Development of wireless control and automation systems

Goh Han Leong

NATIONAL UNIVERSITY OF SINGAPORE 2006

Development of wireless control and automation systems

Goh Han Leong (B.Eng., National University of Singapore)

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING NATIONAL UNIVERSITY OF SINGAPORE 2006

Acknowledgments

I would like to express my sincerest appreciation to all who had helped me during my study in National University of Singapore. First of all, I would like to thank my supervisors Associate Professor Tan Kok Kiong and Dr. Fong Aik Meng for their helpful discussions, support and encouragement. Their vision and passion for research influenced my attitude for research work and spurred my creativity.

I would like to give my gratitude to all my friends in Mechatronics and Automation Lab. I would especially like to thank Dr. Huang Sunan, Dr. Tang Kok Zuea, Ms. Raihana Ferdous, Mr. Tan Chee Siong, Dr. Zhao Shao, Mr. Teo Chek Sing, Mr. Andi Sudjana Putra, Mr. Chua Kok Yong, Mr. Jerry Tai for their inspiring discussions and advice.

Finally, I would like to thank my family for their endless love and support. Specially, I would like to express my deep gratitude to Hui Fern for her understanding and support.

Contents

A	ckno	wledgi	ments	1
Li	st of	`Figur	es	VI
Li	st of	Table	s	XII
Li	st of	Abbr	eviations	XIII
Sı	ımm	ary		XVI
1	Inti	roduct	ion	1
	1.1	Impac	et of Wireless Technologies	. 1
	1.2	Contr	ibutions	. 4
		1.2.1	Remote Access of Computer Devices	. 4
		1.2.2	Development of a Mobile Spreadsheet-based PID Control Simula	-
			tion System	. 7
		1.2.3	Development of Bluewave, a New Wireless Protocol	. 10
	1.3	Organ	nization of Thesis	. 13
2	Ren	note A	access of Computer Devices	15

2.1	Introd	luction	15
2.2	Propo	sed System	17
	2.2.1	System Models	18
	2.2.2	Proposed Model	19
	2.2.3	Agent and Server	20
	2.2.4	Service Setup Sequence	22
2.3	Imple	mentation of the System	24
	2.3.1	Server	25
	2.3.2	Agent	30
	2.3.3	Session tracking	36
2.4	Conne	ecting to Email Server	37
	2.4.1	Hotmail	39
	2.4.2	Reading Email	43
2.5	NAT a	and Firewall	44
	2.5.1	NAT Overview	45
	2.5.2	UDP Bombarding	47
2.6	Secure	e Wake-On-Ring	48
	2.6.1	Ring Signal	49
	2.6.2	Telephone Line and Modem Line	53
2.7	Concl	usions	57

 ${\bf 3} \quad {\bf Development \ of \ a \ Mobile \ Spreadsheet-based \ PID \ Control \ Simulation}$

Sys	$ ext{tem}$		58
3.1	Introd	uction	58
3.2	Target	Course	60
3.3	Mobile	e Excel PID Simulator	64
	3.3.1	Process Model and Controller	65
	3.3.2	Application Models	68
	3.3.3	Architecture of Mobile Excel Simulator	72
3.4	Excel	File Conversion	73
	3.4.1	COM and Clipboard	74
	3.4.2	Excel to Bitmap	75
3.5	Image	Processing	76
	3.5.1	Resize and Resampling	77
	3.5.2	Resampling and Gaussian Blur	78
	3.5.3	Subjective Fidelity Criteria	79
	3.5.4	Zoom Function	80
3.6	Studer	nts' Feedback	82
	3.6.1	Methodology	84
	3.6.2	Procedure	85
	3.6.3	Survey Subject	86
3.7	Result	s and Discussion	86
	3.7.1	Section A: Current Scenario	86

		3.7.2 Section B: Benefits and Improvement Areas	89
		3.7.3 Section C: Resistant Factors	91
		3.7.4 Feedback Summary	92
	3.8	Conclusions	92
4	Blu	etooth Assistive Technology	95
	4.1	Introduction	95
		4.1.1 Bluetooth Protocol Stack	96
		4.1.2 Bluetooth Profile	96
		4.1.3 Piconet	97
	4.2	Design Issues	98
	4.3	Implementation Overview	101
	4.4	Bluetooth Server	102
	4.5	Bluetooth Client (Mobile device)	107
	4.6	Cost Issues	111
	4.7	Conclusions	112
5	Dev	elopment of Bluewave: A Wireless Protocol for Industrial Automa-	
	tion		114
	5.1	Introduction	114
		5.1.1 Bluetooth for M2M	116
		5.1.2 Bluewaye	118

	5.2	Bluew	vave Algorithm	. 119
		5.2.1	Zone Discovery	. 120
		5.2.2	Route Setup	. 125
		5.2.3	Route Reconfiguration	. 126
		5.2.4	Zone Maintenance	. 128
		5.2.5	Comparison Studies	. 128
	5.3	Evalua	ation of Bluewave	. 132
	5.4	Discov	very Phase	. 136
		5.4.1	Connectivity	. 136
		5.4.2	Inquiry Window Length and Discovery Time	. 137
		5.4.3	Conventional Flooding Method	. 139
		5.4.4	Bluewave Method	. 141
	5.5	Data '	Transfer Phase	. 143
		5.5.1	Average Hop	. 143
		5.5.2	Route Setup	. 145
	5.6	Concl	usions	. 147
6	Cor	nclusio	ns	149
	6.1	Summ	nary of Contributions	. 149
	6.2	Sugge	stions for Future Work	. 151
Bi	ibliog	graphy		153

List of Figures

1.1	Wireless device overtake internet (source: Motorola)	5
2.1	Hybrid model	19
2.2	Overall network architecture	22
2.3	Service setup sequence	23
2.4	Registration	27
2.5	Proactive IP address update	28
2.6	Reactive IP update and redirection process	29
2.7	Agent	31
2.8	Agent service	33
2.9	Simulation using an emulator	33
2.10	WAP email service	34
2.11	Effect of timeout	34
2.12	Agent with buffer	36
2.13	Session tracking of Agent	38
2.14	Legend	38
2 15	WebDAV example	40

2.16	Initial request and response	41
2.17	Authentication required respond	41
2.18	Authentication required respond	42
2.19	Location response	43
2.20	NAT table	46
2.21	UDP Bombarding	47
2.22	Authentication black box	49
2.23	Authentication black box block Diagram	50
2.24	Phone Pick-up Circuit	51
2.25	Ring Generation Circuit	52
2.26	Telephone Off-hook Detection Circuit	54
2.27	Modem Online and Offline Detection Circuit	56
3.1	Thermal chamber	64
3.2	The PID control system	65
3.3	Preprogrammed Excel spreadsheet	67
3.4	Response to step changes in setpoint	67
3.5	Mobile spreadsheet architecture	73
3.6	Clipboard conversion. The figure illustrates the use of clipboard to con-	
	vert a chart in an Excel file to bitmap format. Note that besides the	
	enhanced Windows metafile format, the chart is also stored in other clip-	
	board formats.	76

3.7	Mobile Excel Simulator. The figure shows the Excel simulator results	
	displayed on a cellular phone	79
3.8	The convolution radius used is 1	80
3.9	The convolution radius used is 2	81
3.10	The convolution radius used is 3	81
3.11	The convolution radius used is 4	81
3.12	The bar chart shows the results obtained from a survey of 40 persons	
	rating the quality of image according to subjective fidelity criteria	82
3.13	Figure 10 (a) on the left shows the original image being split into four	
	equal regions over two rows and two columns. (b) on the right shows the	
	image being split into four rectangular regions on a single row	83
3.14	Using the zoom function, the processed image is enlarged	83
3.15	Section A of the questionnaire	87
3.16	Section B of the questionnaire	90
3.17	Section C of the questionnaire	91
4.1	Bluetooth AT system architecture	99
4.2	Bus commuter enters bus number into his mobile phone	100
4.3	Bus commuter receives notification message when the bus is approaching	100
4.4	Implementation of AT system prototype	101
4.5	Flowchart for the initialization of server	104
4.6	Flowchart of server polling to find matching Bluetooth address	106

4.7	Address of the bus module detected by the server
4.8	Flowchart of Device Discovery process of client device
4.9	Flowchart of Service Discovery process of client device
4.10	Flowchart demonstrating the sequence of audio files for the client application 11
5.1	Multiple PLCs connected to central computer
5.2	Simplified zone formations
5.3	Detail zone formations
5.4	Finite state machine for Zone x node
5.5	Route setup
5.6	Route Reconfiguration
5.7	CNC milling machine
5.8	Communication pathway
5.9	Fault-monitoring network
5.10	Machinery position
5.11	PC acting as CNC machine
5.12	Probability of success vs trial
5.13	Discovery time vs number of nodes
5.14	Packet end-to-end delay vs number hop
5.15	Transfer time vs file size
5.16	Average route setup time

List of Tables

3.1	Rate Scale of the Television allocations	79
3.2	Summary of survey results	93
5.1	States Table	124
5.2	Comparison of the characteristic of bluewave with existing protocol	131
5.3	Average Connectivity of zone x to zone $(x+1)$	137
5.4	Effect of a inquiry length on the success of a trial	138
5.5	Comparison of initialization time of bluewave with conventional flooding	142
5.6	Percentage of nodes in zone x	144

List of Abbreviations

AODV $Ad-hoc\ On-demand\ Distance\ Vector$

 $API \qquad \qquad Application \ Programming \ Interface$

ASP Active Server Page

AT Assistive Technology

CBM Conditional Based Maintenance

CNC Computer Numerical Control

COM Component Object Model

DCS Distributed Control Systems

 $DSDV \qquad \quad Destination \ Sequence \ Distance \ Vector$

 $DSR \qquad \qquad Dynamic\ Source\ Routing$

DTMF Dual Tone Multiple Frequency

EMF Enhanced Windows Metafile Graphics

etc. et cetera

e.g. exempli gratia

E-learning Electronic Learning

GIF Graphics Interchange Format

GPS Global Positioning System

GPRS General Packet Radio Service

IP Internet Protocol

ISP Internet Service Provider

 $JPEG \hspace{1cm} \textit{Joint Photographic Expert Group}$

LAN Local Parcel Service

LAR Location Aided Routing

MEMS Microelectromechanical Systems

M-learning Mobile Learning

NAT Network Address Translation

NTP Network Time Protocol

PC Personal Computers

PDAs Personal Digital Assistants

PID Proportional Integral Derivative

PIM Personal Information Managers

PLC Programmable Logic Circuit

PSTN Public Switched Telephone Network

RAM Read Access Memory

SCADA Supervisory Control Data; Acquisition Systems

SMS Short Message Service

UART Universal Asynchronous Receiver Transmitter

UDP User Datagram Packet

UPS United Parcel Service

 $URL \hspace{1cm} Uniform \hspace{1cm} Resource \hspace{1cm} Locator$

WAP Wireless Application Protocol

 $WRP \hspace{1cm} Wireless \ Routing \ Protocol$

XML Extensible Markup Language

ZRP Zone Routing Protocol

Summary

The advancement in cellular and microelectronic industries over the past few decades had spurred the growth in wireless technologies tremendously. Together with the creation of smaller and more computationally powerful devices, extensive mobility has been endowed on computing. This new wave of mobile computing trend generates ample opportunities for the development of new mobile inventions. It has revolutionalized the way people communicate with one another. With the new emerging wireless technologies, convergence of voice and data communication via wireless devices is now possible. This allows wireless communication technologies to become the key enabling technology in many applications. This thesis examines and exploits the use of WAN, LAN and PAN wireless technologies in three different arenas; home control, education and industry automation.

Cellular WANs have evolved drastically since 1980s. In the first part of the thesis, work has been done on the use of the 2.5 generation of WAN technology to make possible the remote access of PC using mobile devices. A comprehensive architecture has been proposed to establish a link between the mobile device and the computer system. The architecture adopts a distributed instead of conventional centralized approach in

connecting the mobile devices with the computing devices. The distributive measure reduces the overall cost of the system; however, on the other hand, it increases the complexity of the system. In the distributed architecture, the WAP gateway at the base station connects directly to the online computer device. This induces the problem of NAT and Firewall problem if the computer is within a LAN. A "UDP Bombarding" method is devised to counter this problem. As connection can be lost due to gateway timeout, a WML solution is provided in the architecture for resuming the network gateway connection after the assigned access time on the network gateway maintaining the connection has expired. It is not realistic to assume that the targeted computer device is running at all time. A hardware solution that works on the dial-tone principle to boot up the computer devices via Wake-On-Ring function is constructed to enable the mobile devices to activate the target computer remotely.

Based on the architecture proposed, implementation of a mobile spreadsheet-based PID control simulation system that aims to enhance the learning of the student in the field of control engineering has been developed. A file conversion technique that converts the Microsoft Excel spreadsheet to an image format is introduced. As most mobile devices have small form factor, viewing of the resultant image file is difficult. Image processing techniques are employed to resample the image file so as to create a clearer image of the image file when it is viewed on the mobile device. The PID control simulation system comprises a zoom function to improve readability of the image file. The zoom function may segment the image of the image file when viewed on the mobile

device into one or more selectable regions. A selected region may be cropped from the image of the image file when viewed on the mobile device and magnified for viewing on the mobile device. From the pedagogy aspect, the results of the student feedback yield a clear current scenario of students with regards to their potential and readiness to adopt on new learning modes, and the trend in the learning habits with the proliferation of mobile devices.

Next, recognizing the potential of PAN wireless technologies in providing an extra "sensory" channel for the visually disabled. The use of wireless technologies in assistive technology application is examined. Finally, the thesis presents the development and implementation of a new wireless routing protocol, Bluewave, which caters specially to wireless communication among machines in a factory setting. With the availability of inexpensive wireless short range network such as Bluetooth and 802.11, there is a growing trend in developing the wireless Machine-to-Machine (M2M) communication for factory automation. Most of the current wireless protocols focus on the mobility issue and are mainly suitable for narrowband radio devices. The mechanisms that work well for wireless mobile ad-hoc network prove to be redundant, and instead add to processing overheads in a static wireless factory setting. Bluewave utilizes the features of Bluetooth technology when performing the route setup, it caters specifically to wireless communication among machines in a factory setting. Bluewave contains desirable properties such as loop freedom, demand-based operation, short route setup time and good endto-end data throughput. A case study is conducted on the use of Bluewave in an online Condition-Based Maintenance of computer numerical control milling machine setup.

In this thesis, extensive simulation and experimental results will be furnished to illustrate the effectiveness of the proposed approaches.

Chapter 1

Introduction

1.1 Impact of Wireless Technologies

Wireless technologies have undergone great changes since February 1896, when Gugliemo Marconi developed the first wireless telegraph system. Since then, the wireless spectrum is increasingly getting crowded daily. Wireless communication together with miniaturization of integrated circuits revolutionalized our world. For the past ten years, convergence of telecommunication and computing technologies grasps the attention of many inventors and industrial players. In the short span of ten years, cellular wireless data communications have evolved from second-generation (2G) networks to third generation (3G) networks. The change in the network infrastructure increase the communication bandwidth tremendously, from the low data rate of 10kbps in 2G to 2kbps in 3G. This exponential rise in data rate allows content rich services like MMS to deliver a variety of multimedia data to the user. With the introduction of short range wireless MEMS; it is possible to endow intelligence on tiny gadgets within our surrounding. Wireless MEMS open up a new field of research in wireless sensor network and ubiquitous computing.

