i\Delta\epsilon_2$ . The real and imaginary parts of the dielectric constants can be related with each other by Kramers-Kronig relations.

By first-order perturbation calculation, Aspnes [8] derived a result that  $dR/R$  is proportional to the third derivative of the unperturbed dielectric function, which depends on the joint density of states of a material, in the limit of low-field modulation. Here, the low-field limit means that the electro-optic  $\eta\Omega \ll \Gamma$ , where  $\Gamma$  is the broadening parameter of the optical feature, and  $\eta\Omega$  is given by

$$(\eta\Omega)^3 = e^2 \eta^2 F^2 / 2\mu \quad (2)$$

where  $F$  is the electric field in the sample which could be externally applied (or built-in) electric field. When  $\eta\Omega \gg \Gamma$ , the situation is in a high-field regime, where in bulk materials it appears as FKO, which are described in section 3.

In the low electric field limit, i.e., the PR spectra as a function of photon energy can be analyzed using the Aspnes derivative function [7],

$$\frac{dR}{R} = \text{Re} \sum_{j=1}^p C_j e^{i\theta_j} (E - E_{gi} + i\Gamma_j)^{-n}, \quad (3)$$

Here,  $R$  is the reflectance,  $dR$  is the induced change in the reflectance by modulation light,  $E$  is the photon energy,  $p$  is the total number of spectral structures to be fitted,  $E_{gi}$ ,  $\Gamma_j$ ,  $C_j$  and  $\theta_j$  are transition energy, broadening parameter, amplitude and phase, respectively, of the feature corresponding to the  $j^{\text{th}}$  critical point. The parameter  $n$  is a factor used to specify the critical point dimension.

### 3. Franz-Keldysh Oscillations in PR

When  $\eta\Omega \gg \Gamma$ , the situation is in a high-field regime, where in bulk materials it appears as FKO. In the high-field limit, the FKO in PR spectrum is proportional to the product of Airy function such as  $\text{Ai}(x)$ ,  $\text{Bi}(x)$  and their derivatives such as  $\text{Ai}'(x)$ ,  $\text{Bi}'(x)$ . The function

describing the FKO in PR spectrum taking into account the mixing of both light and heavy hole signals [8]:

$$\frac{dR}{R} = \alpha_{lh} G[(E_g - E)/\eta\Omega] + \alpha_{hh} G[(E_g - E)/\eta\Omega], \quad (4)$$

where  $G(x)$  is given by the Airy functions,  $G(x) = \text{Ai}'(x)\text{Bi}'(x) - x\text{Ai}(x)\text{Bi}(x)$ ,  $(\eta\Omega)^3 = e^2 \eta^2 F^2 / 2\mu_{ih}$ , ( $i=l$  or  $h$ ). Here,  $E$ ,  $E_g$ , and  $F$  are the photon energy, the band-gap energy, and the electric field strength, respectively.  $\mu_{l(h)h}$  is the reduced light (heavy) hole mass:  $\mu_{l(h)h}^{-1} = m_e^{-1} + m_{l(h)h}^{-1}$ . The values for the effective electron and hole masses used here are in units of  $m_0$ . The  $\alpha_{lh}$  and  $\alpha_{hh}$  are coefficients containing the transition oscillation strength of the light and heavy hole.

### 4. Simulation with Airy function

The Airy function in Eq. (4) was used to simulate the FKO in the PR spectrum of a GaAs sample. The Airy function program was written by using C language run on UNIX systems such as Linux UNIX and FreeBSD UNIX. The FKOs of various electric fields from 20 to 80 kV/cm with a fixed  $\alpha_{lh}$  and  $\alpha_{hh}$  are shown in Figure 1. The periods of FKOs increase clearly with electric fields. When the electric fields are fixed at 40 kV/cm, the FKOs in PR spectra depend on the  $\alpha_{lh}$  and  $\alpha_{hh}$ , as shown in Figure 2. The amplitudes of FKOs increase with  $\alpha_{lh}$  and  $\alpha_{hh}$ , while the periods have not change in the PR spectra.

### 5. Fast Fourier transform PR

Fast Fourier transform (FFT) is applied to the PR spectra in the energy ( $E$ ) region higher than the band-gap energy  $E_g$  to obtain the FKO period and then the electric field in the sample. The horizontal-axis variable is transformed from  $\epsilon = (E - E_g)^{3/2}$  to the inverse of  $\epsilon$  as  $t$ , while the vertical-axis variable is from  $dR/R$  to  $G(t)$  that can be written as [9]

$$G(t) = \int (dR/R) \exp(-i2\pi\epsilon t) d\epsilon. \quad (5)$$

The main peak ( $t_0$ ) evaluated from the Fourier transform is related to the electric field by

$$t_0 = (2/3\pi)(2\mu)^{1/2}(1/e\eta F), \quad (6)$$

where  $\mu$  is the reduced effective mass and  $F$  is the electric field [10]. The FFT PR program was written by using C language run on UNIX systems such as Linux UNIX and FreeBSD UNIX.

As seen in Figure 3, the transformed results of the simulated FKOs in PR spectra shown in Fig.1 exhibit clearly two peaks. High- $t$  peak corresponds to the FKOs of heavy hole, while lower- $t$  peak to the FKOs of light hole. The values of the electric fields of light and heavy holes calculated from the peak position are, of course, the same as the electric field values of simulated FKOs in the section 4.

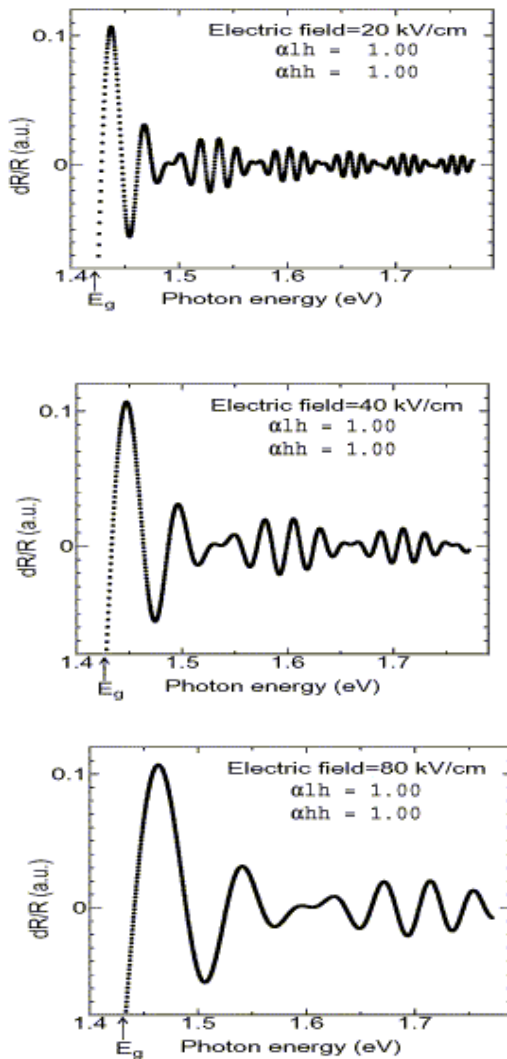


Figure 1. The FKOs due to various electric fields from 20 to 80 kV/cm with Fixed  $\alpha_{lh}$  and  $\alpha_{hh}$  are simulated by using the Airy function.

## 6. Discussions

J. Nukeaw *et al.* [11] have reported the electric fields at the surface and interface of doped GaAs/Si-GaAs structures grown by MBE. The electric fields were investigated using room-temperature PR spectroscopy. In the original PR spectra measured using a He-Ne laser as the pump light that had a modulation effect through out the doped layer, two different FKO periods could not be distinguished from each other. The application of FFT to the PR spectra can be used to calculate the electric field for each of the two periods involved in the original PR spectra. The transformed spectra clearly exhibited two peaks. By using an Ar laser as the pump light, the FFT-PR spectra showed only one peak. Because the Ar laser light was absorbed near the surface region. The resultant surface electric fields increased with increasing carrier concentration, while the change of the interface electric field was small. The behavior agreed qualitatively with the results of the model calculation, although there was a small discrepancy at

higher carrier concentrations. The discrepancy was explained the field strength derived from PR measurement was not their maximum, but the average value for high electric fields.

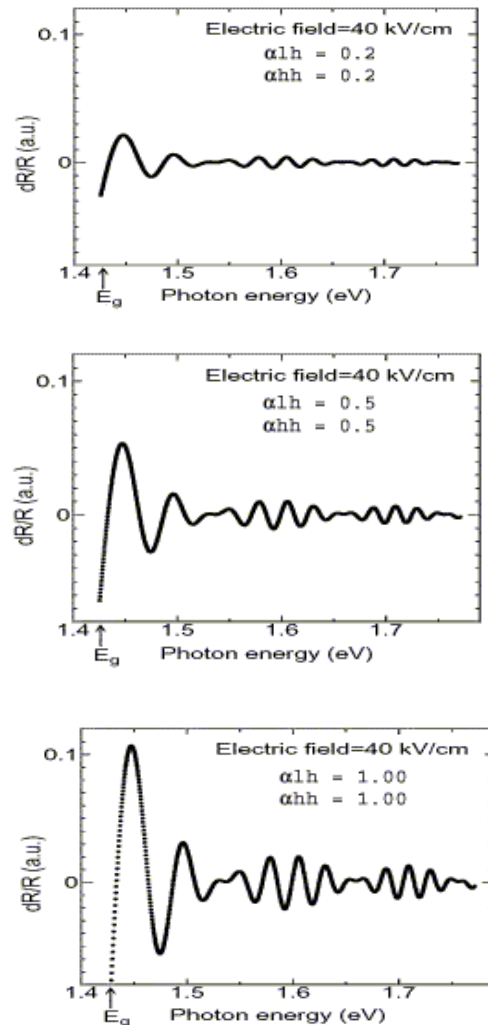


Figure 2. The FKOs for various  $\alpha_{lh}$  and  $\alpha_{hh}$  with a fixed electric field are simulated by using the Airy function.

J. Nukeaw *et al.* [12] also reported the electric fields of InP Er delta-doped grown by OMVPE. The electric fields were systematically investigated by room-temperature PR spectroscopy. The PR spectra were characterized by FKO above the band-gap energy due to an internal electric field in the epitaxial layer. FFT was applied to the PR spectra to calculate strength of the electric field from the FKO. The resultant electric field decreased with increasing Er-exposure duration. The dependence of the electric field on the Er-exposure duration was interpreted by change in size of the ErP clusters with the Er-exposure duration. The dependence of the electric field on the cap-layer was explained by the Er atoms easily reach to the surface with the thin cap-layer and react with O atoms in air to form  $\text{Er}_2\text{O}_3$  after growth, resulting in increase of electric field.

Thus, FFT is successfully applied to the PR spectra to calculate the electric field from the FKO.

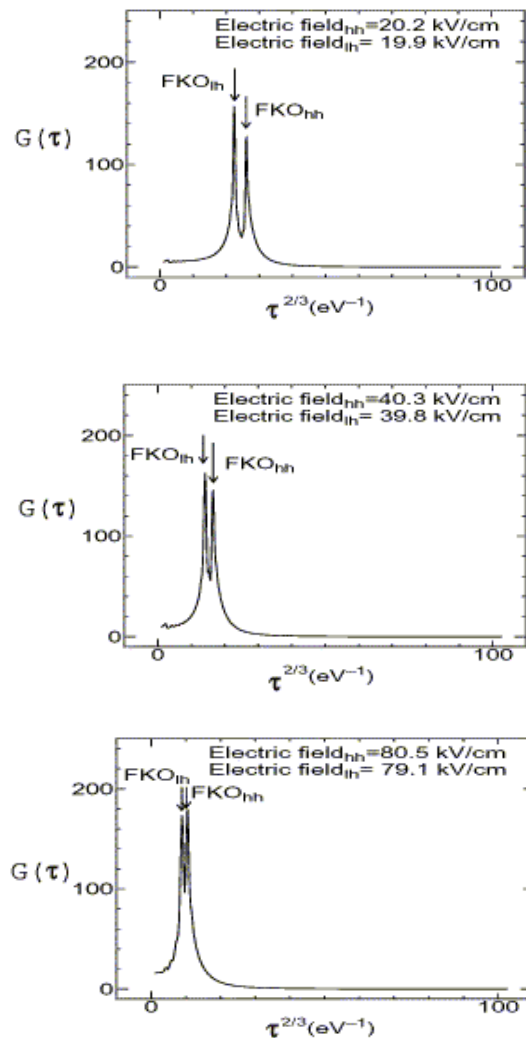


Figure 3. Transformed results of the PR spectra shown in Figure 1. From the peak positions the electric fields were obtained using eq. 6.