Together, these developments bring many potential applications in different area, including m-commerce, m-education and industrial automation that have great impact to our society.

For m-commerce, the use of wireless data communication is mainly in transaction management, digital content delivery and telemetry services. Transaction management applications are commonly found in customer service sector, e.g., air travel, automobile rental and store check-out. There are two major concerns in mobile transactions, ease of use and security. [1] examine the digital right management issue, while [2] and [3] propose methods and algorithm to improve the user's mobile transaction experience. Digital content delivery activities include mobile browser, mp3 downloading and mobile email service. Commercial companies like CNN wireless news and Blackberry invest heavily in these activities. Phillip's HomeLab [4] is a typical example of mobile telemetry service, whereby mobile devices are used to control home appliances.

For m-education, there is a rising trend to use mobile computing to facilitate pedagogy. The education resources can be accessed via a mobile web surfing. An example of accessing a central server via mobile devices for information is the KNOWMOBILE [5]. In the KNOWMOBILE project by University of Oslo, the medical student can access web-based medical information via PDAs. Wireless Internet Learning Devices (WILD) [6], uses handheld devices for Computer Support Cooperative Learning (CSCL). This can be thought of as an extension to the E-Learning version of a laboratory session. MOOsburg++ [7] is an online virtual community which allows different groups of users

to interact in collaborative learning while on the field. SMS are also widely use in disseminating short announcements like examination results release date and change of lecture venue.

For industrial applications, wireless technologies are employed in different areas of the manufacturing chain. Wireless technologies reduce the number of cabling connected to the machinery. This reduces the cost of installation and maintenance, and simplified the data collection process. In Oslo, ABB installed the world's first large-scale industrial application of Bluetooth wireless communications technology at 179 water utility pump stations. New technologies like RFID play an important role in products inventory and unfinished part tracking [8]. [9] gives a detailed description on the impact of location-capable phone on intelligent transport system, specifically on automotive telematics and public transit system.

The attempt made by Mark Weiser and his colleagues in 1991 to create a ubiquitous computing environment that delivers data to user failed due to lack of hardware support. A decade had passed, today, mobile devices like laptop, mobile phone and PDAs are widely available. However, there are still many challenges remaining in creating a ubiquitous wireless data communication world. Small form factor of wearable mobile devices put a constraint on its data presentation. Small display screen size may limit the presentation to text-based or audio-based information. Relatively slow processing speed and communication bandwidth as compared to desktop has lowered the user satisfaction as the user is accustomed to the desktop's fast interaction. On the software aspect,

thickness of client side application, proactivity, cross-layer transparency and security are some of the issues that need to be addressed. Problems encountered in the deployment of wireless technologies in a wireless setting are discussed in [10].

In essence, wireless technologies open up vast technological possibilities that were unimaginable previously. There is nowhere in an urban city that is not covered by radio signal. This signifies the intense market penetration of wireless applications. It is the aim of this thesis to explore the uncharted area of wireless applications and meet the challenges encountered in developing new use of wireless technologies to improve the quality of living.

1.2 Contributions

This thesis aims at developing wireless functionalities in different application areas to achieve improvement in the quality of life. As the focus of the thesis, not only the WAN technologies (e.g., GPRS) are examined but also new emerging short range wireless technologies (e.g., Bluetooth and Zigbee) are investigated of its application in industrial automation.

1.2.1 Remote Access of Computer Devices

According to the prediction of Motorola in 2001, the number of internet capable wireless devices will overtake wired internet devices at the end of 2003 as illustrated in Figure 1.1.

It is true that there is a steady increase in the availability of wireless internet devices

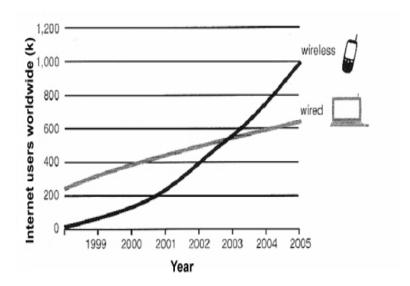


Figure 1.1: Wireless device overtake internet (source: Motorola)

in the market. However, most user still prefer to access internet via wired devices. The internet service offered on mobile devices is limited and is dependent on the network service provider's gateway [11]. This is due to the lack of a unique IP address in the mobile device. Therefore, mobile devices have to access internet via the gateway provided by the network service provider. There are future plans to implement IPv6 network that have intrinsic mobility support [12] for wireless devices. However, before then whereby, all mobile devices have unique IP address, a true mobile computing era has not arrived yet.

The processing speed and the storage space of the handheld mobile devices have posed a great barrier to the development of complex mobile devices software. As such, most mobile devices are unable to run even simple software common in desktop computers. This deters users from mobile computing application.

In this thesis, a solution to the limitation of the processing speed and the storage

space of the mobile devices has been proposed. A comprehensive architecture had been devised to link the mobile devices with the desktop computer. The architecture proposes is a new and novel ways to efficiently link the mobile device with PC in a distributed manner, resulting in resource saving and exploration of exciting application. The novelty of architecture is ensured through patent examination search report conducted by Australia patent office. The architecture is currently patent pending in U.S. (U.S. serial no. 10/549,045).

The applications developed on the proposed platform use the WAP browser. The mobile devices link up with the PC using WAP via a WAP gateway. WAP gateway has timeout mechanisms that will break connection between mobile devices and the PC if it does not receive a response from the PC within a time duration. This often occurs when the PC is running a heavy application and does not send a response in time. In the architecture, a mechanism is devised to solve this premature timeout problem.

LAN is usually protected by firewall and uses a mechanism called NAT to save on the available public address assigned. This poses a problem to the distributed architecture proposed, whereby the base station has to connect directly to the target computer. Without knowing the exact public IP address and the port the target computer is using, the base station will not be able to establish connection with it. Methods to overcome this problem have been devised.

Part of the architecture is dedicated to provide hardware solutions for connection of the mobile device to a power down PC. The hardware solution makes use of the WakeOn-Ring feature in the motherboard, to boot up a shut down Computer devices (e.g. personal computers, laptop). An authentication feature is imbued into the hardware that works on the dial-tone principle to authenticate before booting up the computer devices via Wake-On-Ring function. The overall system is also able to track and backlog the phone number of the phone device that is used to boot up the computer.

To testify the usefulness of the architecture, an application with the aim to improve current mobile email service had been developed. "Complete Mobile email Management" an application that allows a user to access PC email service and send email with attachment via mobile devices is developed based on the architecture. Emphasis is placed on the development of the mobile email service with attachment sending. The ability of sending email attachment is not found in mobile email service.

1.2.2 Development of a Mobile Spreadsheet-based PID Control Simulation System

The electronic spreadsheet was first developed in the 1960s for use in business accounting. An electronic spreadsheet organizes data into columns and rows (also known as cells), yielding a framework which is ideal for data organization and manipulation. The modern electronic spreadsheet provides good formulation functions which allow cells to interact, allowing automated computations, and thus making the spreadsheet an excellent modeling and simulation tool for other fields. Recent years, a clear trend in the engineering fields is to adopt the spreadsheet for modeling and simulation purposes [13], [14] for a number of reasons. First, the spreadsheet is a great modeling tool which is

widely available and, in many cases, is packaged as part of off-the-shelf PC. This makes the spreadsheet a platform which is naturally common to many engineers and facilitates joint development work. Secondly, the learning curve to use the spreadsheet to formulate a model for simulation is considerably shortened when compared to using other dedicated modeling and simulation software, since most engineers would have used the spreadsheet for one purpose or another. Thirdly, the spreadsheet is a good teaching aid. From a pedagogical perspective, the exploratory approach used when programming the spreadsheet improves the student's understanding of the topic. This feature has been highlighted in some recent papers [15], [16].

Concurrently, the advance in wireless technology is enhancing E-learning with mobile functions, facilitating the rise of M-learning. M-learning, in essence, allows the learner the flexibility of time and space to access education resources. However, for mobile education to be pervasive, especially to student on the move, the system has to be adaptable to include affordable, handheld, mobile devices (e.g., cellular phones). Unfortunately, handheld devices typically have small display units and limited memory space, limiting the possibilities of access to educational resources on these devices mainly to text-based messaging. Running simulation on these handheld devices will be unthinkable if not impossible.

Based on "Remote Access of Computer Device" architecture, a system which will allow the learner to access a remote PC that houses an Excel simulator, and convert the result to a standard picture file format was developed. A clipboard function was developed to provide a means to convert the result of Excel simulator to image file easily. When part of the compound document is copied to the clipboard, the object is stored in one or more of the clipboard format. One of the clipboard formats that the object is stored in is the EMF format. EMF as the name suggests is a metafiles graphics that contain both raster and vector data. The object is stored as a vector with bitmap attribute. Procedure is made for the EMF file to be converted to bitmap file. This method is deem to be faster and less intrusive than most current excel to image converter program.

The converted image size in terms of pixel height and width is too large for the cellular phone display screen. After conversion the image size is usually several hundreds pixels in dimension. The cellular phone display screen is only about 200 x 200 pixels, thus resampling is needed. Resampled images have undesirable artifacts, e.g., distortion of straight line. As most alphabets contain fine straight lines, this poses a serious problem to the viewability of the resultant image. The distortion is severe when the scaling factor is high. A gaussian blur image processing technique was developed to smoothen out these artifacts. In the gaussian blurring process, each pixel in the RGB image is mixed with its adjacent neighboring pixel with gaussian probability. Thus, the pixel nearer will have a greater effect on the desired pixel. The amount of affected pixels on either side is determined by the convolution radius specified. By blurring the image before performing resampling, the fine details will be "enlarged" and remain visible after resampling. The size of the convolution radius used is dependent on the resampling scale

factor. An algorithm is used to blur the image according to the resample scale factor size.

A mobile simulator system that enhances the student understanding in PID Control was delivered. The case study for the use of the system in an industrial control engineering course is conducted. The system is evaluated in the form of student feedback. The details on the development of the system and the nature of the course to which it is applied will be useful to readers who are keen to develop a similar M-learning application for their students. The results of the student feedback yield a clear current scenario of students with regard to their potential and readiness to adopt new learning modes and to use the trend in the learning habits with the proliferation of mobile devices.

1.2.3 Development of Bluewave, a New Wireless Protocol

M2M technology is fast gaining popularity in industrial automation [17]. Machines communication is not a new concept. As early as the 1990s, SCADA had already gained prevalent use in distributed control systems. However, the proprietary nature of SCADA limits its widespread use in the industry. Deployment of a SCADA system is time-consuming and usually involves a number of engineers, who are required to design the complex communication network that is industry-specific.

The pervasiveness in internet had spurred the next phase in the development of SCADA. Standardized communication protocol (e.g., TCP/IP) in internet provides the technologies needed for SCADA to be evolved to wired M2M. For example, in [18], it is proposed that ethernet is used as the communication interface for electric drives. M2M

open and collaborative peer to peer properties offer more flexibility compared to the server poll environment typical of SCADA, in which the data cannot be pushed by the field nodes to the server.

Recent advances in wireless communication technologies (e.g., 802.11, Bluetooth) and MEMS have caused a positive spillover effect in M2M. Intelligent sensor network plays an increasingly important role in industrial control systems today. This can be seen from standards, for example, the IEEE 1451 [19] formed by various workgroups. Wireless sensor networks have opened up new possibilities for industry automation; it allows online CBM [20] and wireless PLC control [21, 22].

With the improvement in the performance of the wireless technologies, wired M2M is now ready to be transformed into wireless M2M. However, conventional communication protocol used in wired M2M cannot be used directly on the wireless arena. As pointed out from the research conducted at MIT's Media Lab in Cambridge, point-to-point or point-to-multipoint transmission is not suitable for use under a factory setting, due to the interference caused by the metal parts in the machinery [23].

Wireless mesh network topology aims to solve the interference issue by transmitting the data packet in short hops over a long distance with intermediate nodes acting as relay routers. Currently, there are many wireless mesh routing protocols proposed and each has its own advantages. Till now, there is still no single dominant design adopted by the industry. Most of the existing wireless protocols have assumed that the node is mobile and there is no prior knowledge in the positioning of the nodes. Extensive

mechanisms, for example, to allow frequent beaconing updates have been constructed to ensure the connectivity of the neighbouring nodes. These mechanisms are not necessary in the factory setting as the machines are stationary. Assumptions can also be made on the relative position of the machines.

In this thesis, a new wireless routing protocol, Bluewave, which uses Bluetooth technology is proposed with the objective to perform M2M communication in a factory setting. Bluetooth has several advantages over other wireless technologies (e.g., 802.11) when it comes to wireless M2M for industry automation. In fact, there are already a number of industrial sensing applications that use Bluetooth [24].

Bluewave contains detailed connection set up and maintenance algorithm that mitigate the limitation and constraints in Bluetooth, it consist of four phases 1) Zone discovery, 2) Route setup, 3) Route reconfiguration and 4) Zone maintenance.

Bluewave synchronized an alternate active and dormant state of the nodes in the same spatial area to reduce the interference. Reducing the number of active nodes within the same transmission area and timeframe will enhance the probability of a successful discovery. Experiment is conducted to demonstrate the discovery message process in Bluewave. From the data collected, comparison is made between Bluewave and conventional flooding method.

As a result of the circular geometrical shape of the zone formed in Bluewave and given that the factory floor setting is rectangular the middle zones will occupy the most area. Therefore, most of the nodes are found in the mid zone. This reduces the average

hops distance of a node away from central computer, thus effectively increase the overall throughput of the system.

The experiment also investigates into the difference in performance between memory and memoryless route setup. It is shown that the memory approach adopted in Bluewave increases route setup performance drastically.

The proposed protocol features together with Bluetooth characteristics minimize the route setup and data transfer time. Bluewave is also suitable in other application areas like, Home automation and Healthcare environment, where the node's mobility is low.

1.3 Organization of Thesis

The thesis is organized as follows.

Chapter 2 presents the architecture that enables the mobile device to establish connection with a computer device. A review on the possible system model namely; thick client and thin client is conducted to determine the suitable approach in designing the architecture. Detailed description of the hybrid model proposed for the architecture that leverage on the advantages of thick client and thin client model is given. The communication of the User, Server and Agent in the hybrid model will be examined. Implementation of a mobile email service application based on the proposed architecture is presented. A time line diagram is used to show the service setup sequence. Problem and solution concerning NAT, WAP timeout and session tracking will be elaborated.

Chapter 3 gives a short description of the control engineering course used in the case

study. In this chapter, the development of the mobile simulator will be elaborated in details. The detail of the control process simulated by the mobile simulator will be furnished. A method to convert the Excel spreadsheet to image file efficiently is proposed. The image processing algorithm used to improve the delivered image is shown. Discussion and analysis is made on the feedback collected from students which is used to assess the usefulness of the system.

Chapter 4 demonstrates the use of PAN wireless technologies in AT application. The chapter illustrates the implementation of a "Bus Notification System" using Bluetooth wireless technology that enables automatic notification of the requested bus number to the commuter via mobile devices.

Chapter 5 proposes a new wireless protocol, Bluewave that caters specially for factory setting. The new protocol utilized the features of Bluetooth technologies. In the first part of the chapter, a short introduction on Bluetooth usage in M2M and current wireless protocols is given. The Bluewave algorithm is described and comparison is made between the characteristic of Bluewave and some of the existing wireless protocol. A case study is conducted to examine the performance of Bluewave. The subject of the case study is a CNC milling machine. Experimental results obtained will be duly discussed.

Finally, conclusions and suggestions for future work are discussed in Chapter 6.

Chapter 2

Remote Access of Computer Devices

2.1 Introduction

Since the beginning of early 1990s, computers have infiltrated into many aspects of our lives. It has brought great changes to the way businesses is conducted and has also enhanced the speed at which knowledge propagate. However, the power of computer has been restricted to nearest available terminal. With the advancement in Wireless Communication, the boundary of computing is now expanded; there are numerous handhelp mobile devices (e.g., Personal Digital Assistance and mobile phone) that allow us to do computing on the move. Yet with limited processing power and hard disk space, these mobile devices can so far at best only serve as PIM. Linking mobile devices with the PC enable us to harness its immense processing speed and storage capacity. In addition, for some application such as email service, it is necessary for mobile devices to access the PC. The current WAP email service offered by most telecommunication providers lacks an important Internet email function: the ability to attach files. By providing the mobile device access to user PC, the user will be able to send emails with file attachment using

their mobile phones. Realizing the potential in linking mobile device the PC, intense research and development have been conducted by various institutes as follows:

Mai Messaging Agent [25] - Mai is a software Agent that allows email attachments to be easily and efficiently viewed on virtually any connected device, including most web-enabled mobile phones and PDAs. Mai resides in the data center near the email server. Mai translates and converts the attachment file in an email to a picture format and send it to user mobile devices. It is also able to unpack MIME files and decompress ZIP files. One key difference between the Mai system and the proposed system is in the position of the Agent. In the Mai system, the Agent is residing at a centralized data center, while in the proposed system, the Agent is residing at the user computer. Since the Agent in the proposed system is located at the user's computer, the user is not only able to view an email attachment, the system also allows the user to access the file in the computer and send it as an attachment in an email. In Mai, the user is only able to view an email attachment. The proposed system being a distributed system, also has the added advantage in resource saving as the user computer handles all the processing of data. For a corporate user, the privacy and security offered by the proposed system is an attractive feature as no data is passed to the central data center.

Intellisync [26] - the architecture used in Intellisync is very similar to the proposed system. Like the proposed system, the Agent used in Intellisync is on the user computer. However, unlike the proposed system, the Intellisync Agent does only partial data processing. In the proposed system, the Agent acts as a WAP server and it is ac-

cessed on demand. Formatting required to view an email attachment will be processed at the Agent. The Intellisync Agent, on the other hand, maintains a permanent connection to a central server and the required data is transmitted to the central server to be processed. In this aspect, the proposed system is a better candidate in term of resource saving and privacy.

There are other variations of WAP email applications based on the above mentioned two schemes. To the best of the author's knowledge, the proposed system is the only WAP email system that allows the user's computer to handle all the processing. The central server is only used for redirection in the initial phase. The difficulty in constructing a truly distributed system lies in the NAT and the firewall used in a LAN environment. The NAT and firewall prevents the base station of the user's mobile devices to connect directly to the target computer. The proposed solution to this problem will be discussed in a subsequent section.

2.2 Proposed System

The network architecture used in the implementation of the proposed system is shown in Figure 2.2. In the following subsections, the functions of key components within the system will be briefly presented.

2.2.1 System Models

Possible system models which may be used will be first briefly reviewed.

Client/Server

The Client/Server model was first used in the 1980s. Due to the good scalability and flexibility when compared to mainframe computing, it has become a pervasive computer software model by late 1980s. There are two main types of Client/Server architectures; "two-tier" and "three-tier". In this model, the client will issue request for service to the server and the server will act as the provider for the service. The two-tier architecture consists of three components fulfilling three functions; User System Interface, Processing Management and Database Management. The User System Interface resides on the client's side. Processing Management will be shared between the client and the server. The three-tier architecture is meant for systems with more than 100 simultaneous users. It has an additional middle tier that can perform queuing, application execution and database staging.

Thin client

Thin client is a form of browser-based development model, in contrast to the Client/Server model which can be considered to be an application-based model. The thin client model was first used in 1996. In recent years, it has become a popular software model for the development of mobile applications. Under the thin client model, all the application processing is performed on the server side. This enables the devices on the client's side to work with minimal requirements on hard disk and RAM capacity. This feature is well

suited for mobile applications since mobile devices are limited in terms of such storage capacities.

2.2.2 Proposed Model

A hybrid model, as shown in Figure 2.1, is proposed for the system to leverage on the advantages of each of the schemes.

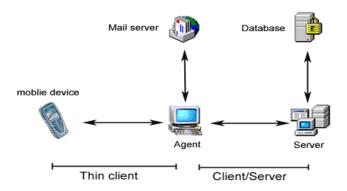


Figure 2.1: Hybrid model

Mobile device and Agent

As shown in Figure 2.1, the mobile device and the Agent exhibit a thin client relationship. Adopting this model has several advantages. First, the mobile device does not need to be installed with any program. This simplifies the system development as it will not pose any additional demands on existing mobile devices. Secondly, the interoperability is enhanced as different mobile devices based on the same type of WAP browser will be able to view the output of the Agent correctly.

Here the Agent acts as a server. To allow the user to read emails remotely, the Agent will retrieve the emails from the email server which can be a POP3, IMAP, or a web-based server. In order to retrieve email from a web-based mail server, the Agent will first

establish a HTTP connection with the server. If the remote PC accesses the Internet via a proxy server, the Agent will have to additionally establish a contact with the proxy server of the ISP before setting up the HTTP connection via the proxy server. Details on connection to web-based email server will be presented in Section 4.

The Agent allows the user to attach any file in the target PC to the composed email. To implement this service, Java I/O API is used to enable the Agent to browse through the file directory in the target computer and to copy the designated file to the email multipart through the file streaming function in Java I/O API. To further facilitate and shorten the time of file transfer from the Agent to SMTP server, a file ZIP feature is developed.

Microsoft PowerPoint attachments can be opened and converted to GIF image file by the Agent. The resultant image file can then be retrieved with a MMS-enabled mobile device. The same facility can be extended to other Microsoft documents such as WORD and EXCEL. Thus, not only remote documents can be attached to the email, but they can also be accessed for their contents.

2.2.3 Agent and Server

To improve on the security and usability of the system, the thin client model is extended to encompass a Client/Server part in the wired network. In the Client/Server portion of the system, the Agent has been switched to act as a client. The functions of the server are to provide authentication and perform "offline" tasks like SMS alert and user profile maintenance. The "offline" tasks are defined as processes which can take place without

the user mobile device connected to the Agent.

If a pure thin client model is used, the user will have to send his/her email account user ID and password to the Agent each time he wants to access his/her email. This can prove to be a potential security hazard. With the use of Client/Server component, the user will first need to log onto the Server for authentication. The Server will retrieve the encrypted userid and password from the database and match against those provided by the user. Upon positive verification, the Server will redirect the user to his/her remote PC. Once the user is linked to the Agent, the Agent will retrieve the encrypted password from the Database and make a direct connection to the specific email server (e.g., Hotmail and POP3 servers), thus relieving the user from a redundant chore of having to enter the same data on the mobile devices since authentication has already been performed.

The Server will monitor the arrival of new emails at the email server of the user. Upon detection of the arrival of a new email, the Server will send a SMS alert via a SMS gateway to the user pre-selected mobile device. The user also has the flexibility to modify his/her profile (e.g., the interval of new mail detection) by logging to the server via any PC terminal. The user may boot up the remote PC via the BIOS wake-on-ring function. However, enabling the wake-on-ring function indiscriminately is not desirable as the PC will boot up whenever there is a call which may well be a normal house call. Thus, an authentication black box will be constructed, based on the dial-tone principle, which will authenticate the user before proceeding to boot up the computer via the

wake-on-ring function. The overall network architecture is shown in Figure 2.2.

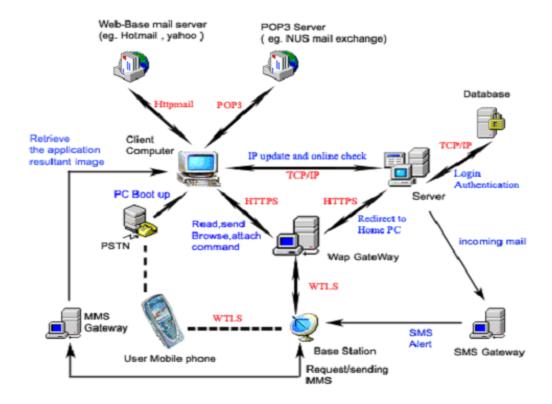


Figure 2.2: Overall network architecture

2.2.4 Service Setup Sequence

To illustrate the message flow within the system architecture, an example of a user accessing the Agent will be presented in this subsection. Figure 2.3 shows the sequence graphically. In this example, the user email server is Hotmail and the target PC accesses Internet via a proxy server. Two key points should be noted in the interpretation of the example:

- The proxy server does not require authentication.
- The redirection process of Hotmail is not shown. In actual fact, upon successful

authentication with the hotmail server, there will be four redirections to different HTTP servers before the email messages can be accessed.

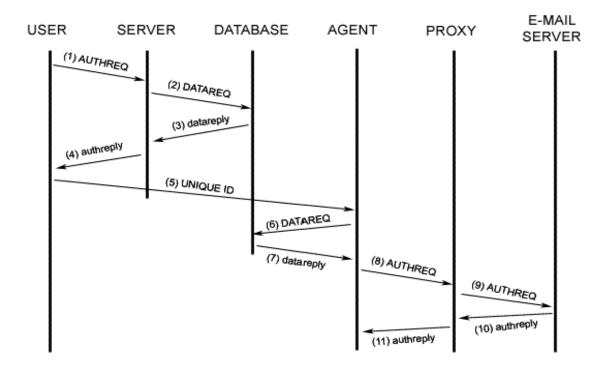


Figure 2.3: Service setup sequence

With reference to the numerical label in Figure 2.3, the steps for the service setup are summarized as follow:

- (1) User sends user ID and PIN to the Server
- (2) The Server searches the Database for a match. If no match is found, the user will be directed back to the login page. Otherwise, the sequence continues.
- (3) The IP address of the corresponding match will be retrieved and transmitted to the Server.
- (4) The current IP address of the target PC and a unique ID will be sent to the user.

- (5) Based on the IP address obtained, the user will establish a connection with the Agent on the target PC. A unique ID is transmitted to identify the user.
- (6) When the Agent is activated, it will make a request to Database for the email server user ID and password.
- (7) The Database will respond to the Agent request with the encrypted email server user ID amd PIN.
- (8) Using the de-encrypted password obtained from the Database, the Agent will establish connection with the web-mail server (Hotmail). The authentication request is first sent to the proxy server of the ISP.
- (9) The proxy server relays the request to the Hotmail HTTP server.
- (10) (11) Assuming the authentication is successful, the Hotmail server will reply to the Agent with the email message location.

2.3 Implementation of the System

Two programming languages, JAVA and PASCAL, are used in the development of the system. JAVA has the virtue of providing vast API support and suitability in developing WAP applications, while Delphi (Developing tools for PASCAL) provides good control of API under windows, thus it is conducive to use to facilitate TCP/IP connection and user interface development.

2.3.1 Server

In previous section, the functions of the Server including authentication, redirection and registration have been presented. JAVA servlet [27] is chosen for server side computing as it offers many attractive features which make it suitable for the application that have been developed. These features will be briefly described as follows.

- A servlet is a JAVA application. By using servlet, it is possible to utilize the vast resources of JAVA API provided by SUN Microsystems. Some of the JAVA APIs used in the development of the system are JAVA I/O, JAVAMAIL [28], sunjce_provider [29] and commons_httpclient [30]. With the use of these APIs, the development time of the application is greatly shortened and reliability of the system is enhanced as each of these APIs has been widely tested and proven.
- The servlet allows for good security features in the system [31]. Unlike ASP or Cgi-perl, the user will not be able to view the source code since the servlet class file is a byte-coded file which is not readable with a text editor. However, the servlet class file can be reverted to its source code using JAVA decompiler. Using the cryptography library provide by SUN Microsystems, sensitive information like user ID and password can be encrypted.
- Session tracking of API presence in a servlet will further enhance the security and efficiency of the system. Session tracking is the capability of a server to maintain the current state of a client's sequential request. Being a stateless protocol,

HTTP does not inherently support session tracking. Without an adequate session tracking, the user will have to log in their email server user ID and password each time they read a different email. This will prove to be highly inefficient, especially for tasks such as forwarding emails, which require the previous state information. Several methods are possible to implement session tracking, including; (a) Hidden form fields, (b)URL rewriting, (c)Cookies, (d)Using the HTTPSession interface available in servlet API. The security issues are cumbersome with methods (a) and (b). If the session information is written in the HTTP header, it is possible for a third party to intercept the packet via an intermediate router and retrieve the session information. Method (c) is not viable with certain WAP gateways which do not accept the use of cookies. A servlet session API keeps the session information at the server's side and this information can be retrieved based on the session ID placed in the cookie file at the WAP gateway. Thus, there is no sensitive session information being transmitted between the user WAP gateway and the Server/Agent. Although some WAP gateways do not accept cookies, this problem can be resolved by using URL rewriting to store the session ID.

In the following subsections, the registration procedure and IP address [32] update method used in the implementation of the redirection function will be elaborated.

Registration

A registration key is required for each Agent installed in a remote computer. If a user wants to install the Agent on his/her office and home computer, two separate registration

keys are needed. The purpose of the registration key is to keep track of the user account and it is also used during a redirection.

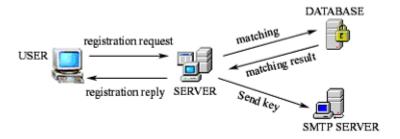


Figure 2.4: Registration

To obtain the registration key, the user is required to visit a designated HTTPS server and choose a user ID and PIN. The user ID and PIN keyed in by the user will be matched against the database to check whether the user ID and the PIN entry already exist. If no match is found, the Server will generate a registration key and store it with the user's ID and PIN into the database. The Server will also send the registration key to the email account provided by the user via a SMTP Server. Upon successful registration, the user will be able to configure his/her email account setting.

Redirection

The Server will redirect the authenticated user to his/her associated Agent. To be able to do this, the Server has to know the current IP address of the Agent. The IP address of the target PC can be a dynamic one, which means that each time the target PC goes online, a different IP address can be assigned to it by the ISP. There is basically two types of mechanisms (proactive and reactive) that can be used to keep track of the dynamic IP address of the target computer.

The steps associated with the proactive IP address updating (Figure 2.5) are as follows:

- a) The Server will open up a many-to-one type of TCP/IP port.
- b) When the target computer goes online, the Agent will establish a TCP/IP connection with the Server.
- c) The Agent will transmit the IP address and its registration key to the Server.
- d) Based on the registration key, the Server will update the current IP address of the Agent with the database.
- e) The Server will also allocate a connect ID to the Agent.
- f) When the Agent goes offline, the Server will be notified and based on the connection ID, the Server will update the Agent's status in the database as offline.