## 7. Conclusions

The FKO was generated using the Airy function at several different electric fields in GaAs. It was confirmed that FKO period increases with electric field. It was demonstrated that the application of FFT to the PR spectra can be used to calculate the electric field from the FKO under uniform electric fields.

## References

- [1] E. I. Wang, W. A. Albers, and C. E. Bleil, II-VI Semiconducting Compounds, edited by D. C. Thomas (Benjamin, New York, 1967), p. 136.
- [2] M. Cardona, Advances in Solid State Physics, Vol. X (Pergamon, Vieweg, 1970), p.125.
- [3] O. J. Glembocki, B. V. Shanabrook, N. Bottka, W. T. Beard, and J. Comas, *Appl. Phys. Lett.* **46**,976 (1985).

- [4] D. E. Aspnes and J. E. Rowe, *Solid State Commun.* **8**, 1145 (1970).
- [5] D. E. Aspnes and J. E. Rowe, *Phys. Rev.* **B5**, 4022 (1972).
- [6] J. E. Rowe, D. E. Aspnes, *Phys. Rev. Lett.* **25**, 162 (1970).
- [7] D. E. Aspnes, *Surf. Sci.* **37**, 418 (1973).
- [8] D. E. Aspnes, *Phys. Rev.* **B12**, 2297 (1975).
- [9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1994) p.496.
- [10] D. P. Wang and C. T. Chen, *Appl. Phys. Lett.* **67**, 2069 (1995).
- [11] J. Nukeaw, Y. Fujiwara, and Y. Takeda, *Jpn. J. Appl. Phys.* **36**, 7019 (1997).
- [12] J. Nukeaw, N. Matsubara, Y. Fujiwara, and Y. Takeda, *Appl. Surf. Sci.* **117/118**, 776 (1997).



**Assist. Prof. Dr. Jiti Nukeaw**

He received B.Ed. (Physics) from Srinakarinwirot University, Songkla in 1983 and M.S. (Physics) from Chiangmai University in 1989 and D.Eng. (Material Sciences and Engineering) from Nagoya University in 1998. At present, his working as lecture at Department of Applied Physics, KMITL. His current research interests include Quantum Well Devices, Semiconductor Physics, and Optical Characterization Tools for Semiconductor.

# Optimisation of Switched Mode Power Supply Feedback Compensation Using HSPICE

Chanin Bunlaksananusorn

Department of Control Engineering, Faculty of Engineering, KMITL,  
Chalongkrung Rd., Ladkrabang, Bangkok 10520

**ABSTRACT** – This paper proposes an optimisation technique for the design of switched mode power supply feedback control, implemented using HSPICE and its incorporated circuit optimiser. The proposed optimisation technique overcomes the drawbacks inherent in available design methods and eases the task of designing the error amplifier compensation circuit. A design example is presented with the results confirmed by experiment. It is expected that the proposed optimisation technique can be applied to design the control loop of power supplies for modern high-speed electronic systems that usually demand excellent dynamic performance from their power supplies.

**KEY WORDS** – Switched Mode Power Supply, Optimisation, Feedback Control

**บทคัดย่อ** - บทความนี้นำเสนอ การนำเทคนิคการหาค่าที่เหมาะสม (optimisation) ไปช่วยออกแบบการควบคุมป้อนกลับของสวิทซ์โหมดเพาเวอร์ซัพพลายโดยใช้โปรแกรม HSPICE เทคนิคที่นำเสนอนี้ ช่วยให้การออกแบบวงจรชดเชยในการป้อนกลับง่ายขึ้น อีกทั้งยังแก้ไขจุดด้อยของวิธีการออกแบบแบบคลาสสิกที่ใช้กันอยู่ บทความนี้แสดง ตัวอย่างการออกแบบโดยใช้เทคนิคการหาค่าที่เหมาะสม และ เปรียบเทียบผลที่ได้ กับผลที่ได้จากวิธีการออกแบบแบบคลาสสิก ผลการทดลองแสดงให้เห็นว่า วิธีการออกแบบที่นำเสนอนี้ให้ผลที่ถูกต้อง

**คำสำคัญ** – สวิทซ์โหมดเพาเวอร์ซัพพลาย, การหาค่าที่เหมาะสม, การควบคุมแบบป้อนกลับ

## 1. Introduction

Switched Mode Power Supplies (SMPS) are a basic building block in all modern electronic systems ranging from the modern office equipment (e.g. personal computers, printers, fax machines, etc.) to highly sophisticated systems (e.g. radar systems, telecommunication systems, etc.). The increasing complexity of electronic systems has resulted in ever more demanding requirements being placed on their power supplies, particularly in terms of power density, efficiency, electromagnetic compatibility, and output voltage control. The trend towards lower operating voltage for integrated circuits has further increased the burden on power supply designers

The output voltage is specified both in terms of steady state voltage regulation, and also its response to transient conditions (e.g. application/rejection of load). The SMPS regulates the output voltage by comparing it with a reference voltage, and using feedback control to modulate the pulse width of the switching transistor(s). To design the SMPS to meet both steady state and dynamic requirements requires careful design of the control circuit, in particular the compensation circuit around the error amplifier in the feedback loop.

### 1.1 Frequency Response Design

The design of linear time-invariant Single Input Single Output (SISO) control systems is normally carried out in the frequency domain using classical control theory [1].

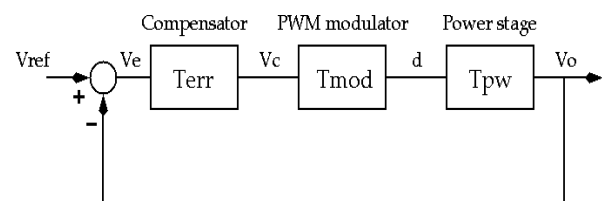


Figure 1. Basic SMPS closed-loop block diagram

Figure 1 shows the basic block diagram of a SMPS from the control system point of view, consisting of the power circuit, PWM modulator, and error amplifier with the compensation circuit, with transfer functions  $T_{pw}$ ,  $T_{mod}$ , and  $T_{err}$  respectively.

The design task is to define  $T_{err}$  to give the desired open-loop gain frequency characteristic, while not violating the stability criterion [2] (typically, a phase margin of at least  $45^\circ$  is required for a stable system). With the aid of a Bode plot, the design is carried out graphically in the frequency domain, simplifying the design process. However, the method has the following limitations:



1. The power supply is assumed to be linear with small-signal perturbations around the operating point; whereas in practice, a converter can be subjected to large disturbances, either from line or load changes.
2. The design is carried out in the frequency-domain, whereas the target specifications are time-domain parameters (voltage regulation, maximum deviation from nominal and recovery time under transient conditions). Therefore, it is common that the desired specifications are not met at the first attempt, with fine-tuning required, usually on a trial-and-error basis.
3. The effects of circuit/device parasitics are usually omitted to simplify the process.

## 1.2 Existing CAD Approaches

Circuit simulation programs such as SPICE and SABER are being used increasingly in SMPS design to validate accurately the initial design results. The SMPS design environment developed in [3] uses SPICE as a simulation platform and the compensation circuit design is carried out by repeatedly simulating the circuit and validating its performance, until a satisfactory result is met. Because a trial-and-error approach is used to adjust the circuit component values for each simulation run, it can be very time consuming, particularly when designing for stringent specifications.

MATSPICE, a compensation circuit design environment, has been proposed in [4,5]. The package incorporates MATLAB, PSPICE, and some C interfacing routines to implement the optimisation design environment. However, the authors mention that the execution time of the program is slow, owing to the considerable amount of data being transferred between MATLAB and SPICE for each optimisation iteration, preventing it from being an efficient design tool.

In this paper, optimisation using HSPICE [6] is proposed as a solution to overcome the drawbacks of the existing tools. Compared with the conventional frequency response design, the proposed method has the following advantages:

1. A non-linear large-signal model, instead of a small-signal model, is used; it hence yields a more accurate design result.
2. The design is carried out in the time-domain by iterative simulation and optimisation. Through this iterative process, automatic fine-tuning of the design is accomplished.
3. The effect of circuit parasitics can be incorporated into a simulation model, if needed.

## 2. Optimisation Design Using HSPICE

The optimisation design scheme in HSPICE (illustrated in Figure 2) is an iterative design environment embracing circuit simulation, performance evaluation, and design variable (compensation circuit component values) optimisation until either the desired specifications are attained, or no solution is found after a pre-defined

maximum number of iterations. In the input HSPICE netlist, the converter under design is modelled by the large signal state-space average model described in section 3. The average model is chosen for this application because it is capable of giving accurate results, while yielding fast simulation time, and is easily implemented for SPICE simulation.

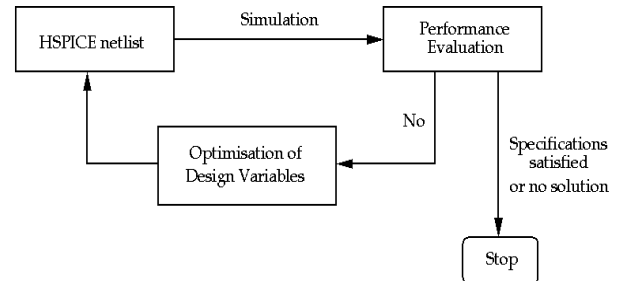


Figure 2. Optimisation design scheme in HSPICE

In addition to the circuit description, information concerning the optimisation task must also be stated in the netlist. The design variables are defined through the **.PARAM** statement in which the initial guess, minimum and maximum values are specified. The design objectives, the desired specifications (e.g. voltage overshoot, maximum voltage drop, settling time etc.) are stated through the **.MEASURE** statement. To perform optimisation in HSPICE, parameters in the analysis statements (**.TRAN**, **.AC** or **.DC**, with the keyname **OPTIMIZE**) must be given corresponding to those in the **.MODEL** statement. Also, the optimisation control options (e.g. the convergent criterion, and the algorithm used), are selected through the **.MODEL** statement. Detailed usage of these statements can be found in reference [6].

## 3. State-space Average Model

In this section, the large-signal state-space average model of a closed loop buck converter, which is used in HSPICE simulation/optimisation, is given.

The averaged model of the buck power circuit in figure 3 can be expressed in the state-space form as [7]:

$$\begin{bmatrix} \frac{di_L}{dt} \\ \frac{dv_C}{dt} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{L} \\ \frac{1}{C} & -\frac{1}{CR} \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix} + \begin{bmatrix} \frac{V_{in}}{L} \\ 0 \end{bmatrix} d \quad (1)$$

$$V_o = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix} \quad (2)$$

where  $i_L$  and  $v_C$  are state variables representing inductor current and capacitor voltage respectively.