Figure 2.5: Proactive IP address update

Under a proactive IP address update scheme, when the user logs on to the Server and requests a redirection to its Agent, it will be immediately notified of the offline status of the target PC. Alternatively, an SMS may be sent to the user mobile device when the Internet connection of the target computer is lost. The disadvantages associated with a proactive IP address update include the following

• As the number of users increases, the load of maintaining TCP/IP connection with

different Agents becomes significant and the performance of the Server will be correspondingly degraded.

• The opening of many-to-one type of TCP/IP port may make the server vulnerable to denial of service of attack.

The steps associated with a reactive IP address updating (Figure 2.6) are as follows: a) When the target PC goes online, the Agent will establish a connection with the Database.

b) The Agent will search the corresponding registration key entry and update its current IP address.

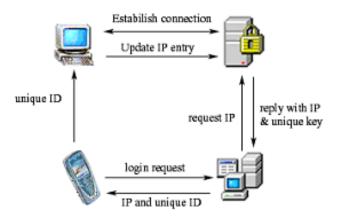


Figure 2.6: Reactive IP update and redirection process

The advantages associated with a reactive IP address update include the followings.

• Each Agent is responsible of updating its current IP address into the database, thus there is no loading on the performance of the Server.

• The reactive IP update mechanism is simple to implement.

The disadvantages associated with the reactive IP address update include the followings:

- When the target PC goes offline, the Server is not informed. The database might contain an outdated IP address.
- A scenario may occur such that when a target PC A with an IP address X goes offline, and when target computer B goes online, the ISP issues PC B with IP address X. Thus, when user A logs into the Server, he will be accessing the PC of user B.

To prevent the scenario painted from happening, an unique key is implemented. As shown in Figure 2.6, when the user logs into the Server, apart from replying with the IP address, the Server also sends a unique key to the user. The user will be redirected to the Agent with the unique key. The Agent will match the unique key to its own to see if the user has the right to access it.

For this system, the reactive method has been selected due to its relatively simpler implementation and better performance when used with a large pool of users.

2.3.2 Agent

In this subsection, the languages used to develop the various functions of the Agent will be elaborated. Specific solutions to key issues will also be elaborated.

Delphi and JAVA

Delphi

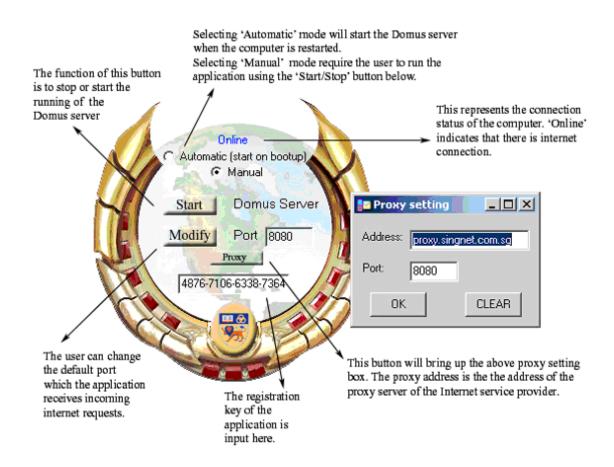


Figure 2.7: Agent

Delphi has been used to develop the following functions:

- Setting the Agent as a service in the Windows OS, thus minimizing interference with the user activity. This service nature of the Agent is depicted in Figure 2.7.
- Checking Internet connection status using Winsock API [33].
- Establishing connection with Database and updating its IP address.
- Recording the proxy information to be used by the servlet.

JAVA

JAVA has been used to develop the WAP functionality behind the system, the database connection and de-encryption. These functions are briefly listed as follows:

- Established connection with the Database.
- De-encrypted the retrieved email server userid and password.
- Established connection with the email server.
- Read/send email.
- Viewing attachment
- File I/O and zip function.
- Attachment access.

Figure 2.9 shows a simulation of the use of the email service provided by the Agent as it appears on the WAP emulator. Figure 2.10 shows the actual display of the proposed system on a mobile device.



Figure 2.8: Agent service



Figure 2.9: Simulation using an emulator

WAP gateway timeout



Figure 2.10: WAP email service

One of the critical problems faced by the Agent is the premature timeout termination by the WAP gateway. When the Agent executes a time consuming task, for example retrieving email from the Hotmail server, the time needed to perform this process can exceed the timeout assigned by the WAP gateway, thus causing an "unknown response" error. This phenomenon is shown in Figure 2.11.

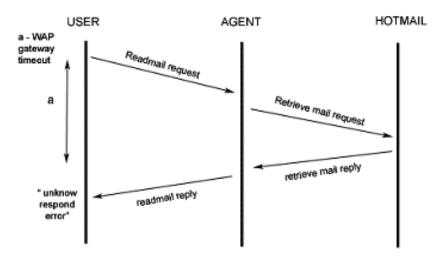


Figure 2.11: Effect of timeout

To resolve this problem, a "time buffer" code is additionally introduced. The following shows a pseudo code readbuffer class:

- < Begin Wireless Markup Language coding >
- < Timer start >
- < On event timeup >
- < Goto actual read class >
- < End Wireless Markup Language coding >

ConnectHotmail

← Timeout occurred

ReadHotmail

When the read class file is run, it will first generate the Wireless Markup Language code. During the first time execution of the connect Hotmail procedure "ReadHotmail", the timeout of the WAP gateway is reached and the connection between the WAP gateway and the Agent will be lost, but the transaction between the Agent and the Hotmail server can still carry on. The WML code generated will be downloaded to the user mobile device and be executed. The WML code will direct the mobile device back to the Agent after 30 seconds. By then, the execution of "ReadHotmail" procedure will be completed and the user will be able to read the email. Figure 2.12 illustrates this process graphically.

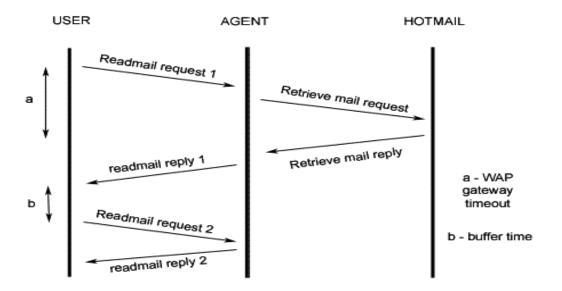


Figure 2.12: Agent with buffer

2.3.3 Session tracking

Session tracking of the user by the Agent is needed to increase the efficiency of the email service. The use of a session tracking API has facilitated this task. However, there may arise a scenario when the use of hidden form fields for session tracking will be more efficient. Session tracking by hidden form fields is achieved by using the HTML "Input tag" by setting the attribute of the "Input" tag to be "Hidden" as the following: < INPUT TYPE="HIDDEN" NAME="session" VALUE="..." >. This will pass the session information to the next servlet visited, thus the current servlet will have session information of past servlet. Each page that the servlet generates can embed these hidden form fields, thereby keeping track of a particular client's state. As shown in Figure 2.13, during certain transactions between the class file, a hidden field in the form of post method is used to transmit the session information. For example, with reference to

Figure 2.13, the transaction between class file inboxes (the top left box in Figure 2.13) and Wapinbox (the top middle box in Figure 2.13), hidden field value acctno is sent by the post method. The purpose is to notify the Wapinbox that it is the first time it has been called and the initialization process is to be carried out. If a session API is used instead to check for initialization, a field for each initialization has to be kept in the Server, thus incurring an additional load on the resources and subjecting the program to errors, as the field may fail to reset when the user leaves the service.

2.4 Connecting to Email Server

Currently POP3, IMAP and web-based mail are the more popular ways of accessing email over the Internet. POP3, the oldest among the three methods, works well for users who want to keep all their mails on one computer. IMAP provides server-side storage and manipulation of emails. It works better when the user has multiple computers and needs to be able to access his/her mails from any of them. Hotmail and Yahoo!mail implemented the use of web-mail to access the email servers provided by Microsoft and Yahoo respectively.

In this system, the Agent is responsible for connecting to the different types of email servers (categorized by the methods of access) to retrieve emails and then send the data to the user's mobile phone using WAP. Connecting to a POP3 server is a relatively easier task, since there exists a set of APIs in JAVA that allows the connection to POP3 email server. Retrieving email from a web-based server such as a Hotmail server proves to be

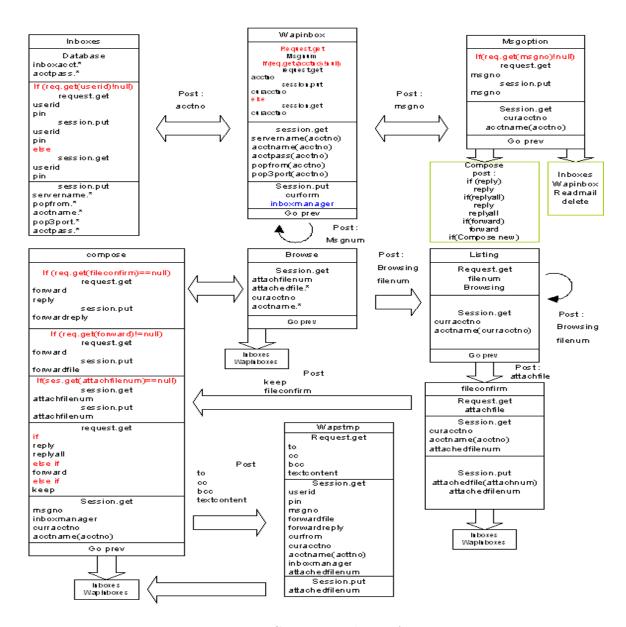


Figure 2.13: Session tracking of Agent

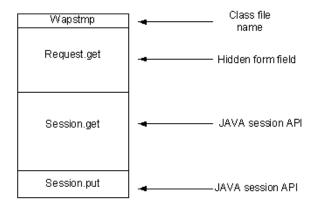


Figure 2.14: Legend

a tougher job. Hotmail, for example, uses a protocol called WEDAV [34] to negotiate the retrieval of emails. A HTTP connection must first be set up between the Agent and the Hotmail server, after which a series of redirections will follow before the Agent retrieve his/her email. A separate subsection will be devoted to a fuller discussion on the connection to the Hotmail server. The connection to Yahoo!mail is based on the same principle as a connection to Hotmail server.

To display the retrieved email in mobile devices, there is another obstacle that has to be overcome. Due to the processing power and the storage space limitation of the devices, an additional processing step has to be performed on the retrieved email before the message can be displayed.

2.4.1 Hotmail

Hotmail server is accessed using HTTPmail protocol. HTTPMail is really a set of Web-DAV queries/properties implemented in Hotmail. To access Hotmail email it is necessary to establish a HTTP connection with the hotmail server and exchange information according to the WebDAV protocol to retrieve the Email.

WedDAV

WedDAV define the HTTP extensions necessary to enable distributed web authoring tools to be broadly interoperable, while supporting user needs. To put it simply WebDAV provide HTTP extension that allow document to be modify over web.

To-date, the WebDAV working group has defined a set of extensions to the base Hypertext Transfer Protocol for six capabilities. The first three-over-write prevention, properties, and name-space management are fully implemented. Figure 2.15 shows how a typical Web client could use these extensions to edit the source of a Web resource.

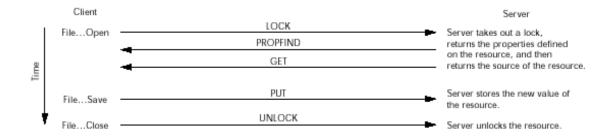


Figure 2.15: WebDAV example

In retrieving Email from Hotmail server, PROPFIND and PROSTAT are used to find out the location of the Email message.

Connecting Hotmail

The default HTTPMail gateway for Hotmail inboxes is located at:

http://service.msn.com/svc/hotmail/Httpmail.asp

However, this website is only an access point. The real server is implemented somewhere else on the MSN network. In order to connect to a secure server, the WebDAV protocol requires HTTP/1.1 authentication. The initial request as shown in Figure 2.16 is sent by the Agent using the WebDAV PROPFIND method to query for a set of properties. The aim is to find the redirection path to the real server at:

http://oe.hotmail.com/cgi-bin/hmdata.

The same request, as shown in Figure 2.16, will then be sent to http://oe.hotmail.com/. The Hotmail gateway will respond with a HTTP error 401 message. This error message signifies that authentication is required. The HTTP error 401 response also indicate the

Request	Respond
<pre><?xml version="1.0"?> <d:propfind xmlns:d="DAV:" xmlns:h="http://schemas.microsoft.com/hotmail/" xmlns:hm="urn:schemas:httpmail:"> <d:prop> <h:adbar></h:adbar></d:prop></d:propfind></pre>	HTTP/1.1 302 Object Moved Server: Microsoft-IIS/5.0 Date: Mon, 25 Feb 2002 18:26:10 GMT P3P:CP="BUS CUR CONo FIN IVDo ONL OUR PHY SAMo TELo" Connection: close Location: http://oe.hotmail.com/cgi-bin/hmdata Content-Length: XXXX
<hm:contacts></hm:contacts> <hm:inbox></hm:inbox> <hm:outbox></hm:outbox> <hm:sendmsg></hm:sendmsg>	Content-Type: text/html Expires: Mon, 01 Jan 1999 00:00:00 GMT Cache-control: private
<hr/> hm:sentitems/> <hm:deleteditems></hm:deleteditems> <hm:drafts></hm:drafts> <hm:msgfolderroot></hm:msgfolderroot> <h:maxpoll></h:maxpoll>	<pre><head><title>Object moved</title></head> <body><h1>Object Moved</h1> This object may be found here </body></pre>
<h:sig></h:sig>	

Figure 2.16: Initial request and response

type of the authentication scheme (Basic or Digest [35]) needed.

```
HTTP/1.1 401 Authorization Required
Server: Microsoft-IIS/5.0
Date: Wed, 27 Feb 2002 02:12:56 GMT
P3P:CP="BUS CUR CONo FIN IVDo ONL OUR PHY SAMo TELo"
Connection: close
Content-Type: text/html
WWW-Authenticate: Digest realm="hotmail.com",
nonce="MTAxNDc3NTk3Njo4Y2RmOWQyZjZIMzVjMWVhNzM2NTcxM2NlNjc0N2ZiOA==", qop="auth"
X-Dav-Error: 401 Whong email address...
HMServer: H: DAV33 V: WIN2K 09.04.15.0012 i D: Feb 25 2002 15:33:22
```

Figure 2.17: Authentication required respond

A series of response and request follows that direct the Agent to the user mailbox

location. The following in Figure 2.18 is one of the intermediate redirection steps. The request is now redirected to one of the multiple hotmail servers, and the URI has been modified to contain identification for the user. This information has been picked-up by the server from the authorization header field.

HTTP/1.1 302 Redirected Server: Microsoft-IIS/5.0

Date: Wed, 27 Feb 2002 02:33:48 GMT

P3P:CP="BUS CUR CONo FIN IVDo ONL OUR PHY SAMo TELo"

Connection: close

Expires: Mon, 01 Jan 1999 00:00:00 GMT

Pragma: no-cache Cache-Control: no-cache Content-Type: text/html

Location: http://law3.oe.hotmail.com/cgi-bin/hmdata/USERNAME@hotmail.com?