The output of the error amplifier,  $V_e$ , is given by:

$$V_c = \frac{Z_1(s)}{Z_2(s)} (V_{ref} - V_o) \quad (3)$$

where  $V_{ref}$  is a referenced voltage and  $Z_1(s)/Z_2(s)$  is the error amplifier's gain due to the compensation circuit. It should be noted that the gain is expressed in the s-domain for ease of presentation. During the time-domain simulation, the gain will be automatically calculated by SPICE from the given compensation circuit component values.

The comparator compares  $V_c$  with a reference sawtooth voltage ( $V_{saw}$ ) to produce the duty ratio (d), which is represented by:

$$d = \frac{V_c}{V_m} = \frac{Z_1(s)(V_{ref} - V_o)}{Z_2(s)V_m} \quad (4)$$

where  $V_m$  is a peak of the sawtooth signal.

Equations (1), (2), and (4) thus form a model of the closed loop buck converter with output voltage control.

#### 4. Design Example

The application of HSPICE optimisation to the design of a SMPS error amplifier compensation circuit is demonstrated by the design example of a buck regulator depicted in figure 3.

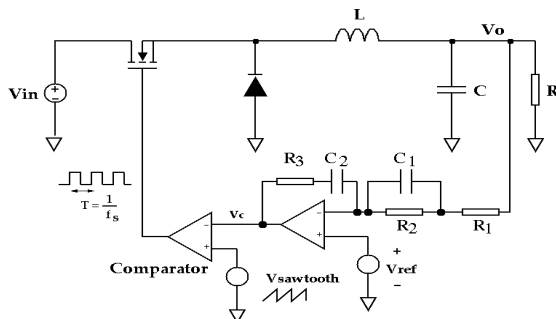
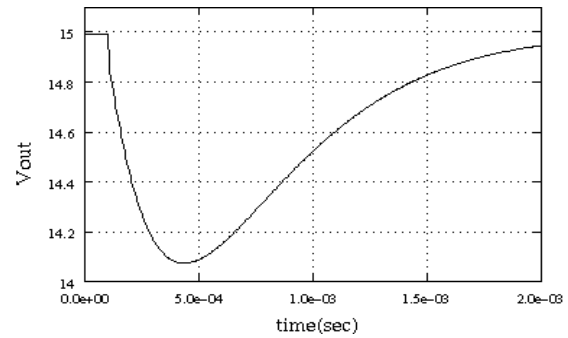


Figure 3. A buck switching regulator

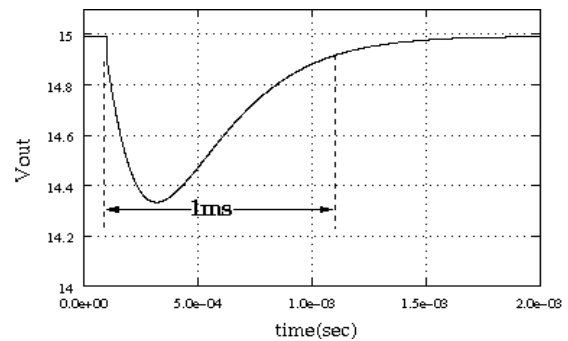
The buck converter has the following specifications:  $L=50 \mu H$ ,  $C=470 \mu F$ ,  $R=20 \Omega$ ,  $V_i = 35V$ ,  $V_o=15V$ ,  $I_o(min)=1A$ , and  $I_o(max)=5A$ , switching frequency  $f_s=100KHz$ . The two-pole, two-zero compensation circuit (comprises of  $R_1, R_2, R_3, C_1, C_2$ ) surrounding the error amplifier is designed to satisfy the following objectives:

1. The required regulation limits are  $\pm 0.5\%$  of nominal output voltage (i.e.  $15 \pm 0.075V$ )
2. A settling time less than 1ms is required, when the converter is subjected to a 3A step load change (60% of maximum load current).

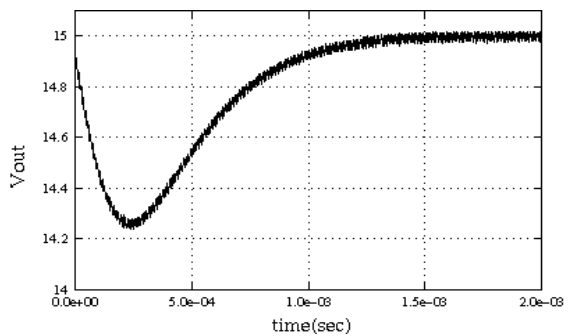
The compensation circuit was first designed using a classical frequency response method, resulting in the component values given in column 1 of Table 1. The associated transient response for a 3A load change is depicted in figure 4(a).



(a)



(b)



(c)

Figure 4:

(a) Output voltage transient response due to a 3A load change, before optimisation (simulated)

(b) Output voltage transient response due to a 3A load change, after optimisation (simulated).

(c) Experimental result after optimisation

It can be seen that a rather slow response is obtained and the design target is not yet met. From the result, it is apparent that this sluggish response is caused by the low loop-gain's bandwidth. The speed of the response can be enhanced to meet the

desired specification, by optimising the error amplifier's gain frequency characteristic (thus optimising the loop-gain's bandwidth).

It is unnecessary to select all five compensation circuit components as design variables, as the error amplifier can be tuned by varying  $R_3$  and  $C_2$ , while keeping the rest of

the components constant. Having too many design variables makes the optimisation more difficult and may cause a convergence problem. For these reasons, only  $R_3$  and  $C_2$  are selected as design variables, with the results from the frequency response design assigned as the initial guess values.

HSPICE is run on a SUN SPARC4 workstation. After a few iterations of simulation and optimisation, the optimised solution is found, with the results listed in column 2 of Table 1. It can be seen that the optimiser has met the specification by increasing  $R_3$  and decreasing  $C_2$ , thus increasing the error amplifier gain. Using the optimised component values in column 2 of Table 1, the output voltage transient response, when the regulator is subjected to a 3A step load change, is simulated and shown in figure 4(b). It can be seen that the optimised compensation circuit yields the output voltage response with the settling time of 1ms, as desired by the design objectives.

To confirm the validity of the design result, an experiment has been performed on the prototype buck regulator, whose specifications are identical to those used in the simulation and compensation circuit component values are the optimised values in column 2 of Table 1. The experimental transient response due to a 3A step load change is shown in figure 4(c). The experimental result behaves very closely to its simulated counterpart (figure 4 (b)), confirming the accuracy of the state-space averaged model in SMPS modelling. A slight discrepancy between the two results might be due to tolerances of various components used in the prototype regulator and circuit parasitics.

Table 1. Compensation circuit component values before and after optimisation

| Classical Design    | Optimisation Design |
|---------------------|---------------------|
| $R_1 = 470\Omega$   | $R_1 = 470\Omega$   |
| $R_2 = 4.70k\Omega$ | $R_2 = 4.70k\Omega$ |
| $R_3 = 16k\Omega$   | $R_3 = 27k\Omega$   |
| $C_1 = 0.068\mu F$  | $C_1 = 0.068\mu F$  |
| $C_2 = 0.022\mu F$  | $C_2 = 0.01\mu F$   |

## 5. Discussion

It should be noted that the results in column 2 of Table 1 found by the optimiser are just one set of solutions that satisfy the design objectives: it is likely that there will be many other sets of solutions which also satisfy these objectives, some of which may yield better transient responses (e.g. faster settling time with small voltage overshoot) than that in figures 4(b). The optimiser stops as soon as the original design objectives have been met: if the designer wishes to investigate whether further improvements in performance can be achieved, tighter design objectives should be entered, and optimisation re-run. The sets of solutions that satisfy the design objectives are called *non-inferior solutions* [8]. HSPICE does not necessarily produce the optimum solution.

## 6. Conclusion

The application of circuit optimisation feature in HSPICE to design SMPS compensation circuits has been described. Unlike traditional frequency response design, the proposed method carries out the design by means of iterative time-domain simulation and optimisation, automatically fine-tuning the design to meet the target specifications. The design example demonstrates the capability of this technique, with a significant improvement in the transient response such that the design target is met. It is expected that the proposed optimisation technique can be applied to design the control loop of power supplies for modern high-speed electronic systems that usually demand excellent dynamic performance from their power supplies.

## References

- [1] G. Ellis, "Control system design manual", Academic Press Inc., 1991.
- [2] L.H. Dixon, "Closing the feedback loop", Unitrode power supply design seminar, Unitrode Corporation, 1983.
- [3] D.E. Macpherson and E.A. Reddy; "Power supply control system design: an expert system approach", Proceedings IEE Power Electronics and Variable Speed Drives Conference (PEVD), 1998, pp. 586-591.
- [4] W.K. Tam, K.W. Sui, and Y.S. Lee, "A computer-aided design environment for switch-mode power supplies", Power Conversion Conference, 1993, pp. 496-50.
- [5] W.K. Tam, K.W. Sui, and Y.S. Lee, "Computer-aided design of feedback compensation networks for switching regulators", Proceedings IEEE Applied Power Electronics Conference and Exposition, 1994, pp. 1066-1071.
- [6] HSPICE user's manual, Meta-software Inc., 1992.
- [7] A.J. Forsyth and S.V. Molloy, "Modelling and control of DC-DC converters", Power Engineering Journal, October 1998, pp. 229-236.
- [8] P.J. Fleming and A.D. Pashkevich, "CACSD using a multiobjective optimisation approach", Proceeding IEE Conference Control'85, 1985, pp. 174-179.



**Chanin Bunlaksananusorn** received a PhD in electrical engineering from the University of Edinburgh in 1998. Since then, he has been with the department of control engineering, KMITL. His research interests are in the areas of switched mode power supplies.

# การสร้างเครื่องเชื่อมโลหะอิเล็กทรอนิกส์โดยใช้หลักการฟูลบริดจ์ พิสัยบลิวเอ็มแบบเลื่อนเฟสพร้อมสวิตช์ที่แรงดันศูนย์\*

## An Implementation of Electronic Welder by Using Full-Bridge ZVS Phase-shifted PWM Method

นายณัชพงศ์ หัตถิ, นายประสิทธิ์ จำปา และนายพงศ์พิชญ์ วิชาสุรณทล

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC)

หน่วยปฏิบัติการเทคโนโลยีคอมพิวเตอร์และระบบอัตโนมัติ (CTL)

ชั้น 5 อาคารสถาบันคั่นคว่ำและพัฒนาเทคโนโลยีการผลิตทางอุตสาหกรรม (RDIPT)

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพฯ 10900

**ABSTRACT:** The objective of this paper is to present an implementation of electronic welder by using full-bridge zero voltage switching phase-shift-PWM converter. The benefits of this method are not only size and weight reductions but also high efficiency and low EMI/RFI interference. The implemented prototype has rated power of 3 kilowatts (at 120 Amps 25 Vdc) and operated at 60 KHz frequency. Efficiencies at various load conditions and some kinds of waveforms measured have been reported.

**KEY WORDS :** Zero Voltage Switching, PWM Converter, Welder, Switched Mode Power Supply

**บทคัดย่อ:** บทความนี้นำเสนอแนวทางการสร้างเครื่องเชื่อมโลหะอิเล็กทรอนิกส์โดยใช้ คอนเวอร์เตอร์ชนิดฟูลบริดจ์พิสัยบลิวเอ็มแบบเลื่อนเฟสพร้อมสวิตช์ที่แรงดันศูนย์ ผลลัพธ์ที่ได้รับคือนอกจากจะทำให้เครื่องเชื่อมมีขนาดเล็กและน้ำหนักเบาแล้ว ยังให้ประสิทธิภาพสูง และ ลดสัญญาณรบกวนจากการสวิตช์ของตัวมอสเฟตกำลังด้วย ชุดต้นแบบที่สร้างขึ้นมีขนาดพิกัดกำลัง 3 กิโลวัตต์ (120 แอมป์ และ 25 โวลท์ ที่ฝั่งขา-ออก) ทำงานที่ความถี่ 60 กิโลเฮิรท์ส ผลการทดสอบประสิทธิภาพและรูปคลื่นบางส่วนได้ถูกรายงานในบทความนี้

**คำสำคัญ :** การสวิตช์ที่แรงดันศูนย์, คอนเวอร์เตอร์, เครื่องเชื่อมโลหะ, สวิตช์โหมดเพาเวอร์ซัพพลาย

### 1. บทนำ

ปัจจุบันเครื่องเชื่อมชนิดที่ให้เอาต์พุตเป็นไฟตรงมีการพัฒนาขึ้นอย่างรวดเร็วเนื่องจากสามารถนำไปประยุกต์ใช้งานได้หลากหลาย โดยเฉพาะการนำหลักการของ คอนเวอร์เตอร์แบบการสวิตช์ที่ความถี่สูง(High frequency switching converters) ในแบบต่างๆ มาประยุกต์ในการสร้างอย่างแพร่หลาย ทั้งในแง่ของผลิตภัณฑ์และงาน

วิจัย ข้อดีที่เห็นได้ชัดคือ การทำให้เครื่องเชื่อมมีขนาดเล็กและน้ำหนักเบา และยังให้ประสิทธิภาพที่สูง

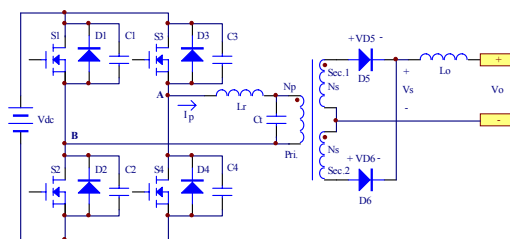
คอนเวอร์เตอร์แบบการสวิตช์ที่ความถี่สูงแบบหนึ่งคือคอนเวอร์เตอร์ชนิดพิสัยบลิวเอ็มแบบเลื่อนเฟส (Phase-shift PWM converter) เริ่มเป็นที่นิยมนำมาประยุกต์ใช้มากขึ้นเพราะมีคุณสมบัติที่น่าสนใจ คือ สามารถควบคุมได้ง่ายโดยการปรับความกว้างพัลส์เช่นเดียวกับชนิด

\* บทความนี้ตีพิมพ์ครั้งแรกในเอกสารประกอบการประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 23 หมวด A38 หน้า 149-152

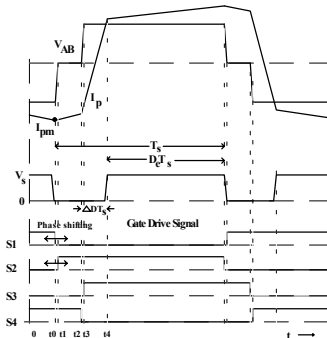
## 2. หลักการทำงานของเครื่องเชื่อมโลหะอิเล็กทรอนิกส์

ในสภาวะการเชื่อมปกติ กระแสอ้างอิง  $I_{ref}$  จะมีค่าตามค่ากระแสที่ตั้งไว้ แต่เมื่อฟังก์ชันออกเกิดสภาวะลัดวงจรหรือใกล้ลัดวงจรเนื่องจากสภาวะ

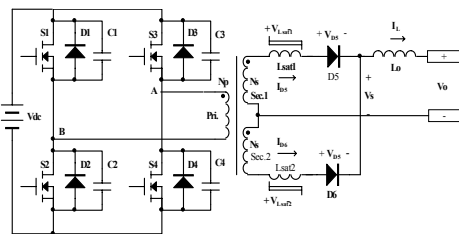
The diagram illustrates the current control system for a high-frequency converter. It features a feedback loop where the output voltage is divided by a 'Volume' resistor and compared with a reference voltage 'Vref'. The resulting error signal is summed with a 'Welding Current (No. sense)' signal at a summing junction. This signal then passes through a 'Current sense' resistor 'Lshort' and a 'Delay 1 Sec.' block. The output of the delay block is summed with a 'Boost Current' signal at another summing junction. The resulting signal is compared with a reference current 'Iref' at a third summing junction. The error signal from this comparison is processed by a 'PI Controller' and a 'Phase-locked PWM' block. The PWM output drives the 'To High-Frequency Converter'. The converter's output is monitored by an 'Over Temp. Detect' block. The system also includes an 'Output short circuit detect' block and a 'Boost Current' block.

$$\Delta D = \frac{2N L_r f_s I_L}{V_{dc}} \quad (1)$$


รูปที่ 3 วงจรพื้นฐาน ของ คอนเวอร์เตอร์ชนิดพีคดับลิเอ็ม  
แบบเลื่อนเฟส



รูปที่ 4 การทำงานของ คอนเวอร์เตอร์ชนิดพีดับบลิวเอ็มแบบเลื่อนเฟส



รูปที่ 5 วงจรคอนเวอร์เตอร์ที่ถูกปรับปรุงแก้ไข

เมื่อ  $f_s = 1/T_s$  เป็นความถี่การสวิตช์,  $N = N_s/N_p$  เป็นอัตราส่วนของรอบหม้อแปลง, และ  $I_L$  คือกระแสที่ไหลผ่านตัวเหนี่ยวนำ  $L_o$  ดังนั้นหากจะแก้ปัญหาการสวิตช์ที่แรงดันศูนย์ที่แคบโดยการเพิ่มค่าเหนี่ยวนำของ  $L_o$  นอกจากจะมี การสูญเสียขณะนำกระแส (Conduction loss) สูงขึ้นแล้วยังทำให้เกิด duty Loss เพิ่มสูงขึ้นอีกด้วย ข้อเสียอีกประการของคอนเวอร์เตอร์นี้คือ ไดโอดทางฝั่งทุติยภูมิมีการสูญเสียจากการสวิตช์สูงเช่นเดียวกับ คอนเวอร์เตอร์ชนิดพีดับบลิวเอ็มทั่วไป และการสูญเสียพลังงานอาจเพิ่มมากขึ้นเมื่อมีการเพิ่มวงจรสับเบอร์ เพื่อแก้ปัญหาแรงดันสไปค์ ตกรวมไดโอดและลดการเกิดสัญญาณรบกวน

### 3.2 การปรับปรุงแก้ไขวงจรคอนเวอร์เตอร์ชนิดฟูลบริดจ์พี-ดับบลิวเอ็มแบบเลื่อนเฟสพร้อมสวิตช์ที่แรงดันศูนย์

ปัญหาในหัวข้อ 3.1 สามารถแก้ไขโดยการใส่ตัวเหนี่ยวนำชนิดอิ่มตัวได้ (Saturable Inductor)  $L_{sat1}$  และ  $L_{sat2}$  เข้าไปอนุกรมกับหม้อแปลงทางด้านทุติยภูมิดังรูปที่ 5 [4] รูปสัญญาณบางส่วนของวงจรนี้แสดงให้เห็นดังรูปที่ 6 โดยมีการทำงานดังนี้ ช่วง  $t_0 < t_1$  เป็นช่วงฟรีวีลโดยแรงดัน  $V_{AB}$  เท่ากับศูนย์ กระแส  $I_L$  ไหลผ่าน  $D_6$  กับ  $L_{sat2}$  (ซึ่งอยู่ในสถานะอิ่มตัว) ที่  $t_1$   $S_4$  หยุดนำกระแส ตัวเก็บประจุ  $C_4$  และ  $C_3$  จะถูกชาร์จและดิสชาร์จตามลำดับ โดยที่  $I_p = NI_L$  ทำให้  $V_{AB}$  เริ่มเพิ่มสูงขึ้นขณะเดียวกัน เริ่มมีแรงดันตกคร่อม  $L_{sat1}$  และเริ่มมีกระแสไหลผ่าน  $I_{D5}$  และเมื่อ  $V_{AB}$  สูงขึ้นเท่ากับ  $V_{dc}$  กระแส  $I_p$  จะไหลผ่าน  $D_2$  และ  $D_3$  แทน ต่อจากนั้น  $S_3$  จึงถูกสั่งงานให้ปิดวงจรที่แรงดันศูนย์ ในช่วงเวลานี้ ( $t_1 < t < t_2$ ) แรงดัน  $V_{Lsat1} = 2NV_{dc}$  ดังนั้น กระแส

$I_{D5}$  จะค่อยเพิ่มขึ้นอย่างเป็นเชิงเส้นด้วยความชัน  $2NV_{dc}/L_i$  พร้อมๆกับ กระแส  $I_{D6}$  ลดลงด้วยความชัน  $-2NV_{dc}/L_i$  และ  $V_s = -NV_{dc}$  ด้วย จากนั้น ที่  $t_2$   $L_{sat1}$  เริ่มอิ่มตัว กระแส  $I_{D5}$  จึงเพิ่มขึ้นอย่างฉับพลันจนมีค่าเป็น  $I_L - I_c$  พร้อมๆกับนั้น  $L_{sat2}$  ก็เริ่มคลายการอิ่มตัว และ กระแสผ่าน  $I_{D6}$  ลดลงอย่างฉับพลันจนมีค่าเท่ากับ  $I_c$  ต่อมา (ช่วง  $t_2 < t_3$ ) แรงดันตกคร่อม  $L_{sat2}$  มีค่าเป็น  $-2NV_{dc}$  และกระแส  $I_{D6}$  จึงลดลงอย่างเป็นเชิงเส้นด้วยความชันเท่ากับ  $-2NV_{dc}/L_i$  ไปสู่ศูนย์ พร้อมๆกับกระแส  $I_{D5}$  เพิ่มขึ้นด้วยความชัน  $2NV_{dc}/L_i$  ไปสู่ค่า  $I_L$  หลังจากนี้  $I_{D6} = 0$  แล้ว (ที่  $t_3$ ) แรงดัน  $V_{D6}$  จะมีแรงดันไบแอสกลับเป็น  $-NV_{dc}$  ในช่วงเวลานี้ ไดโอด  $D_6$  จะทำการคืนตัวอย่างช้าๆเพราะถูกอนุกรมด้วยคาอิมพีแดนซ์ที่สูงของตัวเหนี่ยวนำ  $L_{sat2}$  ทำให้แรงดันสไปค์ที่เกิดขึ้นในคอนเวอร์เตอร์เดิมจึงถูกขจัดออกไป นอกจากนี้ค่า duty Loss สูงสุด  $\Delta D_{max}$  เป็นไปตามสมการต่อไปนี้

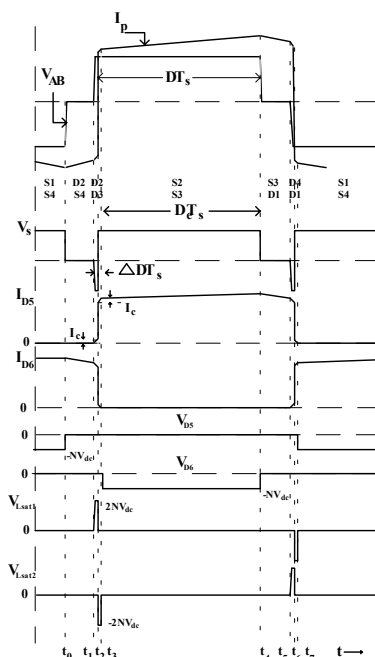
เมื่อ  $I_{LC}$  คือกระแสโหลดที่ทำให้  $L_{sat}$  อิ่มตัว,  $L_i$  คือค่าความเหนี่ยวนำ

$$\Delta D_{max} = \frac{L_i f_s I_{LC}}{NV_{dc}} \quad (2)$$

ของ  $L_{sat}$  ขณะไม่อิ่มตัว,เมื่อเปรียบเทียบกับ สมการ (1) จะเห็นว่าการใช้  $L_{sat}$  จะเป็นการจำกัดค่าของ  $\Delta D$  เมื่อกระแสโหลดเพิ่มขึ้น แรงดันขาออกที่โหลดสูงๆจึงสูงกว่ากรณีไม่ใส่  $L_{sat}$