X-Day-Error: 200 No error

HMServer: H: DAV17 V: WIN2K 09.04.15.0012 i D: Feb 25 2002 15:33:22

Figure 2.18: Authentication required respond

Finally, after several redirections, the correct URI has been identified and the user has been authenticated. Once connected, the server will request to set various cookies, valid for the current session. From the first response (as shown in Figure 2.19) after successful connection with the mailbox, the URL locations of the folders within the mailbox can be obtained.

After obtaining the URL of a mailbox folder, the Agent will send a WebDAV request to the folder's URL. The server will then respond with an XML stream containing information for each email contained within the folder. The mail message can then be extracted from the mail response.

Figure 2.19: Location response

2.4.2 Reading Email

The retrieved email message cannot be displayed on the user mobile device directly. Due to its limited processing power, WAP browsers are not able to support most image file and HTML display. There are basically two types of email message format; text/plain and HTML. If the mail is plain text, i.e., it has a MIME type of "text/plain", then it can be displayed directly. If, however, the MIME type is multipart or HTML, additional processing must be carried out on the message. The following shows the partial code for distinguishing between the two type of messages.

```
// Determine Message contents data type

Object messageContent = message.getContent();

if (message.isMimeType("text/plain") && messageContent instanceof String) {

// Show message
```

} else if (message.isMimeType("HTML") && messageContent instanceof String) {
parse(messageContent) }

Before displaying message content that is "HTML", it is necessary to first parse off the HTML tag. The HTML tag is identified by the < > symbols enclosing it, for example . By searching line by line for the opening of the < and remove the content until the detect of the closing of tag > , all the HTML tag will be able to remove. After the HTML tag has been removed, only the message content will be left. This plain text message content will then be suitable to be displayed on the mobile devices.

2.5 NAT and Firewall

Due to rapid depletion of public IP address, most LAN networks today use NAT. NAT allows a single device, such as a router, to act as an Agent between the public network (Internet) and a local (LAN) network. This means that only a single, unique IP address is required to represent an entire group of computers. Thus in simple term, it allows multiple PC to share one IP address.

Besides allowing IP reuse, NAT have created another effect. Implementing NAT automatically creates a firewall between public network (Internet) and a local (LAN) network. NAT only allows connections that originate inside the LAN. Essentially, this means that a PC on an external network cannot connect to a LAN computer unless the LAN computer has initiated the contact. A PC in the LAN can browse the Internet and connect to a site, and even download a file; but a computer from the internet cannot latch onto

LAN PC's IP address and use it to connect to its port.

This pose a problem to the system and architecture proposed which uses the WAP gateway to initiate connection to the PC via updated IP address. The following describes the method which had been devised to solve this problem.

2.5.1 NAT Overview

Below is a summary of different form of NAT.

- Full Cone: A full cone NAT is one where all requests from the same internal IP address and port are mapped to the same external IP address and port. Furthermore, any external host can send a packet to the internal host, by sending a packet to the mapped external address.
- Restricted Cone: A restricted cone NAT is one where all requests from the same internal IP address and port are mapped to the same external IP address and port.
 Unlike a full cone NAT, an external host (with IP address X) can send a packet to the internal host only if the internal host had previously sent a packet to IP address X.
- Port Restricted Cone: A port restricted cone NAT is like a restricted cone NAT, but the restriction includes port numbers. Specifically, an external host can send a packet, with source IP address X and source port P, to the internal host only if the internal host had previously sent a packet to IP address X and port P.
- Symmetric: A symmetric NAT is one where all requests from the same internal

IP address and port, to a specific destination IP address and port, are mapped to the same external IP address and port. If the same host sends a packet with the same source address and port, but to a different destination, a different mapping is used. Furthermore, only the external host that receives a packet can send a UDP packet back to the internal host.

Each connection from the internal (private) network to the external (public-Internet) network, and vice versa, is tracked and a special table called NAT table is created to help the router determine what to do with all incoming packets on all of its interfaces.

The NAT table works differently depending on the NAT mode. Figure 2.20 shows the NAT table for Network Address Overload Mode.

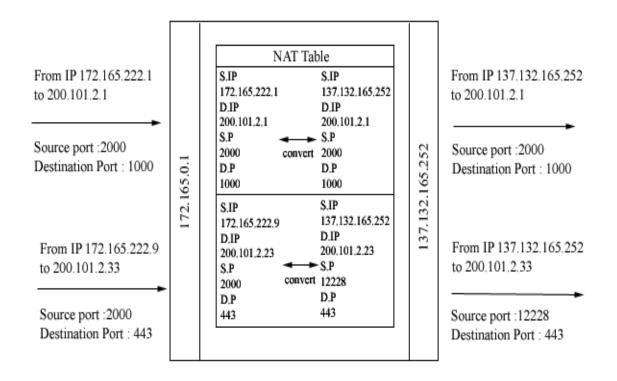


Figure 2.20: NAT table

2.5.2 UDP Bombarding

In this method, the Agent will send UDP at every one minute interval to the server that is located outside the LAN network. At the NAT device, a NAT table as shown in Figure 2.20 is created. The Agent will operate at the port number assigned in the UDP packet Source port. This will allow the Agent to be contactable by an external computer, as the LAN PC and the port at which the Agent run have an entry in the NAT Table. From the UDP packet received, the Server will be able to determine the translated address given to the LAN PC. The WAP gateway will contact the server in the same manner, as describe earlier, to retrieve the translated IP address to the Agent. Figure 2.21 shows the working of this mechanism.

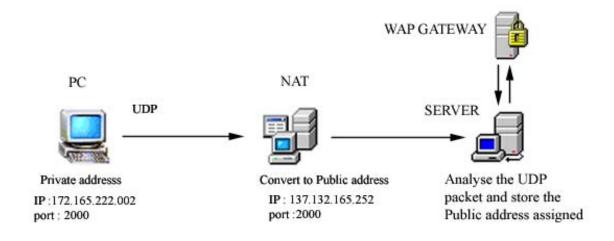


Figure 2.21: UDP Bombarding

A few points to note about this method are elaborated below:

a) The NAT table will keep the entry for certain duration of time known as timeout

period. If the NAT device after the timeout period did not receive new packet from the PC it will remove the entry. Thus there is a need to keep sending UDP packet at regular interval.

- b) For this method UDP is used instead of TCP. This is because UDP is a connectionless protocol thus it does not bind to the source port that is indicated in the UDP packet send out. This allows the Agent to run on the source port number of the UDP packet.

 TCP connection would have prevented the Agent from running at the same port.
- c) This mechanism will only work for the LAN that has NAT operate at the Full cone mode. This is due to the fact that only in Full cone mode does the NAT allows transmission of packet from a third party external host (WAP gateway).
- d) Depending on the vendor, some WAP gateway only allows connection on port 80, 8080 and 443. And in LAN there might be a firewall to block this port number due to security reason thus this approach will not be able to work in this situation.

2.6 Secure Wake-On-Ring

As mentioned, it is not realistic to assume that the target PC is power up at the time of access. Therefore, provision has to be made for such situation. The solution comes in the form of an authentication black box as shown in Figure 2.22, that make use of the Wake-On-Ring feature in the motherboard, to boot up offline computer devices. The authentication black box works on the dial-tone principle to authenticate before booting up the computer devices via Wake-On-Ring function. The authentication black box

provides the power to discriminate all incoming calls and the flexibility to implement certain courses of actions that a user desires. With the box, only authenticated calls will be allowed to boot up a computer.



Figure 2.22: Authentication black box

The general functional block diagram for the authentication black box is illustrated in Figure 2.23.

2.6.1 Ring Signal

The ring-detection module is necessary for the box to detect an incoming phone call so that the ring pick-up circuit can be activated to get the phone off-hook. When a ringing signal is detected on the line, a small current, this circuit will output a HIGH signal to inform the microcontroller of an incoming call.

Once a ring is detected, the microcontroller will hold an "Answer Phone Signal" high, which is used to drive a relay as depicts in Figure 2.24. When the relay is turned on, a

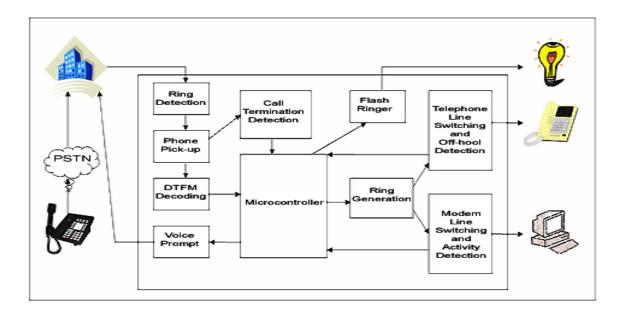


Figure 2.23: Authentication black box block Diagram

line across a 1:1 transformer connects, thereby drawing sufficient current from the phone to get the phone off hook. This is equivalent to getting a telephone handset off its cradle. The phone signal is then passed on to a DTMF detector through the 1:1 transformer. This is akin to listening through a handset. This transformer serves to isolate the rest of the circuitry from the telephone line which might potentially carry high voltage signals that will affect the circuitry. A line is tapped out at the input to the transformer to feed into a call termination detection circuit, which will detect when the caller puts down the phone. At the other side of the transformer, the output from a voice chip is fed in so that voice prompts can be heard at the caller end.

Since the caller can terminate the call at any moment, the box needs to detect this event and promptly end the session and prepare for the next call. When the box is connected to the caller, the voltage between the ring and tip will be around 6V. Upon call termination of the user, this voltage will dip to 0V briefly before rising to 6V again.

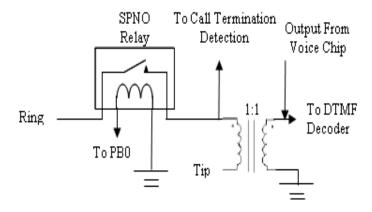


Figure 2.24: Phone Pick-up Circuit

Thus, by detecting this dip in voltage, the microcontroller can be informed of when the call has terminated and perform the necessary tasks.

The box receives commands from the caller by listening to the DTMF tones from the output of the 1:1 transformer as previously stated. When a tone is detected, a DTMF decoding chip will decode the tone and send signals to the microcontroller. The microcontroller then reads and interprets these signals. The strobe signal output from the decoding chip goes high whenever a tone has been decoded and is held high until the input tone has been removed. By feeding this signal to the microcontroller, it can be used to monitor when a caller presses and releases a key.

The ring generation circuit, as shown in Figure 2.25, is used to ring the telephone, as well as to create a ringing pulse to be detected by the modem to boot up the computer. After the phone pick-up circuit, the call is established and the line will be held by the authentication box. Since the line has already been established, it is not possible for another ringing signal to be generated by the telephone company.

It is necessary to replicate this AC 75V and 24Hz ringing signal in order to make a telephone ring. This is achieved by passing an AC 9V signal through a step-up transformer thereby producing an AC 75V signal. Since the mains frequency in Singapore is 50Hz, the signal output from the transformer will also be at 50Hz. This higher frequency has no noticeable impact on the ring produced.

The microcontroller uses an output port to decide whether to ring the phone of the modem. To create the pauses in between the ringing tone, another output port is used to control a relay. Closing this relay sends the ringing signal to the computer or the modem. Similarly, opening this relay creates the short pause between 2 rings and a longer pause after every 2 rings.

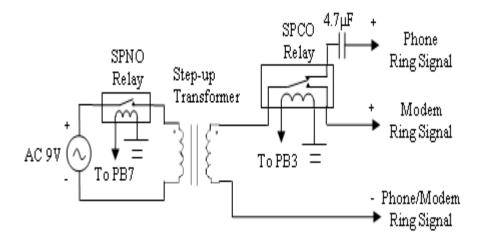


Figure 2.25: Ring Generation Circuit

2.6.2 Telephone Line and Modem Line

For a normal call, the ring generation circuit will ring the telephone. Since the ring signal is controlled by the microcontroller, there is a need for a mechanism to detect when the telephone is answered so that the microcontroller can terminate this ring signal. This is achieved by the "Telephone Off-hook Detection Circuit" as shown in Figure 2.26.

In the telephone Line Switching circuit, the telephone is normally connected to the PSTN through the box. This allows the telephone to be used normally when the box is not taking any calls. When the caller chooses to ring the telephone; the microcontroller will set an output port high, causing a relay to switch. The telephone is now disconnected from the PSTN and connected to the phone off-hook detection circuit. The phone off-hook detection also takes in the phone ring signal from the ring generation circuit and it has the dual task of ringing the phone and detecting when the phone is picked up. Once the telephone is picked up, the call will be connected.

When a telephone gets off-hook, it is equivalent to adding a resistance across the telephone, as its internal circuitry will now be connected. Given this fact, a DC 5V is fed into this circuit through a resistor. This causes a potential drop to be detected across the resistor whenever the telephone goes off-hook. After the microcontroller is informed of the phone off-hook status, it connects the phone to the PSTN and gets the box off-hook. The box will now be ready for the next phone call once the current call is terminated. The call termination in this case is handled by the telephone company since the caller now has a direct connection to the phone.

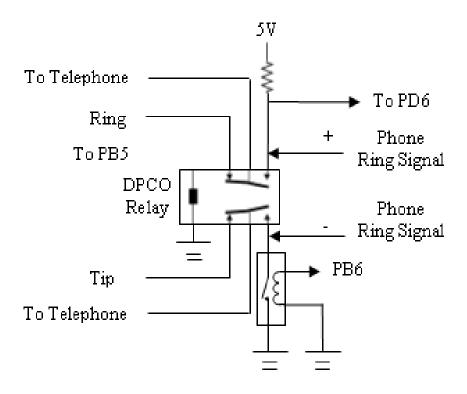


Figure 2.26: Telephone Off-hook Detection Circuit

The modem line module switches the modem between the PSTN and the ring generation circuit. The modem activity detection circuits consist of a modem online and offline detection. Their function is to detect the respective activities and switch the modem connections accordingly.

The modem listens to the telephone line and boots up a computer as soon as it senses the ring signal. This process occurs at a very high speed and before the authentication box can pick up the call, the modem would have sensed the ring signal and boot up the computer. Since the main function of the authentication box is to screen incoming calls, the computer booting up before the box can start operation will make the box obsolete. Thus there is a need for the box to isolate this ring signal from the modem.

When the box is not taking any calls, the modem will be cut off from the telephone line. This ensures that any incoming call will not reach the modem. In this state, the modem will be connected to the output from the ring generation circuit, which supplies the modem ring signal. When the caller chooses to boot up the computer and upon successful authentication, the modem ring signal will pass through the relay and ring the modem.