จากที่กล่าวมาสรุปได้ว่า การใช้  $L_{sat}$  ในคอนเวอร์เตอร์ที่ทำการปรับปรุงแก้ไขนี้ นอกจากทำให้แรงดันขาออกเพิ่มสูงขึ้นแล้ว เนื่องจากค่าของการสะท้อนกระแสจากฝั่งทุติยภูมิยังทำให้การสวิตช์ที่แรงดันศูนย์ใน

ช่วงเวลานี้เป็นไปได้ง่ายขึ้นด้วย และยังทำให้การหยุดนำกระแสของไดโอดเรียงกระแสทางด้านขาออกที่กระแสเป็นศูนย์อีกด้วย



รูปที่ 6 ลักษณะสัญญาณของ วงจรในรูปที่ 5



### 3.3 การปรับปรุงในสภาวะการลัดวงจรทางด้านขาออกและโหลดต่ำ

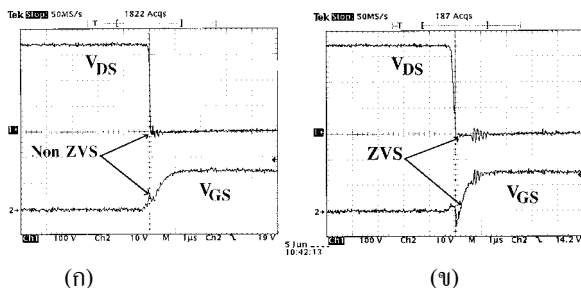
ตามปกติแล้ว มอสเฟตกำลังทั่วไปจะมีไดโอดภายในที่มีความเร็วไม่สูงมากนัก ดังนั้นการคืนตัวของไดโอดนี้จะใช้เวลานานากระแสของมอสเฟตเอง อย่างไรก็ตามในสภาวะโหลดต่ำๆ[2]มีกระแสไหลผ่านตัวมอสเฟตน้อยจึงเกิดแรงดันไบแอสกลับตกรวมไดโอดน้อยมาก จนอาจไม่พอที่จะทำการคืนตัวได้สมบูรณ์ ก่อนที่มอสเฟตกำลังอีกตัวที่อยู่ในกิ่งเดียวกันจะทำการนำกระแส ดังนั้นการลัดวงจรภายในกิ่งจึงอาจเกิดขึ้นได้ และในสภาวะเกิดการลัดวงจรที่ขาออกของคอนเวอร์เตอร์[1] ซึ่งอาจเกิดขึ้นได้ขณะเชื่อม มอสเฟตกำลังที่อยู่ในช่วงนำกระแสจะถูกบังคับให้เปิดวงจรอย่างรวดเร็ว เพื่อป้องกันความเสียหาย จนอาจไม่มีเวลาให้กับไดโอดภายในมากพอที่จะทำการคืนตัวได้สมบูรณ์ เช่นเดียวกับในสภาวะโหลดต่ำ

การแก้ปัญหาดังกล่าว สามารถทำได้โดยการ เพิ่มตัวเก็บประจุภายนอกมอสเฟตกำลังแต่ละตัว และปรับช่วงเวลา dead time ให้กว้างขึ้น การทำเช่นนี้ทำให้ไม่มีการสวิตช์ที่แรงดันศูนย์(Non-ZVS) ในสภาวะโหลดต่ำ ไดโอดจึงไม่มีการนำกระแส จึงทำให้ไม่เกิดการคืนตัวขึ้น ส่วนในสภาวะลัดวงจร การทำเช่นนี้ทำให้ความชัน (dv/dt) ที่ตกคร่อมมอสเฟตที่เริ่มหยุดนำกระแสลดลง และมีเวลาพอสำหรับการคืนตัวให้สมบูรณ์ของไดโอดภายในก่อนที่มอสเฟตกำลังในกิ่งเดียวกันจะนำกระแส การลัดวงจรภายในกิ่งจึงไม่เกิดขึ้น

### 4. การทดลองและผลการทดลอง

ต้นแบบเครื่องเชื่อมโลหะอิเล็กทรอนิกส์ได้ออกแบบไว้ที่ข้อกำหนดดังนี้ แรงดันขาเข้า  $220V_{rms} \pm 15\%$ , 50Hz, กำลังงานทางด้านขาออก (สูงสุด) 3KW, กระแสทางด้านขาออกสูงสุด 120A, ความถี่การสวิตช์ 60KHz

ในสภาวะโหลดต่ำๆ สัญญาณแรงดัน เกรน-ซอร์ส และ เกต-ซอร์ส มีลักษณะดังในรูปที่ 7(ก) ซึ่งแสดงให้เห็นการสวิตช์ที่แรงดันไม่เป็นศูนย์ เมื่อทำการเพิ่มโหลดมากขึ้น (ประมาณ 25% โหลดปกติ) การทำงานของเครื่องเชื่อมเริ่มเข้าสู่สภาวะสวิตช์แรงดันศูนย์ รูปที่ 7 (ข)

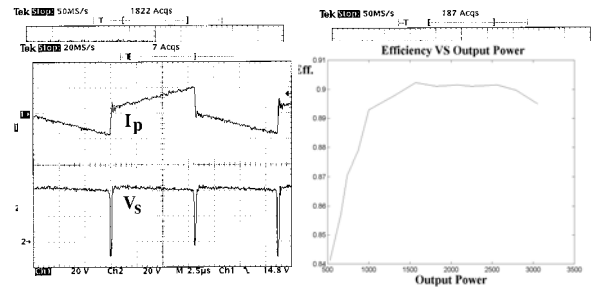


รูปที่ 7 เปรียบเทียบแรงดันระหว่าง  $V_{DS}$  กับ  $V_{GS}$

ก) กรณีที่มีการสวิตช์ที่แรงดัน  $V_{DS}$  ไม่เป็นศูนย์

ข) กรณีที่มีการสวิตช์ที่แรงดัน  $V_{DS}$  เป็นศูนย์

รูปที่ 8 แสดงให้เห็นความสามารถในการเพิ่มความกว้างพัลส์และความชันของกระแสขณะกลับขั้ว ของคอนเวอร์เตอร์ ที่ทำการปรับปรุง จากการเพิ่ม  $L_{sat}$  ประสิทธิภาพของชุดต้นแบบที่สร้างขึ้น แสดงในกราฟ รูปที่ 9



รูปที่ 8 แรงดัน  $V_s$  เทียบกับ  $I_p$   
(20A/ช่อง) ที่ Duty cycle สูงสุด

รูปที่ 9 ประสิทธิภาพ(Eff.)  
เทียบกับกำลังงานขาออก

### 5. สรุป

ธรรมชาติของเครื่องเชื่อมโลหะนั้นมีพิสัยโหลดที่กว้างมาก ตั้งแต่สภาวะไร้โหลดไปจนถึงลัดวงจร แม้ว่าวิธีการ พัดดับลิ่วเอ็มแบบเลื่อนเฟสพร้อมสวิตช์ที่แรงดันศูนย์ จะมีข้อดีหลายประการคือ มีการสูญเสียจากการสวิตช์ต่ำ สัญญาณรบกวนน้อย ไม่ต้องมีสับเบอร์เพิ่มเติม และง่ายต่อการควบคุม แต่ยังมีข้อเสียในการนำไปประยุกต์ใช้ในเครื่องเชื่อมโลหะอยู่บางประการอย่างเช่น ในกรณีที่โหลดเบาและสภาวะใกล้เคียงหรือลัดวงจร สามารถแก้ปัญหาโดยการปรับ dead time และเพิ่มตัวเก็บประจุคร่อมตัวมอสเฟตกำลัง ตามลำดับ สำหรับปัญหาย่านการสวิตช์ที่แรงดันศูนย์แคบ การหยุดนำกระแสของไดโอดเรียงกระแสทางด้านขาออกที่กระแสไม่เป็นศูนย์ และการเกิด duty loss สูง สามารถแก้ไขได้โดยเพิ่มตัวเหนี่ยวนำแบบอิมิตัวได้เข้าไปทางด้านทุติยภูมิของหม้อแปลง

### เอกสารอ้างอิง

- [1] H. Aigner, K. Dierberger, and D. Grafham, "Improving the Full-Bridge Phase-Shift ZVT Converter for Failure-Free Operation under Extreme Conditions in Welding and Similar Applications," IEEE IAC 1998, vol.2, pp. 1341-1348.
- [2] L. Saro, K. Dierberger, and R. Redl, "High - Voltage MOSFET Behavior in Soft-Switching Converters: Analysis and Reliability Improvements," IEEE INTELEC 1998, pp. 30-40.
- [3] M. M. Walters, and W. M. Polivka, "Extending the Range of Soft-Switching in Resonant -Transition

DC-DC Converters, " IEEE IN TELECOM 1992, pp. 343-350.

- [4] B.-H.Kwon, J.-H.Kim, and G.-Y.Jeong, "Full-bridge soft switching PWM converter with saturable inductors at the secondary side, " IEE Trans.Power Appl., Vol.146.No.1,pp.116-122,January 1999.



**ณัชพงศ์ หัตถิ** จบการศึกษาด้านวิศวกรรมศาสตรบัณฑิต จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อ พ.ศ. 2537 เริ่มเข้าทำงานใน ศูนย์เทคโนโลยี

อิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ เมื่อปี พ.ศ. 2538 ขอบเขตที่สนใจคือ การควบคุมทางอิเล็กทรอนิกส์กำลัง ระบบการแปลงผันรูปพลังงาน และการประยุกต์ใช้ไมโครคอนโทรลเลอร์



**ประสิทธิ์ จำปา** จบการศึกษาด้านวิศวกรรมศาสตรบัณฑิต จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อ พ.ศ. 2541 เริ่มเข้าทำงานใน ศูนย์เทคโนโลยี

อิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ เมื่อปีพ.ศ. 2537 ขอบเขตที่สนใจคือ การควบคุมทางอิเล็กทรอนิกส์กำลัง และการประยุกต์ใช้ไมโครคอนโทรลเลอร์



**พงศ์พิชญ์ วิภาสุรณทล** จบการศึกษาด้านวิศวกรรมศาสตรบัณฑิต เมื่อ พ.ศ. 2530 จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง หลังจากนั้นเข้าทำงาน

ตำแหน่ง วิศวกรไฟฟ้าใน บริษัท เบอเล็คเกอร์ จำกัด หลังจากจบการศึกษาด้านวิศวกรรมศาสตรมหาบัณฑิตจากสถาบันเดิม ในปีพ.ศ. 2537 ได้เข้ามาทำงานในตำแหน่งนักวิจัย ที่ห้องปฏิบัติการเทคโนโลยีคอมพิวเตอร์และระบบอัตโนมัติ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติจนถึงปัจจุบัน ขอบเขตที่สนใจคือ การประยุกต์ใช้ไมโครคอนโทรลเลอร์ในงานอุตสาหกรรม ระบบการแปลงผันรูปพลังงาน และระบบขับเคลื่อนต่างๆ ทางไฟฟ้า

## จาก ISO 9000 สู่ QS-9000 : กลยุทธ์แห่งความสำเร็จ

วราพจน์ ทัศนเมธี

*Senior Department Manager, Circuit Electronic Industries Public Co.,Ltd.*

ดร. สุวรรณ ตรงจิตวิกรัย

*Vice President, Circuit Electronic Industries Public Co.,Ltd.*

**ABSTRACT** In today globalized economic world, business strategic competition has become increasingly severe. There have been changes in terms of technology, international legislative and commercial regulations as well as standard systems all of which are aiming at highest level of customer satisfaction. On the other hand, these regulations and standard systems have proven to be trade barriers which put resource-scarce businesses in disadvantage position. As a result, small and medium sized businesses are now operating under increasingly higher and higher pressure and challenge.

This article outlines a case study of implementation of the QS-9000 quality system at Circuit Electronic Industries Public Co., Ltd. (CEI) <http://www.cei.co.th>, a subcontracting manufacturer and final tester of integrated circuit packaged products. Currently, the company has approximately 2,300 employees and a monthly production capacity of 40 million units. All the finished- goods products are exported. Since its first being certified for ISO9002 in 1995, CEI was certified for the second time in 1998. Furthermore, CEI also received QS-9000 certification in 1999. This article intends to share the experience in QS-9000 implementation after ISO-9002 system from the point of view of a company that has applied for QS-9000 registration which should be a useful reference for other organizations.

**KEY WORDS** -- QS-9000, ISO 9000, Quality System, IC Packaging

**บทคัดย่อ** ในเศรษฐกิจยุคโลกไร้พรมแดนทุกวันนี้, ซึ่งมีการมีเพิ่มการแข่งขันในเชิงกลยุทธ์ทางธุรกิจที่รุนแรงอย่างไม่มีที่สิ้นสุด ได้มีการเปลี่ยนแปลงทางเทคโนโลยี, ข้อกำหนดของกฎหมาย การค้า และ ระบบมาตรฐานต่างๆระหว่างชาติ ให้มีการตอบสนองความพึงพอใจของลูกค้าอย่างสูงสุด แต่ในอีกทางหนึ่งก็เป็นกำแพงทำให้เสียเปรียบทางธุรกิจสำหรับองค์กรที่ขาดแคลนทรัพยากร ซึ่งทั้งหมดนี้ได้เพิ่มแรงกดดันและท้าทายแก่องค์กรขนาดย่อมและขนาดกลางให้เพิ่มขึ้นอย่างมาก

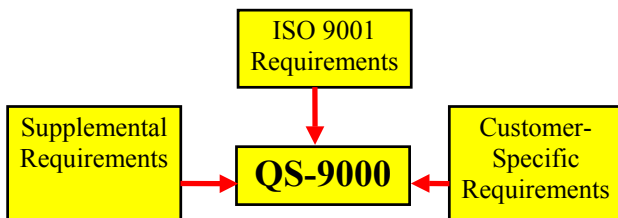
บทความนี้เป็นกรณีศึกษาของบริษัท Circuit Electronic Industries Public Co.,Ltd. (CEI) <http://www.cei.co.th> ในการจัดทำระบบคุณภาพ QS-9000 ซึ่งเป็นบริษัทรับช่วงผลิตและทดสอบวงจรครั้งสุดท้ายของผลิตภัณฑ์แผงวงจรรวม ( IC Packaging ) บริษัทได้มีพนักงานประมาณ 2,300 คน ปัจจุบันบริษัทมีกำลังการผลิตกว่า 40 ล้านชิ้นต่อเดือน ผลิตภัณฑ์ทั้งหมดถูกจำหน่ายไปยังต่างประเทศ บริษัทได้รับการรับรองระบบคุณภาพมาตรฐานอุตสาหกรรม ISO 9002 ครั้งแรกใน พ.ศ. 2538 ครั้งที่สองใน พ.ศ. 2541 และ ได้รับการรับรองระบบคุณภาพ QS-9000 ใน พ.ศ. 2542 ความตั้งใจในการจัดทำบทความนี้เพื่อเป็นการแลกเปลี่ยนประสบการณ์ของการจัดทำระบบคุณภาพ QS-9000 ต่อจากระบบ ISO 9002 ในมุมมองหนึ่งของผู้ขอการรับรอง เพื่อเป็นประโยชน์กับองค์กรอื่นๆ ที่กำลังพิจารณาจัดทำระบบคุณภาพ QS-9000

**คำสำคัญ** -- QS-9000, ISO 9000, ระบบคุณภาพ, IC packaging

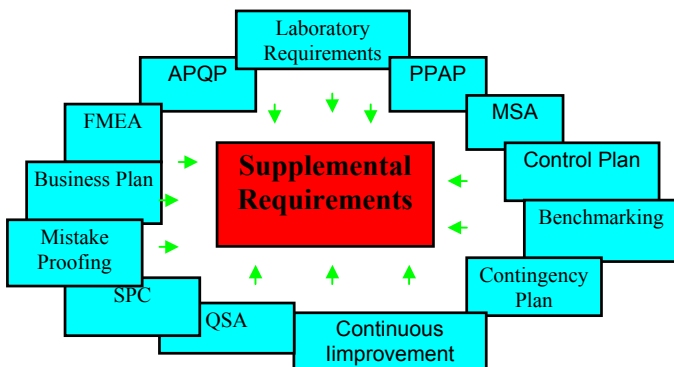
## 1. บทนำ

เป็นเวลากว่า 10 ปีมาแล้ว ที่อนุกรมมาตรฐาน ISO 9000 ได้เริ่มประกาศใช้อย่างเป็นทางการ และเป็นที่ยอมรับกันอย่างแพร่หลาย เพราะสามารถใช้ได้กับธุรกิจอุตสาหกรรม ตลอดจนผู้ให้บริการทุกประเภท ไม่ว่าจะเป็นธุรกิจขนาดเล็กหรือใหญ่ก็สามารถนำมามาตรฐานนี้ไปประยุกต์ใช้ตามความเหมาะสมของกิจการ

หลังจากนั้นก็ได้มีมาตรฐานต่าง ๆ ตามออกมาเพื่อพัฒนาและกำหนดลักษณะเฉพาะเจาะจง ตามประเภทอุตสาหกรรมมากขึ้น หนึ่งในจำนวนนั้นก็คือ QS-9000 ซึ่งเป็นตัวอย่างของการนำ ISO 9000 มาประยุกต์ใช้เพื่อเน้นการปรับปรุงอย่างต่อเนื่อง, การป้องกันของเสีย และกลยุทธ์เชิงรุกต่าง ๆ มาตรฐานจัดทำโดยกลุ่มยักษ์ใหญ่ของวงการรถยนต์ ประกอบด้วยบริษัท Ford, General Motors, Chrysler โดยเริ่มประกาศใช้ใน พ.ศ. 2537 จุดเด่นของ QS-9000 ก็คือ การนำเอาข้อกำหนดของมาตรฐาน ISO 9001 ทั้งหมด มาเป็นพื้นฐานในการจัดทำ โดยเพิ่มข้อกำหนด, เทคนิคต่างๆ และ ความต้องการในแต่ละลูกค้า (รูปที่ 1) เพื่อให้ข้อกำหนดเดิมของ ISO 9001 เฉพาะเจาะจงมากขึ้น (รูปที่ 2) นอกจากนั้นยังบังคับให้ผู้ส่งมอบผลิตภัณฑ์ให้กับกลุ่มโดยตรงจะต้องผ่านการรับรอง QS-9000 ตามเวลาที่กำหนดด้วยโดยมีการผ่อนผันตามกรณี แต่ก็มีบริษัทจำนวนไม่น้อยที่ไม่ได้อยู่ในข้อบังคับแต่นำระบบ QS-9000 มาใช้เช่นกัน



รูปที่ 1 แสดงองค์ประกอบหลักของ QS-9000

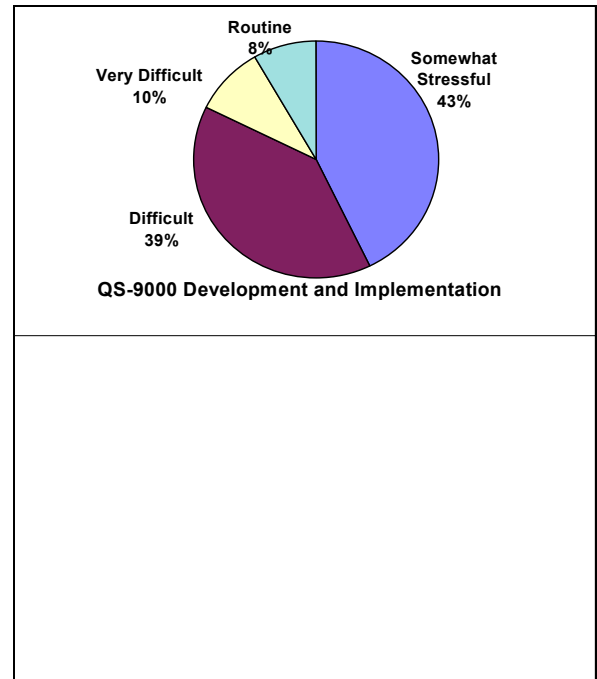


รูปที่ 2 แสดงตัวอย่างองค์ประกอบของความต้องการส่วนเพิ่มเติมที่สำคัญของ QS-9000

## 2. แนวคิด / วิธีการที่นำเสนอ QS-9000

### 2.1 ผลการสำรวจบริษัทที่จัดทำระบบ QS-9000

จากผลการสำรวจบริษัทที่จัดทำระบบ QS-9000 ของหน่วยงานแห่งหนึ่งพบว่าการดำเนินการจัดทำระบบนั้นมีความยากมาก, ยากและมีความเครียด 10%, 39% และ 43% ตามลำดับ นอกจากนั้นยังบอกว่ามีค่าใช้จ่ายสูงในการขอการรับรองมากกว่า 80% ของบริษัทที่สำรวจทั้งหมด (รูปที่ 3)



รูปที่ 3 แสดงผลการสำรวจบริษัทที่จัดทำระบบ QS-9000 ที่มาของข้อมูล [3]

### 2.2 การที่นำระบบ QS-9000 มาใช้

CEI ก็เป็นบริษัทหนึ่งที่ได้รับเอาระบบ QS-9000 มาใช้หลังจากการได้รับใบรับรอง ISO 9002 แล้ว โดยผู้บริหารระดับสูงได้มีความเห็นให้จัดทำระบบ QS-9000 ในการตัดสินใจนั้นเกิดขึ้นจากปัจจัยสิ่งแวดล้อมต่างๆ

ปัจจัยภายนอก : จากที่กล่าวมาแล้วข้างต้นว่าในโลกปัจจุบันได้มีการแข่งขันในเชิงธุรกิจอย่างรุนแรง ซึ่งลูกค้ามีความต้องการหลากหลายหนึ่งในจำนวนนี้ที่ลูกค้าตามถึงก็คือ ระบบคุณภาพ QS-9000 ถึงแม้ว่าลูกค้ายังไม่ได้บังคับให้บริษัทต้องจัดทำระบบคุณภาพ QS-9000 ก็ตามแต่ก็นับได้ว่าเสียงจากลูกค้าเป็นปัจจัยหนึ่ง ที่เป็นแรงจูงใจให้เกิดแรงผลักดันให้พิจารณาระบบ QS-9000 มาใช้

ปัจจัยภายใน : ความพึงพอใจของลูกค้า ในวันนี้ไม่ได้หมายความว่าลูกค้าจะพอใจในวันพรุ่งนี้หรือในอนาคต เพราะฉะนั้นการหยุดนิ่ง หรือ

พอใจในกิจกรรมและผลงานที่ทำอยู่จึงเท่ากับเป็นการถอยหลัง เนื่องจากลูกค้ามีโอกาสเลือกทำธุรกิจได้มากขึ้น เพื่อให้ได้สินค้าที่ตอบสนองความพึงพอใจมากที่สุด เพราะฉะนั้นการพัฒนาอย่างต่อเนื่องจึงเป็นกลยุทธ์สำคัญอย่างหนึ่งในการพัฒนาระบบคุณภาพในองค์กรทุก ๆ ด้านให้พร้อมอยู่เสมอ