Although isolating the ring signal solves the problem of the modem sensing the ring signal before the box does, cutting off the modern from the telephone line means that the modem will be unusable for normal operation when the box is not authenticating any calls. This is undesirable as one of the design goals is the usability of the modem. The modem activity detection thus comes into play by sensing when the modem needs to be used, and subsequently connect the modem to the telephone line. Since the modem needs to go off-hook whenever it needs to be used, sensing this off-hook state will inform the microcontroller when to connect the modem to the telephone line. Similar to the phone off-hook detection, the circuit shown in Figure 2.27 makes use of a DC 5V supply across a resistor. When the modem is online, the drop in voltage across the resistor will be detected by the microcontroller, which then switches the connection of the modem to the telephone line. Since the modem needs to be disconnected to the PSTN when it is not in use, there is a need to detect when the modem goes offline so that the microcontroller can disconnect the modem.

When the modem is online, the voltage across the ring and tip lines is 7V. When the

modem goes offline, the voltage across the ring and tip lines will rise to 70V. This circuit thus senses for this rise in voltage using a relay, and informs the microcontroller, which will then subsequently disconnect the modem from the PSTN.

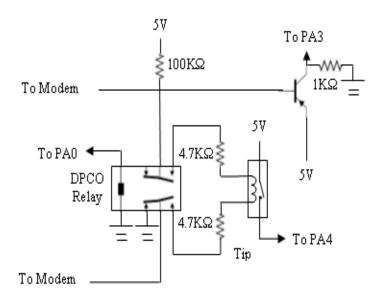


Figure 2.27: Modem Online and Offline Detection Circuit

2.7 Conclusions

In this chapter, a comprehensive architecture had been proposed to establish link between the mobile device and computer device. The architecture adopted a distributed instead of conventional centralized approach in connecting the mobile devices with the computing devices. Comparison had been made between the proposed system and some of the current methods. Solutions for WAP timeout and LAN problems have been described in detail. A hardware solution for power off PC that works on the dial-tone principle to boot up the computer devices via Wake-On-Ring function is constructed to enable the mobile devices to activate the target computer remotely. A mobile email with attachment application based on the architecture had been implemented to verify the viability of the proposed system.

Chapter 3

Development of a Mobile Spreadsheet-based PID Control Simulation System

3.1 Introduction

M-learning applications may be classified into different categories. For example, in [36], M-learning activities are classified into six broad groups: a) Behaviorist - promote learning as a change in learner action; b) Constructivism - learners actively construct a new concept based on previous or current knowledge; c) Situation- promote learning within an authentic context and culture; d) Collaboration - promote learning through social interaction; e) Informal and lifelong - support learning outside a dedicated learning environment; and f) Learning and teaching support - assist in the coordination of learners and resources for learning.

Because of the inclination towards collaborative learning in the field of learning theory, most of the M-learning applications developed belong to the Situated and Collaborative category. Researchers in collaborative M-learning projects [37]-[39] make use of the

wireless connectivity of the mobile handheld to facilitate the sharing of data and communication among the team members. These collaborative projects commonly involve the use of expensive high-end handhelds (e.g., PDAs) [40], and usually the handheld is a loan set rather than an owned one. Ownership of the handheld will provide a sense of personal commitment to the device, thus increasing the willingness of the user to explore its potential.

The behaviorist learning approach still remains the popular approach utilized in many E-learning systems. Also a trend exists in transferring the legacy system from E-learning to M-learning. A major obstacle faced is in the limited processing speed and data storage of the handheld. A specific M-learning project adopting the Behaviorist approach is the BBC Bitesize [41]. Bitesize is a Java-based application that provides revision materials via the mobile phone. However, a prominent implementation problem highlighted in the deployment of Bitesize is the compatibility issue. Despite Java offering generally a cross platform environment, is still difficult to operate Bitesize on all common handhelds.

In this chapter, the proposed system will aim to act as an optional short pre-experiment and revision exercise. Most of the work is expected to be performed individually. Therefore, the Behaviorist learning approach is the most suitable implementation approach for the system and is duly adopted. The proposed system uses a Thin Client model as opposed to a Java-based model. The merits of the two models will be discussed in a later section. The proposed system conforms to the guidelines suggested in [42], [43] for a good M-learning application in terms of costs, choice of technology, service/applications,

and adaptability of mobile technologies.

3.2 Target Course

"Industrial Control Systems" is an elective control engineering course offered in the National University of Singapore to third and fourth year electrical undergraduate students. The aim of this course is to equip the students with the necessary background and skills for industrial control engineering and practice. A wide range of topics useful for practicing control engineers includes sensors and instrumentation, industrial control practices and advanced control, and discrete event control using *Programmable Logic Controllers*, all covered in this course. A brief description of the course syllabus is as follows.

1. Input/Output Conversion and Signal Processing

Application-oriented description of transducers and actuators; signal-conditioning techniques applicable to measurement and control of most commonly used process variables; conversion techniques focusing on sample rates, accuracy, resolution, linearization, isolation, and noise filtering.

2. PID Control and Tuning

PID Control, Implementation of digital PID controllers; integral action and antiwindup strategies; derivative action and derivative gain limits; manual tuning methods and applicability; Cohen-Coon step response refinements; Ziegler-Nichols frequency response method; limit cycle (relay) method; PID implementation aspects.

3. Evolution and Types of Control Systems

Evolution of control configurations; centralized control, direct digital control, DCS; fieldbuses.

4. Advanced Control Techniques

Applications and implementation of feedforward control, cascade control, ratio, selective.

5. Programmable Logic Controllers

Introduction to sequence control; relay and PLC; relay ladder diagrams; PLC ladder program; function block diagrams. Sequential function chart.

For such a course, a practical-oriented delivery of the course materials is important and necessary. A series of four experiments are formulated to enable the students to consolidate the key concepts covered by the lecturers through realistic hands-on sessions in the laboratories. These experiments involve real-time instrumentation and control design for real pilot setups in the laboratories.

Traditionally, the course has a yearly intake of about 160 students. With limited laboratory resources, this means the laboratory schedule is rather tight, and laboratory classes run one after another throughout the day to enable the large class of students to complete the intensive laboratory work. Therefore, preparatory exercises prior to the actual experiments are extremely useful to ensure that the most learning experience on real-time control can be reaped from each session. Such a phenomenon is observed in

most courses with a high level of student enrollment or in courses for part-time students where only the evenings are available for laboratory work. For these courses, E-learning can possibly offer an excellent alternative to physical presence in the laboratory.

One such useful preparatory exercise in the context of the "Industrial Control Systems" course will be on the modeling and simulation of control systems. As field reports show, control operators are having problems tuning the PID controller to perform adequately. Therefore, students, encountering PID control tuning for the first time, must receive adequate training in the controller tuning and are able to appreciate the effect of changes caused by each of the PID parameters on the control performance. A key consideration governing the design of such a preparatory exercise in a large class setting is that students should be constrained by time and space as little as possible, so that they are able to initiate the exercise anytime they are available and anywhere they may be at that time, with a mobile device that is commonly owned.

This serves as the motivation for the development of a mobile simulator for the course to allow the student to tune a PID controller iteratively based on a plant model to reach a desired level of control performance using only a preprogrammed spreadsheet. The spreadsheet can receive parameters input from the student in the form of model and control parameters, and give a graphical output in the form of control performance results. The requirements to run this self-preparatory exercise appear modest at the time of development, but real student feedback will be sought upon implementation to verify this assumption.

The duration of the entire course spans over 13 weeks. The laboratory experiments begin from the seventh week, and students are encouraged to use the mobile PID simulator for pre-experiment preparation on PID tuning prior to attending the laboratory classes on real-time control. At the end of the course, the students will be surveyed on the usefulness of the system in enhancing learning. The survey results will be duly furnished in later Section.

Laboratory Experiment

In an experiment which is part of the course, the students are required to control the temperature of a thermal chamber to within provided specifications.

The experiment is conducted using the thermal chamber as shown in Figure 3.1, which typifies a system with a time delay. The thermal chamber is a transparent plastic chamber mounted on top of a National Instruments Signal Condition Unit (SC-2345). A thermocouple inside the chamber is used to sense the temperature, and it is connected to a thermocouple input module inside the SC-2345 unit. In addition to this, two feed-through modules for the PWM control of a lamp and a fan inside the chamber are connected via a 6-position terminal plug cable. The temperature inside the chamber can be varied by controlling the light intensity of the lamp or the speed of the fan. In this experiment the focus is on single input and single output configuration, thus the fan speed is kept constant, and the temperature fluctuation inside the chamber will largely depends on the light intensity of the lamp.

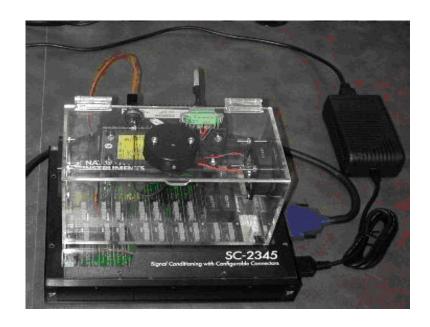


Figure 3.1: Thermal chamber

Both feedforward and feedback controllers are to be designed and tuned by the controller to yield a specified set-point tracking performance. As the time available for a thorough PID tuning is limited for such a large class, a pre-experiment session for students to hone their skills independently will be very useful, prior to the actual experiment. Therefore such sessions need to be initiated by students according to their own schedules, on an anytime and anywhere, low cost basis. A mobile PID simulator will serve this purpose, and the development of such a system will be elaborated in the next section.

3.3 Mobile Excel PID Simulator

In this section, the development of the mobile simulator will be elaborated in detail.

The problems encountered and proposed solutions presented will enable the final and viable on-the-go simulation exercise for students using handheld mobiles. This section

should be useful to readers who would like to develop a similar system for their courses.

3.3.1 Process Model and Controller

Simulation with a spreadsheet falls into two main categories [44]: 1) sampling or Monte Carlos simulation, and 2) process simulation. In this section, a process simulation in which Excel is used to simulate a PID controller is described.

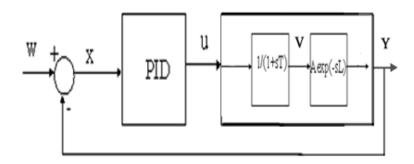


Figure 3.2: The PID control system

Figure 3.2 illustrates the PID control system. W represents the setpoint signal; Y is the process output; X = W - Y represents the tracking error; and U is the control effort.

The Excel simulator [45] models a digital PID controller and simulates its performance in a simple control loop. The process model used in this experiment is a standard "first order lag plus delay" model used in many process control applications:

$$Y(s) = A \frac{e^{-sL}}{1 + sT} U(s), \tag{3.1}$$

where T is the time constant; A is the gain; and L is the time delay of the process. The PID controller transfer function is as follows.

$$U(s) = G(1 + \frac{1}{sT_i} + sT_d)X(s), \tag{3.2}$$

where G is the proportional gain; T_i is the integral time constant; and T_d is the derivative time constant.

The model and controller functions can be converted to difference equations in the discrete time domain as given below.

$$y_k = y_{k-1} + (u_{k-n} - y_{k-1}) \frac{h}{h+T}, (3.3)$$

$$u_k = Gx_k + \sum_{j=1}^k \frac{Gh}{T_i} x_j + \frac{GT_d}{h} (x_k - x_{k-1}), \tag{3.4}$$

where h denotes the sampling interval and k represents time at t = kh. In these forms, the process and controller models are ready to be programmed into the spreadsheet. The mobile Excel simulator should allow the user to set the system input A, L, and T to fix a certain process model. Then, by varying the controller parameters (G, Ti, and Td), the spreadsheet, as illustrated by Figure 3.3, has been preprogrammed to compute automatically and display the response of the control system to step changes in the setpoint. As depicted in Figure 3.3, each column of the spreadsheet is devoted to computation of a signal (or a part of it) based on the difference equations. Each row of the spreadsheet will thus contain the various signals in the loops at a certain time

instant. As time progresses, the spreadsheet will expand row-wise accordingly.

Different performance characteristics in terms of response speed and overshoot of the process model to step changes in the setpoints can be observed because of each combination of PID parameters input by the students. A typical response is shown in Figure 3.4. The students are expected to adjust the PID parameters to achieve a good compromise between speed and stability.

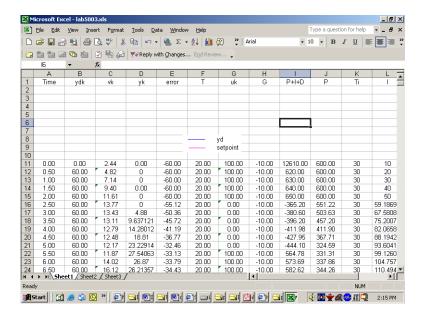


Figure 3.3: Preprogrammed Excel spreadsheet

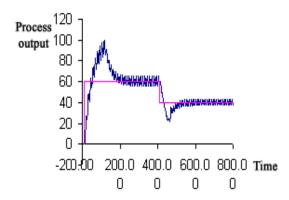


Figure 3.4: Response to step changes in setpoint

3.3.2 Application Models

In this sub-section, possible application models for such a system will be briefly reviewed and the final selection will be justified with respect to the objective application.

Thin Client

The Thin Client is a form of browser-based development model; it was first introduced in the 1990s as an alternative to PC [46]. Under the Thin Client model, all the application processing is performed on the server-side. This enables the devices on the Client's side to work with minimal requirements on disk storage and RAM capacity. This feature makes Thin-Client a prevailing model adopted by the first wave of mobile application developers.

Thick Client

With the increase in the disk storage and processing power of handheld devices, there is a steady shift of applications moved from the Thin Client to the Thick Client model. The Thick Client model, the application caches more data on the handheld devices and utilizes more of its processing power. Thick Client can be broadly classified into three types: 1) native application 2) smart client 3) mobile application framework all of which can run in one or more of the following three modes: a) standalone b) wireline, and c) wireless. The smart client application development using J2ME is the most popular implementation of the Thick Client. In the subsequent section, without loss of generality, the Thick Client will be considered to be implemented as J2ME Smart Client.

Comparison

The current trend indicates that there is a steady increase in the number of developers who use the Thick Client model [47]. On the other hand, a substantial number of developers who still rightly advocate the use of the Thin Client model for the development of certain mobile applications [48], where the model is still the most appropriate. In this sub-section, a comparison of the two models will be presented, and the final selection of the Thin Client model will be justified with respect to the objective application of the proposed mobile, spread-sheet simulator.

1. Administrative Effort

The Thin Client model has a very efficient deployment mechanism; any handheld devices with a WAP browser will be able to use the Thin Client application without any prior software installation. The Thick Client application, on the other hand, requires the download and installation of software. With a number of downloading channels now available (e.g., Bluetooth, infar-red, WAP), this additional deployment step may be manageable if the software is stable. The user will only need to follow these steps once. However, the mobile spreadsheet simulator is a new development; addition of new functionalities and correction of bugs are widely expected to occur. Frequent update of the Thick Client software will become a hassle to the student user and further deter the use of this application. Other significant problems, such as a different version of installation will likely occur. The Thin Client approach will be spared from these problems since the application resides on the server. Therefore, a Thin Client approach will incur a much lower

administrative effort than a Thick Client one.

2. Portability

"Write once, Run Anywhere" is usually associated with applications based on Java. However, this has not turned out to be true for applications which will run on different makes of devices. Different device vendors will have a slightly different implementation of MIDP. Thus, Java programs suitable for the Nokia phone can be unusable in an O2 phone. Moreover, different handhelds will have different form factors; a Thick Client approach will not be able to optimize the full screen. With a Thin Client model, which is browser based, customizing the display to a different phone will be easier through the use of a database and a user profile.