ด้วยเหตุผลข้างต้น ทางผู้บริหารระดับสูงจึงได้เล็งเห็นความสำคัญของการนำระบบ QS-9000 มาใช้ ต่อจาก ISO 9002 ซึ่งการขอการรับรองต้องมีค่าใช้จ่ายต่างๆเกิดขึ้น เวลาของบุคลากรที่ต้องอุทิศให้กับการดำเนินการนอกเหนือจากงานประจำที่ทำอยู่ แม้ว่าผลตอบแทนที่บริษัทได้รับยังไม่ชัดเจนก็ตาม

### 3. การดำเนินการ

**3.1 การสรรหาผู้ดำเนินการ** ทางคณะผู้บริหารได้พิจารณาสรรหาวิธีการดำเนินการที่เหมาะสมหลายๆ ด้าน และในที่สุดก็ตัดสินใจใช้พนักงานภายในบริษัทโดยแต่งตั้งผู้ดำเนินการหนึ่งคนจากพนักงานภายในบริษัท โดยเลือกผู้ที่มีความเข้าใจระบบภายในองค์กรเป็นอย่างดีและมีทักษะในการบริหารงาน เพื่อเป็นแกนหลักในการจัดทำระบบ QS-9000 ให้บรรลุตามเป้าหมายอย่างมีประสิทธิภาพเพราะทางบริษัทได้มีระบบคุณภาพ ISO 9002 อยู่แล้ว ปัญหาที่ต้องพิจารณาคือทำอย่างไรที่จะบูรณาการระบบที่มีอยู่แล้วให้เข้ากันกับระบบ QS-9000 ได้อย่างกลมกลืน และมีประสิทธิภาพ

**3.2 การสำรวจและศึกษาเบื้องต้น** การทำ QS-9000 ไม่มี ทางลัดที่ทำให้ประสบความสำเร็จได้ นอกจากต้องศึกษาข้อกำหนดของ QS-9000 รวมทั้งคู่มือต่างๆ อีกหลายเล่ม เช่น APQP [4], PPAP [5], FMEA [6], MSA [7], SPC [8], QSA [9] เป็นต้น เพื่อนำเอาองค์ความรู้เหล่านั้นมาเป็นประโยชน์และสามารถประยุกต์ใช้เทคนิคต่าง ๆ ให้เหมาะสมกับธุรกิจ ดังนั้นจึงต้องศึกษาข้อกำหนดของ QS-9000 และสำรวจระบบเดิมที่มีอยู่ เพื่อจัดเตรียมแผนดำเนินงาน ( รูปที่ 4 ) ทรัพยากรต่างๆ ที่ต้องการใช้ เช่น ค่าใช้จ่าย, เวลาของพนักงานที่ต้องอุทิศมากน้อยขนาดไหน, การฝึกอบรม, และด้านอื่น ๆ เพื่อนำเสนอต่อผู้บริหารให้เข้าใจถึงแผนดำเนินพอสังเขย และขอรับการสนับสนุนจากหน่วยงานต่างๆ ให้บรรลุตามเป้าหมาย

| กิจกรรม                   | ระยะเวลาการดำเนินการ |   |   |   |   |   |   |   |   |    |    |
|---------------------------|----------------------|---|---|---|---|---|---|---|---|----|----|
| เดือนที่                  | 1                    | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| การสรรหาผู้ดำเนินการ      |                      |   |   |   |   |   |   |   |   |    |    |
| การสำรวจและศึกษาเบื้องต้น |                      |   |   |   |   |   |   |   |   |    |    |
| การสร้างทีมงาน            |                      |   |   |   |   |   |   |   |   |    |    |
| การจัดทำเอกสาร            |                      |   |   |   |   |   |   |   |   |    |    |
| เลือกหน่วยงานให้การรับรอง |                      |   |   |   |   |   |   |   |   |    |    |
| การทำฝึกอบรมและปฏิบัติ    |                      |   |   |   |   |   |   |   |   |    |    |
| การตรวจสอบภายในและแก้ไข   |                      |   |   |   |   |   |   |   |   |    |    |
| การขอการรับรอง            |                      |   |   |   |   |   |   |   |   |    |    |

รูปที่ 4 แสดงตัวอย่างของแผนดำเนินงาน

**3.3 การสร้างทีมงาน** ทีมงานประกอบด้วยตัวแทนจากแต่ละหน่วยงานโดยคำนึงถึงผู้ที่เกี่ยวข้องในข้อกำหนด QS-9000 เป็นหลัก ซึ่งส่วนใหญ่เป็นระดับผู้จัดการของแต่ละแผนก เพื่อให้เกิดแนวคิดหรือมุมมองที่หลากหลาย ในการสรรหาแนวทางที่ดีที่สุดในการนำข้อกำหนดไปประยุกต์ใช้ ในขั้นตอนนี้ได้มีการฝึกอบรมให้กับทีมงาน เพื่อให้เข้าใจในข้อกำหนดและเป้าหมายของบริษัท นอกจากนั้นยังได้กำหนดให้ทีมงานมีการประชุมกันทุก 2 สัปดาห์ เพื่อเปิดโอกาสให้มีการปรึกษาหารือในประเด็นปัญหาและปรับปรุงวิธีการที่จัดทำขึ้นให้สอดคล้องและไปในทิศทางเดียวกัน

**3.4 การจัดทำเอกสาร** หลังจากได้มีการทบทวนเบื้องต้นแล้วทีมงานพบว่าระบบเอกสารที่มีอยู่เดิมนั้นสอดคล้องกับ QS-9000 ในส่วนของพื้นฐานที่ได้จากการทำ ISO 9002 มาแล้ว จึงแก้ไขคู่มือคุณภาพให้สอดคล้องกับข้อกำหนดส่วนที่เพิ่มเติมก่อน ต่อจากนั้นจึงวางแผนงานในการจัดทำเอกสารระดับรองลงมา และกำหนดผู้ที่รับผิดชอบในแต่ละเรื่อง เอกสารบางฉบับอาจไม่ต้องแก้ไขเพิ่มเติม, บางฉบับเพียงแก้ไขเพิ่มเติมบางส่วน หรืออาจต้องจัดทำเอกสารขึ้นมาใหม่ขึ้นอยู่กับการข้อกำหนด โดยต้องมีการเชื่อมโยงซึ่งกันและกัน นอกจากนั้นในช่วงที่ได้มีการจัดให้มีการประชุมยังได้ให้ผู้รับผิดชอบในการจัดทำเอกสารมาบรรยายข้อกำหนดใหม่ที่จัดทำหรือแก้ไขให้แก่สมาชิกทุกคนรับไปปฏิบัติ

**3.5 การเลือกหน่วยงานให้การรับรอง** การตรวจประเมินไม่ได้ทำครั้งเดียว แต่ต้องผูกพันตลอดระยะเวลาของใบรับรอง และต่อไป ทางบริษัทได้สัมภาษณ์หน่วยงานให้การรับรองเพื่อคัดเลือกหน่วยงานที่ดีที่สุดในการตรวจประเมิน ทางทีมงานได้ใช้โอกาสนี้ซักถามข้อกำหนดต่างๆ เพื่อให้มั่นใจในทิศทางที่กำลังดำเนินการอยู่รวมทั้งในประเด็นปัญหาที่มีความคิดเห็นไม่ตรงกันในระหว่างการทำงาน ซึ่งพวกเขายินดีอธิบายข้อกำหนดให้แก่ทีมงาน และตอบข้อสงสัยต่างๆ ในการตัดสินใจเลือกหน่วยงานให้การรับรองทีมงานได้ใช้แนวทางดังนี้

- ความต้องการของลูกค้า
- คุณสมบัติและความรู้ในข้อกำหนด
- สัดส่วนของผู้ให้บริการ
- เงื่อนไขข้อผูกพันของการประเมิน
- Recognized Accreditation bodies[10]
- ค่าใช้จ่ายต่างๆ
- การให้บริการ

**3.6 การฝึกอบรมและปฏิบัติ** ทางบริษัทได้ใช้บุคลากรภายในบริษัทในการฝึกอบรมแก่พนักงานซึ่งเป็นการประหยัดค่าใช้จ่ายไปได้มาก

อีกทั้งบุคลากรภายในยังเข้าใจระบบที่ซับซ้อนภายในบริษัทได้ดีกว่าบุคลากรภายนอก และสามารถยกตัวอย่างตามสภาพความเป็นจริงในบริษัททำให้มองเห็นภาพได้ชัดเจน การฝึกอบรมกระทำอย่างต่อเนื่องตามความจำเป็นแก่พนักงานในแต่ละระดับ และได้้นำข้อกำหนดไปลงมือปฏิบัติตามเอกสารที่ประกาศใช้

**3.7 การตรวจสอบภายในและแก้ไข** ทางบริษัทได้ตรวจสอบข้อกำหนดภายใน โดยเน้นข้อกำหนดส่วนที่เพิ่มเติมจาก ISO 9002 ซึ่งใช้ทีมผู้ตรวจสอบภายในซึ่งมีประสบการณ์จากการทำ ISO 9002 และรายงานข้อบกพร่องต่าง ๆ เพื่อทำการแก้ไข ให้ผู้บริหารได้รับทราบตลอดเวลา การทำการตรวจสอบภายในของบริษัทได้กระทำทั้งที่เป็นทางการและตามแต่กรณี โดยทางทีมงานยังได้เชิญผู้บริหารแต่ละฝ่ายมาช่วยในการตรวจประเมินเป็นระยะๆ ซึ่งทำให้ฝ่ายบริหารได้มีโอกาสสัมผัสสภาพความเป็นจริง และเป็นการกระตุ้นให้พนักงานทุกระดับเห็นความสำคัญในการจัดทำระบบ QS-9000 ตัวอย่างแนวทางการตรวจสอบภายในให้ดูจากรูปที่ 5

| Section 1 : QS-9000 Requirements          | Quality Manual | Other Associated Documents                        |
|---|----------------|---|
| 4.1 Management responsibility             | 4.1            | QA-01-01  |
| 4.1.1 Quality policy                      | 4.1.1          | QA-01   |
| 4.1.2 Organization                        | 4.1.2          | Organization chart                                |
| 4.1.2.1 Responsibility and authority      | 4.1.2.1        | Job description                                   |
| 4.1.2.2 Resources                         | 4.1.2.2        | Training, PM, Calibration                         |
| 4.1.2.3 Management representative         | 4.1.2.3        | Appointed by President                            |
| 4.1.2.4 Organization interfaces           | 4.1.2.4        | Business plan, Customer perception, turn time etc |
| 4.1.2.5 Information to management         | 4.1.2.5        | PE-12-02, QA-16-01                                |
| 4.1.3 Management review                   | 4.1.3          | Minutes of meeting                                |
| 4.1.4 Business plan                       | 4.1.4          | MK-01-02  |
| 4.1.5 Analysis and use company level data | 4.1.5          | PPM, yield, cust complaints, delivery.            |
| 4.1.6 Customer satisfaction               | 4.1.6          | MK-19-01  |

#### Supplemental requirements

รูปที่ 5 แสดงตัวอย่างการตรวจสอบข้อกำหนดของ QS-9000 โดยเชื่อมโยงกับระบบภายใน

**3.8 การขอการรับรอง** หลังจากการใช้เวลาในการดำเนินการเป็นเวลา 1 ปี ทางบริษัทได้รับการตรวจสอบจากหน่วยงานแห่งหนึ่งที่มีชื่อเสียง ผลการตรวจสอบจริงพบข้อบกพร่องเล็กน้อย ซึ่งได้ถูกแก้ไขให้ลุล่วงได้หมดก่อนสรุปการประเมินในวันสุดท้าย บริษัทจึงได้การเสนอการรับรองระบบ QS-9000 ได้สำเร็จ

#### ปัญหาและอุปสรรค

- เนื่องจากในช่วงเวลาดำเนินการได้เกิดสภาวะวิกฤตทางเศรษฐกิจทั้งภูมิภาค บริษัทก็ได้รับผลกระทบนี้เช่นกัน ค่าใช้จ่ายในการดำเนินงานจึงต้องใช้อย่างมีประสิทธิภาพซึ่งก็ต้องทำความเข้าใจกับทีมงานและพนักงานให้รับทราบถึงความจำเป็น
- เนื่องจากข้อกำหนดบางอย่างไม่ได้ชี้ชัดว่าต้องทำอะไร ทำให้มีการตีความไปคนละแนวทางบ้าง, ไม่มั่นใจบ้าง, เข้าข้างตัวเองบ้าง ฉะนั้นจึงต้องการพิจารณาหาแนวทางที่เหมาะสมอย่างรอบคอบ โดยยึดหลักการของข้อกำหนด QS-9000 เป็นหลัก ประกอบกับการคำนึงถึงประโยชน์ต่อลูกค้าและบริษัทที่ได้รับ
- การจัดทำเอกสารพบว่า การเขียนข้อกำหนดซับซ้อนมากเกินไปจนทำให้ผู้ปฏิบัติไม่สามารถทำได้ถูกต้องและเกิดข้อบกพร่องได้ง่าย หรือเขียนข้อกำหนดน้อยเกินไปจนทำให้การปฏิบัติตามประสิทธิภาพ

#### กลยุทธ์ความสำเร็จ

จากการดำเนินการรับรองระบบคุณภาพ QS-9000 ข้างต้น อาจพอสรุปแนวทางที่จะประสบความสำเร็จในการขอการรับรองดังนี้

- ⇒ การมีผู้ดำเนินการในทิศทางที่ถูกต้องทำให้การนำเอาข้อกำหนดมาใช้ได้อย่างเหมาะสมกับการปฏิบัติในองค์กรนั้นๆ
- ⇒ การทำงานร่วมกับหน่วยงานให้การรับรองทำให้การปฏิบัติงานเกิดความมั่นใจมากขึ้นและขจัดความขัดแย้งลงได้
- ⇒ การมีตัวแทนฝ่ายบริหารที่ดีทำให้การสื่อสารงานระหว่างผู้บริหารระดับสูงและระดับกลางเป็นไปด้วยดี และกระตุ้นให้มีความมุ่งมั่นในการบรรลุเป้าหมาย
- ⇒ การให้ทุกคนมีส่วนร่วม ทุกคนมีความสำคัญในการทำงาน เพราะฉะนั้นจึงต้องมีการประสานงานกันในทิศทางที่ถูกต้อง เพื่อบรรลุวัตถุประสงค์หลักขององค์กร

#### 4. บทสรุป

นับได้ว่าบริษัทได้ประสบความสำเร็จในการขอการรับรอง QS-9000 ซึ่งใช้เวลา 1 ปี ตามแผนดำเนินงานและใช้งบประมาณอย่างมีประสิทธิภาพ จากการบริษัทได้ใช้บุคลากรภายในดำเนินการจึงทำให้เป็นองค์กรแห่งการเรียนรู้โดยใช้ QS-9000 เป็นศูนย์กลางดึงดูดศักยภาพของบุคลากรมาใช้อย่างมีคุณค่า และเป็นหัวใจสำคัญในการปฏิบัติตามข้อกำหนดและคงไว้ซึ่งระบบ เพราะพวกเขามีส่วนร่วมและเกิดความภูมิใจในการจัดระบบขึ้นมาด้วยตัวของเขาเอง การได้ใบรับรอง QS-9000 เป็นแค่การเริ่มต้นบนปัจจัยพื้นฐานที่บริษัทสามารถทำได้ภายใต้สิ่งแวดล้อมที่เปลี่ยนแปลงตลอดเวลาในการดำเนินธุรกิจ เช่น การ



เปลี่ยนแปลงของตลาด, การเปลี่ยนแปลงของลูกค้า, การเปลี่ยนแปลงของผลิตภัณฑ์, การพัฒนาของคู่แข่งและนโยบายการค้า เป็นต้น แต่อย่างน้อยที่สุดก็นับว่าเป็นการปรับปรุงระบบภายในบริษัทในทิศทางที่ดีขึ้นจากเทคนิคต่าง ๆ ตามข้อกำหนดของ QS-9000 ในก้าวต่อไปบริษัทได้เริ่มดำเนินการพัฒนาในด้านการจัดการสิ่งแวดล้อม หรือ ISO 14000 เพื่อเป็นประโยชน์ต่อสังคมในการรักษาสิ่งแวดล้อมซึ่งคาดว่าจะแล้วเสร็จในเร็ววัน

## เอกสารอ้างอิง

- [1] ISO 9001 : 1994 Quality Systems-Model for quality assurance in design, development, production, installation and servicing, International Organization for Standardization, Switzerland.
- [2] Quality System Requirements QS-9000 : Third Edition, Automotive Industry Action Group ( AIAG ), U.S.A.
- [3] AEC International Inc., "Survey : QS-9000 is expensive and difficult", Quality, vol. 38/No. 11, p. 19, October 1999. <http://qualitymag.com>
- [4] Advanced Product Quality Planning and Control Plan Reference Manual : First Edition, June 1994, AIAG, U.S.A.
- [5] Production Part Approval Process : Third Edition, September 1999, AIAG, U.S.A.
- [6] Potential Failure Mode and Effect Analysis Reference Manual : Second Edition, February 1995, AIAG, U.S.A.
- [7] Measurement Systems Analysis Reference Manual : Second Edition, February 1995, AIAG, U.S.A.
- [8] Statistical Process Control Reference Manual : First Edition, 1992, AIAG, U.S.A.
- [9] Quality System Assessment : Second Edition, March 1998, AIAG, U.S.A.
- [10] <http://www.asq.org/9000> : List of the recognized accreditation



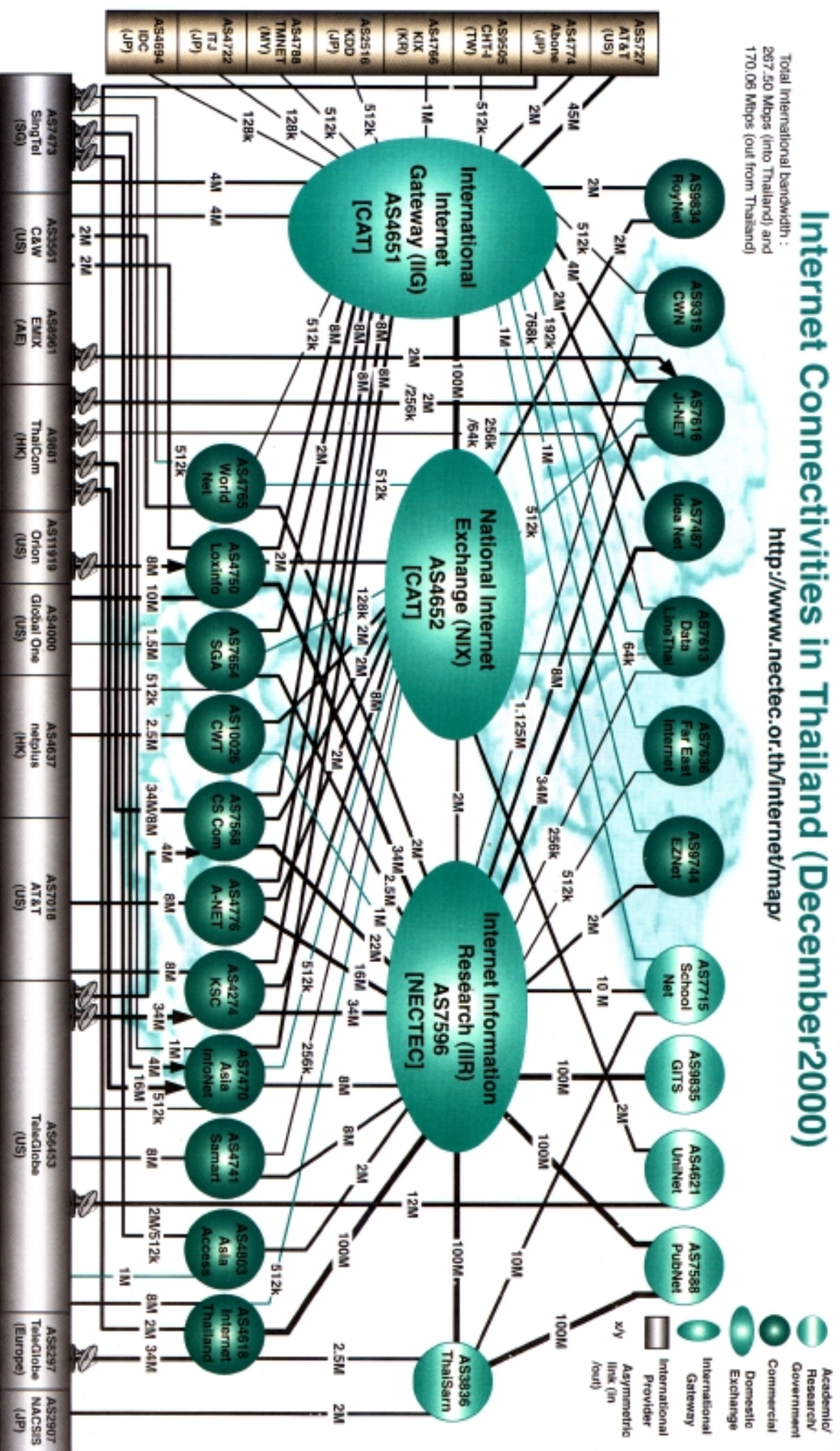
**Mr. Woraphot Thusanametee** received B.S. degree in industrial electrical technology from Pranakorn Teacher College, Thailand in 1992. He is currently studying for Master of Public Administration(MPA) at National Institute of Development Administration (NIDA), Bangkok, Thailand. Since 1985, he has been with Circuit Electronic Industries Public Company Limited in Thailand. His current title is Senior Department Manager and is responsible for Quality System, Quality and Reliability Assurance.



**Suwan Trongjitwikrai** received the B.S., M.S. and Ph.D. degrees in electrical engineering from the University of Arizona, U.S.A., in 1984, 1986 and 1991, respectively. He also received his executive MBA degree from the Asian Institute of Technology, Thailand, in April 1999. Since 1993, Dr. Trongjitwikrai has been with Circuit Electronic Industries Public Company Limited in Thailand. His current title is Vice President and he is responsible for new products development, quality and reliability assurance departments. He is an Adjunct Associate Professor in Microelectronics at the Asian Institute of Technology where he teaches a course in "Fundamentals of IC Packaging, Assembly and Test". He was a senior resident engineer for 3M (Thailand) in 1992. From 1985 to 1991, he was a research associate with the Center of Electronic Packaging Research at the University of Arizona. Between May and August, 1990, he worked as an intern engineer for Bull HN Information System, U.S.A., in modeling and characterization of IC packages. Dr. Trongjitwikrai has published numerous technical articles in the field of electronic packaging. Dr. Trongjitwikrai is a senior member of IEEE Components, Packaging, and Manufacturing Technology Society and a member of Tau Beta Pi and Eta Kappa Nu honor societies.

<http://www.nectec.or.th/internet/map/>

267.50 Mbps (into Thailand) and 170.06 Mbps (out from Thailand)



## Chart Date: 2000-12-01

This chart is designed, maintained and copyrighted by Pinas Teachshong, Kitiya Sringampong and Thaweesak Koonrakool NTL, NECTEC. All rights reserved. The information contained in this chart is based on actual measurements and estimation. We welcome update information, but reserve the rights to verify the accuracy of the given information. Please contact us at [netadmin@ntl.nectec.or.th](mailto:netadmin@ntl.nectec.or.th). For authoritative information please contact Communications Authority of Thailand.