3. Development Effort

A spreadsheet with good formulation and cell interaction function is much easier to program for simulation as compared to J2ME which is only armed with basic math functions. Mobile Client-side development is also more tedious and tougher to debug than developing programs on desktop computers. In most cases, the Thin Client model will have a much shorter development cycle when compared to the Thick Client model.

4. Disconnected Access

For Thin Client application to be used, the handheld has to be connected within the coverage area. On the other hand, Thick Client applications can work offline as a stand-alone. Stand-alone bestows upon users a greater convenience while saving in cost. However, in certain countries, such as Singapore, GPRS coverage is island-wide and the cost in using GPRS is low and becoming lower.

5. User's Experience

Java-based Thick Client applications provide richer graphic contents than browser-based Thin Client applications. Image processing techniques (which will be described in detail later) can be used to mitigate the difference. The responsiveness of the application is also an important consideration. Since a Thick Client application runs on the mobile device locally, it usually allows faster user interaction than a Thin Client counterpart. However, in this case of a mobile spreadsheet simulator, the converse is true because a large amount of recursive floating point calculations are expected to be executed in the simulator. In an experiment conducted in [49], repeated floating point calculation uses a significant amount of time in most cellular phones.

6. Memory and Storage Requirement

Most new generation handhelds have a sizable memory and disk storage allocation that is sufficient for a Thick Client implementation of the mobile spreadsheet simulator. However, for future expansion, a more complicated and comprehensive control simulation will pose a problem since the amount of raw data will be much higher than what is implemented in this project. Thin Client implementation will not pose this bottleneck to future developments, since the data resides on the

server side.

7. Scalability

The Thick Client allows the application to be used on a standalone mode, which will result in fewer users accessing the server at the same time. This reduction in the amount of resource sharing increases the performance of the application. The Thin Client application has to be connected to the server for usage and is, therefore, less scalable than the Thick Client. As the number of users executing the application at the same time is not expected to be large in this case, scalability is not a crucial problem.

In summary, in the context of the mobile spreadsheet simulator and the objective of encouraging a big pool of students to use the system even if they own vastly different handheld devices, the advantages and convenience offered by a Thin Client approach outweighs any associated with a Thick Client one. Thus, the proposed mobile spreadsheet simulator is implemented using a Thin Client model.

3.3.3 Architecture of Mobile Excel Simulator

Figure 3.5 shows the simplified architecture of the mobile spreadsheet simulator. In Figure 3.5, the handheld mobile device and agent exhibit a Thin Client relationship through the use of a WAP browser. Adopting the Thin Client model has several advantages. First, the handheld mobile device does not need to install any program; the system development is simplified since the constraint imposed by the low hard disk size

of the handheld mobile device is removed. Secondly, the interoperability is enhanced since different handheld mobile devices with the same type of WAP browser will be able to view the output of the server correctly. Thirdly, in Thin client there is no active connection established between the server and the handheld mobile device that will hog the resources of the server. Therefore, many users can log in the system simultaneously as if it is a web server. Two novel methods on converting Excel to image and image processing will be elaborated in the next section.

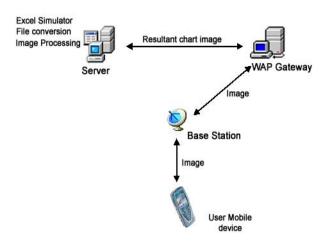


Figure 3.5: Mobile spreadsheet architecture

3.4 Excel File Conversion

Based on the statistics obtained from [50], the conclusion that more people use mobile phones than PDAs. Most handheld mobile devices have limited memory space and are unlikely to be able to display the Excel document in the form of a chart. Therefore, the Excel document needs to convert to an image format that can be displayed in all

handheld mobile devices.

An initial approach adopted has been to use the commercial printer driver software to convert the Excel document to an image file. However, the conversion time is slow, and the WAP gateway timeout [51] becomes a serious problem. Instead, a faster and more efficient method has been devised to convert the resultant chart from an Excel simulator to an image format for display on the handphone. The method will be elaborated in details in the following subsections.

3.4.1 COM and Clipboard

Excel is a compound document. Compound documents have the OLE feature that allows raw data to be copied to the clipboard and enables parts of two documents to be merged into one [52]. Through the clipboard, there a means exists to convert an Excel file or part of it to an image file easily. When part of the compound document is copied to the clipboard, the object is stored in one or more of the clipboard formats. One of the clipboard formats that the object is stored in is the EMF format. EMF, as the name suggests, refers to metafile graphics which contain both raster and vector data. The object is stored as a vector with bitmap attributes. Using Windows graphics and multimedia API, the EMF file can be converted to a bitmap file.

To automate this clipboard copying process, the COM server is used. COM defines a standard communication interface between the server and client module.

3.4.2 Excel to Bitmap

Two methods control Excel from the server application: early binding and late binding. Binding refers to the process of linking the function calls in the developed application and the code in the Excel application. Binding replaces the function names in the developed application with the memory addresses in the table of pointers known as v-table. Programming the pointers in the v-table is complicated and error prone. A programming environment, such as Delphi and Visual Basic, provide matching of the function calls behind the scene, thus greatly reducing the complexity and development time. The main advantage of early binding over late binding is the speed of execution. Early binding executes at least twice as fast as late binding. This advantage is the result of direct linking of the application and function address in the COM, thus there is no overhead in runtime lookup as in the case of late binding. Late binding, on the other hand, is version independent, since the exact COM object need not be known at design time. For example, to use EXCEL2002, EXCEL95-type library must be used in early binding, but the version does not matter for a late binding approach. The type of binding to use will thus depend on the nature of the application. After the application has completed the binding to the Excel COM, the application will be able to use the functions present in the Excel application. The application can then specify the parts of the Excel spreadsheet that is to be copied to the clipboard. The parts that are copied to the clipboard are in several different types of clipboard formats, e.g., CF_METAFILEPICT and CF_TEXT. Using the graphics device interface API, the EMF representation of the copied part can be converted to a bitmap file. By iteratively checking the number of active workbooks and worksheets in the Excel file, all the worksheets and charts in the file can be converted to a bitmap file. Figure 3.6 shows the concept of using the clipboard to convert the Excel file to a Bitmap file diagrammatically.

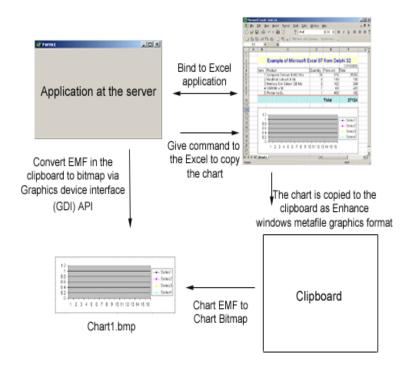


Figure 3.6: Clipboard conversion. The figure illustrates the use of clipboard to convert a chart in an Excel file to bitmap format. Note that besides the enhanced Windows metafile format, the chart is also stored in other clipboard formats.

3.5 Image Processing

In the raw form, the bitmap file obtained from the file conversion process is not adequate to be displayed on a mobile phone for the following reason.

1. The image size, in terms of pixel height and width, is too large for the cellular phone display screen. After conversion, the image size is usually several hundreds

pixels in dimension. The average cellular phone display screen is only about 200 x 200 pixels. Thus, resampling is needed.

- Because of resampling, the image will become less refined and the words may be distorted beyond recognition. Thus, further image processing is needed to refine the resampled image.
- 3. Although the cellular phone can support a wide range of image format, the service provider usually only allows the JPEG or the GIF to be transmitted.

3.5.1 Resize and Resampling

In image processing, the resampling approach is preferred over a resizing approach. When the image is enlarged using resizing, some of the pixels are duplicated. When the image is reduced using resizing, some pixels will be discarded. Hence, thin lines and fine features may be lost in the resized image. In the case of resampling, the pixels are recalculated using mathematical interpolation techniques. The resultant image produced by resampling is of better quality than that produced by the resizing method.

Nearest Neighbour, Bilinear, and Bicubic approaches are three interpolation methods commonly used in a resampling method [53].

- 1. The nearest neighbour method determines the pixel at the resultant image from the neighbour pixel at specified coordinates of the input image.
- 2. The bilinear method uses the colors from four of the nearest neighbour pixels of a specified point.

3. The bicubic method uses eight surrounding pixels colors, taking the average of it for the pixel at the resultant image.

Of the three methods, the Nearest Neighbour interpolation method is fastest. However, the nearest neighbour interpolation method produces a lower quality image. The Bicubic method produces the best resultant image, but requires at a longer processing time. For the target application, the Bicubic interpolation is used because the delay incurred is acceptable.

3.5.2 Resampling and Gaussian Blur

Although resampling using Bicubic interpolation has produced relatively good quality images, undesirable artifacts (e.g., distortion of a straight line), still exist. As most alphabets contain fine straight lines, the viewing of the resultant images poses a serious problem. The distortion can be critical when the scaling factor is high.

Gaussian blur is adopted to smooth these artifacts. In the gaussian blurring process, each pixel in the RGB image is transformed using the characteristics of its adjacent neighbouring pixel with a gaussian probability based on proximity. Thus, the nearer pixel will have a greater influence on the final transformed pixel. The amount of affected pixels on either side is determined by the convolution radius specified. By blurring the image before performing resampling, the fine details are "enlarged", and they remain visible even after resampling. The size of the convolution radius used is dependent on the resampling scale factor. Figure 3.7 shows the display of the Excel simulator chart on a mobile phone.



Figure 3.7: Mobile Excel Simulator. The figure shows the Excel simulator results displayed on a cellular phone

The subjective fidelity criterion is used to find the best fit convolution radius size. Using Table 3.1 as the rating scale, a group of 40 people were interviewed to rate the image displayed on a cellular phone (Figure 3.8, Figure 3.9, Figure 3.10 and Figure 3.11). The details are discussed in the next section.

Table 3.1: Rate Scale of the Television allocations

Value	Rating	Description
1	Excellence	An image of extremely high quality, as good as you could desire
2	Fine	An image of high quality, providing enjoyable viewing, interference is not objectionable
3	Passable	An image of acceptable quality, interference is not objectionable
4	Marginal	An image of poor quality, you wish you could improve it. Interference is somewhat objectionable
5	Inferior	A very poor image, but you could watch it
6	Unusable	An image so bad you could not watch

3.5.3 Subjective Fidelity Criteria

There are two general classes of criteria to assess the quality of a image. (1) Objective fidelity criteria and, (2) subjective fidelity criteria. Objective fidelity criterion which

involves the use of complex mathematical model, that is beyond the scope of this thesis. Below describe how the subjective fidelity was performed and the result is shown.

For the survey, 720 x 540 pixels of a converted PowerPoint image is blurred using gaussian blur before it is resampled to an image size of 130 x 90 pixels. The resultant image is displayed on a cellular phone with a small display screen. A bar chart, as shown in Figure 3.12, was plotted based on the survey results. From the subjective fidelity criteria test, it is concluded that a convolution radius of three produces the best viewable image in this example. As the convolution radius increases beyond four, the displayed image begins to blur. The convolution radius is dependent on the original image size and the resampling scale factor. In general, a smaller image size and scale factor requires a smaller convolution radius.

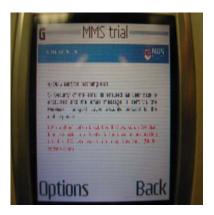


Figure 3.8: The convolution radius used is 1

3.5.4 Zoom Function

Despite the image processing effort, the readability of the resultant image can still be further improved. To this end, a zoom function is introduced. The zoom function allows the user to enlarge the area of interest by selecting one of the regions 1, 2, 3 or 4 from



Figure 3.9: The convolution radius used is 2



Figure 3.10: The convolution radius used is 3

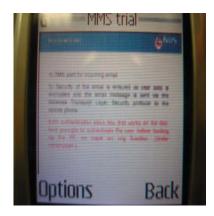


Figure 3.11: The convolution radius used is $4\,$

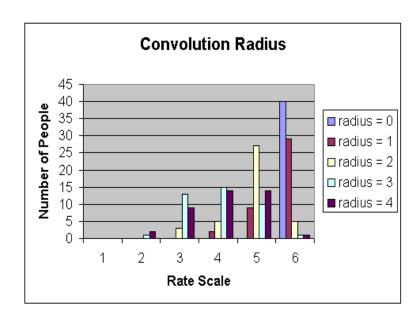
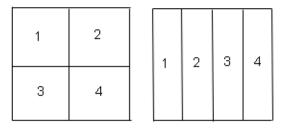


Figure 3.12: The bar chart shows the results obtained from a survey of 40 persons rating the quality of image according to subjective fidelity criteria.

Figure 3.13(a) or Figure 3.13(b). The selected region will then be cropped from the original image. As the cropped image has a lesser number of pixels, the resampled scale factor will be small and the image will appear magnified as shown in Figure 11. Figure 3.13 shows 2 possible ways to split the original image. It is found that the method depicted in Figure 3.13(a) is preferred when dealing with picture-based image, while the method depicted in Figure 3.13(b) is more suitable for text-based image.

3.6 Students' Feedback

In this section, the feedback from students after the experiment will be presented and analyzed. From the feedback of the students, the contribution of the system to the student understanding in PID controller is evaluated. Valuable information pertaining to the general background of students and the benefits reaped from the system are thus obtained which are useful to refine the system and to encourage more students in the



[a] [b]

Figure 3.13: Figure 10 (a) on the left shows the original image being split into four equal regions over two rows and two columns. (b) on the right shows the image being split into four rectangular regions on a single row.



Figure 3.14: Using the zoom function, the processed image is enlarged.

subsequent cohorts to use it.

3.6.1 Methodology

The feedback study is conducted using an electronic questionnaire. The questionnaire includes three sections.

- (A) In this section, questions relating to the background information, such as the type of handheld mobile devices owned, frequency of WAP usage, and the study modes of the students, are asked. The questions are designed to evaluate the potential of the students with respect to resources available to use the system.
- (B) The questions in this section are designed to evaluate the adequacy and the performance of the system. In particular, feedback is sought on the quality of the delivered image and the usefulness of the system with respect to the course objectives.
- (C) Open questions are fielded in this section, targeting the students who have not used the system, urging them to explore the deterrent factors discouraging the students from using the system.

Eleven questions are in the questionnaire with most involving Yes/No answers or quantitative based on a five-point Likert-type scale, although open questions on qualitative aspects of the system are also present.

3.6.2 Procedure

The mobile PID simulator is introduced to the students immediately after the concepts have been covered in the lecture. Although the simulator represents an optional exercise, the students are strongly encouraged to make use of the simulator to consolidate PID tuning concepts and pre-preparation and investigation for the course laboratory experiment. The instructions for the simulator are sent to all the students registered for the course via email. To encourage mass usage of the system, the steps to initiate the exercise are direct, intuitive, and clearly instructed.

1. Enter the following WAP site into the mobile devices.

http://www.domus.nus.edu.sg/excelsim/servlet/input

2. Enter non-zero value for the following parameters.

System

System Gain (A)

Delay (L)

Time constant (T)

Controller

Controller Gain (G)

Ti

Td

3. Click on **Submit** to view the simulated graph.

4. To enter new values, click on "Input parameter".

Following the completion of all the laboratory experiment sessions, an electronic questionnaire is sent to all the registered students via email. The data collected is analyzed with statistical tests using MATLAB [54].

3.6.3 Survey Subject

There are 160 students registered for the course. Seventy-two students respond to the questionnaires sent to them. These 72 students who respond form the subjects of this study. The subjects are categorized into two main groups, students who have used the system and students who have not used the system.

3.7 Results and Discussion

3.7.1 Section A: Current Scenario

Figure 3.15 shows the questions contained in Section A.

Questions in Section A are to be answered by all the students, regardless of whether they have used the system. The survey results show that 75% of the students own mobile devices with configurations suitable to access and use of the mobile Excel simulator. However, from the survey although the survey that although the students have the

Questionnaires on Mobile PID Simulator Based on EXCEL

The mobile PID simulator is a first step towards more mobile e-education tools which we may develop in future where it can be useful to our students. If you can kindly take a few minutes to feedback on your experience and suggestions, it will help us in these works. You can email the completed form directly to me or drop it in the lab's pigeon hole together with your E1 or E3 report. Thank you.

Section A (for everyone even if you do not try the mobile simulator)

-	u currently owr Yes	_	y handheld m	obile device (e.g., PDA:	s and cellular			
2) How o	ften do you stu	dy while takin	g public trans	port? (on a scale of 1 – :	5)			
	2 At least 1							
	per month			Dany				
3) Do you use the WAP function on your mobile phone? Yes No,								
If yes, how often you use it. (on a scale of $1-5$)								
1	2	3	4	5				
	At least 1 per month			Daily				
If no, wh	ıy?							
	have suggestion se do share wit		obile device c	an be used in our educa	tion system			

Figure 3.15: Section A of the questionnaire

means to use the system, a sizable portion of the students still have not attempted the exercise. One needs to know the reasons behind the resistance to this learning mode.

A t-test (similar treatment is used in [55]) is used to investigate the characteristics of the students who use the system. The two characteristics examined are the frequency of study while traveling and the frequency of WAP usage. Students who do not have the means to use the system are first filtered off, thus reducing the number of subjects to 54. The subjects are divided into two groups: 1) students who have used the system and 2) students who have not used the system. One-tail independent sample t-test is performed on the data from Q2 with the expectation that the mean for group 1 is higher than group 2. According to the t-test result, the p-value of 0.00084 obtained is lower than the standard alpha value of 0.05. The p-value obtained suggests that the mean for frequency of studying while traveling for group 1 is higher than group 2. A higher mean for frequency of studying indicates that students who study while traveling have a higher tendency to use the system. This result is encouraging, as reflected by a high mean score of 4.08 with regard to Q2 for all students. It shows that studying while traveling is a study mode inculcated in the students. In conclusion, the students do have a high potential to be able to use such a system.

A similar one-tail independent sample t-test is performed on data from Q3. The p-value of 0.0057 obtained implies that the mean for frequency for WAP usage of the students who have used the system is higher than those who have not. In other words, students who use WAP frequently are also likely to use the system.

From the above statistics, the main deterrent factor for using the system appears to be the level of proliferation of WAP usage among the campus students, a usage which can be and will be further enhanced with time. This is also reinforced via the qualitative feedback collected, whereby the reasons cited for not using the WAP service include the cost associated with WAP service and the technical support needed to setup the WAP service for the first time.

3.7.2 Section B: Benefits and Improvement Areas

After answering the questions in Section A, instructions are given for students to be categorized into two groups according to whether or not they have used the system. Section B contains questions to be answered by students who have used the system and Section C contains questions for those who have not. From the survey results pertaining to questions in Section B, evaluation is made on the adequacy and performance of the system. Figure 3.16 shows the questions in Section B of the questionnaire.

The impact of the delivered image quality on the perceived usefulness of the system is analyzed using the data collected from Q5 and Q6. Pearson correlation analysis on Q5 and Q6 revealed that the correlation coefficient is 0.8165. The high positive correlation coefficient suggests that there is a strong relationship between the image quality and the perceived usefulness of the system. This ascertains the image processing effort, discussed previously, and used to enhance the delivered image quality.

Affirmative results on the image quality delivered is evident from the response to Q5, a mean score of 3.5/5 indicates that the students are able to view the resultant graph

Figure 3.16: Section B of the questionnaire

clearly. For Q6 and Q8, the mean is 3.75 and 3.5, respectively. The response collected shows that the system has helped to improve the students understanding on tuning the PID controller. In general, most of the students made use of the system when traveling home on public transport. This discovery further supports the correlation in the study habit of students and the tendency to use such a mobile system.

However, feedback is given for improvement too. Despite using fast Excel conversion techniques, some students expect a faster response speed and a reduction in input parameters to improve the speed. Other qualitative suggestions include removing the log-in requirement, because the students do not want their access to the system to be tracked.

Section C (For those who do not use the PID simulator)

10) What is the main reason that you do not try the mobile excel simulator? (e.g., WAP is too expensive, no color display cellular phone, no time etc.)

11) Although you do not try it, do you think the mobile PID simulator is useful and will help to improve your understanding on PID controller?

Yes _____ No _____,

Figure 3.17: Section C of the questionnaire

3.7.3 Section C: Resistant Factors

Section C of the questionnaires in Figure 3.17 is mainly to focus on students who have not used the system, and the possible reasons for not doing so. Some of the commonly

cited reasons for not using the system is the perception that WAP usage is expensive; low ownership of a color display mobile device; and that time and effort is needed to resolve difficulties when setting up a WAP connection for the first time. However, a generally encouraging feedback is received from this group of students as to the benefits of the system to improve learning. A strong 82% of the students agree on the usefulness of such a mobile learning system.

From the response, deduction shows that most of the students are motivated to use the system, although some of them are not yet accustomed to using WAP services and believe that WAP usage is expensive. Since the telecommunication company has adopted GPRS infrastructure several years ago, the cost of WAP usage has been maintained at 0.53cents (Singapore dollar) per kB. With the inauguration of 3G in Singapore from 2005, the WAP usage cost is expected to drop further.

3.7.4 Feedback Summary

Table 3.2 below shows the summary of the survey results for all the closed format questions posed in the questionnaire. In general, the feedback to this pilot project to inculcate more M-learning experience for students has been encouraging and positive.

3.8 Conclusions

A mobile PID simulator, which is able to run on most current color display mobile devices, has been developed, implemented, and evaluated. The system uses innovative methods to improve the image qualities and increase the processing speed to enable it to

Table 3.2: Summary of survey results

Question	Mean out of 5	Yes	No
Do you currently own a color display handheld mobile device (e.g., PDAs and cellular phone)?	N.a.	75%	25%
How often do you study while taking public transport?	4.08	N.a.	N.a.
Do you use the WAP function on your mobile phone?	N.a.	80%	20%
If you have the WAP function on a mobile phone how often is used?	2.57	N.a.	N.a.
How is the quality of the image displayed?	3.5	N.a.	N.a.
Does the use of the mobile excel simulator improve your understanding of tuning a PID controller?	3.75	N.a.	N.a.
Does using the system help you to understand the difficulties faced by process operators when they have to tune thousands of loops everyday?	3.5	N.a.	N.a.
Although you do not try it, do you think the mobile PID simulator is useful and will help to improve your understanding on a PID controller?	N.a.	82%	18%

be used for M-learning applications. With the deployment of the system, two important results have been achieved or observed based on the evaluation by the end users.

First, through the use of the system the students have gained a deeper understanding in the function and tuning aspects of a PID controller without the need to schedule further classes and laboratory exercises for this purpose. This capability is a key and useful feature of the system, especially for large classes of students and part-time students who are time-tight to balance work and study. After the exercise the students can better relate the effect of the variation of each of the PID parameter on the overall control performance, and they are in a better position to take on real-time control work and laboratory sessions.

Secondly, through the feedback on questionnaires gathered, valuable information on how to further improve the current system has been obtained. WAP surfing culture is identified as one area to improve in order for such M-learning system to be used even more widely.

Through this pilot launch of mobile simulator, one observes and concludes that the students are generally receptive to mobile education, which encourages the authors to improve on the current system and develop other mobile educational applications suitable for use in engineering education.

Chapter 4

Bluetooth Assistive Technology

4.1 Introduction

This chapter focused on an AT application, leveraging on wireless voice and data communication, to assist the visually-impaired, physically-disabled or elderly public bus commuters to be able to access these transportation services independently and effectively. This class of commuters faces difficulties to resolve the correct oncoming bus and prepare themselves for boarding the bus. Thus, they often have to rely on other commuters to help in this aspect of their lives and this greatly lowers the motivation to use the public transport service freely.

A Bus Notification System, using Bluetooth wireless technology, can be developed to enable automatic and audio notification of the requested bus number that the commuter is planning to board via his/her mobile phone or other mobile device. A similar application has been proposed by a group from the University of Colorado [56].

This chapter elaborates on the development of a "Bluetooth Notification System for the Disabled Bus Commuter" as a specific example of the use of wireless technologies in AT application. The implementation and the development challenges will be duly discussed.

4.1.1 Bluetooth Protocol Stack

The Bluetooth protocol stack comprises a series of layers. It is similar to the Open System Interconnect model. The stack is important to allow one to access and control a Bluetooth device programmatically. The layers of the stack can be realised in one of the following forms: hardware, software, firmware or a combination of them. The stack is used to implement the Bluetooth specifications.

With the stack in place, program applications can then pinpoint the types of Bluetooth devices in the vicinity of another, discover the services they offer and use the services appropriately. The layers that are relevant to this AT application which will be introduced later are as follows:

Host Controller Interface - Handles communications between Host and Bluetooth.

L2CAP(Logical Link Control and Adaptation) - Handles data from higher layers.

SDP(Service Discovery Protocol) - Discovers services in other Bluetooth devices.

RFCOMM - Provides a RS232-like serial interface.

OBEX(Object Exchange) - Sends and receives objects like files.

4.1.2 Bluetooth Profile

A Bluetooth profile refers to a designed set of functionalities for Bluetooth devices.

This definition is to ensure interoperability by providing a structured set of procedure

calls instructions for the higher layers of the Bluetooth stack to communicate with the lower layers of the stack. As a result, Bluetooth profiles provide a way for Bluetooth technology to be incorporated into different devices and applications and yet, still work in a unified standard way. The following profiles are used for the AT application:

Generic Access Profile - it is the most basic and essential profile, upon which all other profiles are built. It is used for a basic connection establishment.

Service Discovery Application Profile - it provides the means for client applications to discover the existence of services provided by server applications as well as the attributes of those services.

Serial Port Profile - it interacts with the RFCOMM layer in the Bluetooth stack and creates a virtual serial port (RS-232 serial cable emulation) for data transmission.

4.1.3 Piconet

Bluetooth employs a Master-Slave architecture, in which a device that initiates the connection is the master and the device that receives the connection is the slave. Piconet is a star topology net that consists of 1 Master and 1-7 slaves. The Bluetooth devices in the Piconet share the same timing sequence adopt from the master device. Thus, there can only be one master in a Piconet. The slave devices in the Piconet cannot transmit data directly to other slave devices in the same Piconet. Communication between the slaves devices have to be initiated via the Master. Piconet can be linked to form Scatternet, a larger scale network. A slave can participate in more than 1 Piconet.

4.2 Design Issues

The target application of "Bluetooth Notification System for the Disabled Bus Commuter" relates to the category of communication aid for the visually-impaired, physically-disabled or elderly who are generally facing difficulties when using the current bus transport system. It can also be extended to those who are facing difficulty in using other public transport, such as the taxis.

Under the current transport system, such a commuter will experience difficulty in flagging down an approaching bus on time. At the same time, a visually-impaired commuter or an elderly with poor vision is unable to know if the right bus is approaching him. In addition, if a visually-disabled commuter has boarded the bus, he is unable to know when the bus has reached the desired destination. He has to rely constantly on fellow commuters around him for guidance. Thus, his ability to travel independently is greatly hampered.

System Architecture

A solution based on Bluetooth technology is proposed to alleviate the mentioned problem.

The proposed system as shown in Figure 4.1 consists of three Bluetooth components:

- 1. A Bluetooth embedded system installed on the bus
- 2. A Bluetooth server module located at the bus stop
- 3. A Bluetooth-enabled phone carried by the commuter pre-installed Java application

software.

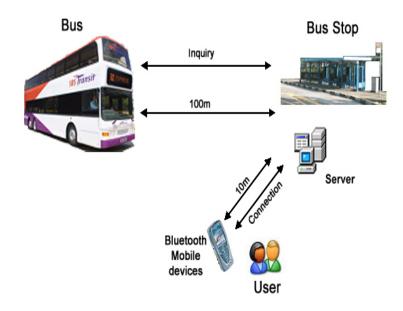


Figure 4.1: Bluetooth AT system architecture

When the user approaches the bus stop, the mobility agent (a pre-installed mobile software) automatically established a connection between the user's mobile and the bus stop server via Bluetooth connection. The automatic connection setup is necessary as elder and the visually impaired is not expected to know how to initiate a bluetooth connection manually. Each bus stop is associated with a unique identification. The bus-stop server will notify the user with the ID of the bus stop. Other information (bus route, frequency etc.) can be accessed from this server as well.

Via his mobile phone, the user will key in the bus number that he is waiting for as shown in Figure 4.2. Upon receiving this message, the bus-stop server will, via Bluetooth, detect the bus once it is within the Bluetooth messaging range and transmit a notification message to the user's mobile phone to alert the user. An audio message is then played to inform the user of his approaching bus as illustrated in Figure 4.3. The bus-stop server

can only maintain up to seven active concurrent connections. Therefore, if there are more than seven users at the bus stop, the bus-stop server will put some of the initial connections to park mode and will reactivate these connections on-demand. Through the Bluetooth system on the bus, the bus driver can also be notified of commuters waiting for it at the bus stop so that he will not miss them.

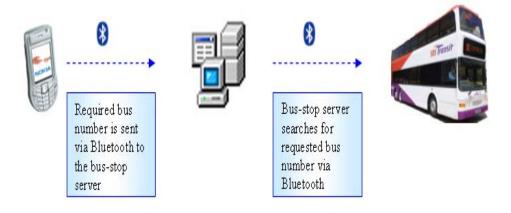


Figure 4.2: Bus commuter enters bus number into his mobile phone

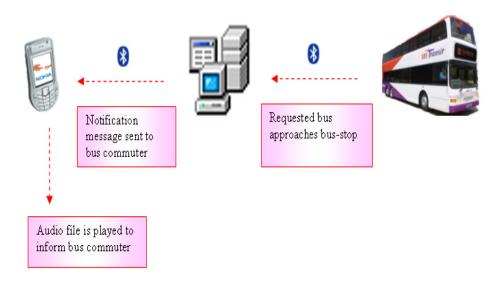


Figure 4.3: Bus commuter receives notification message when the bus is approaching

On the other hand, when the commuter (who has boarded the bus already) wants to

alight from the bus at a specific bus-stop, he can enter the destination bus-stop ID. Once the bus has reached the designated bus-stop, the bus driver is alerted to it and is able to let the commuter alight.

4.3 Implementation Overview

The realisation of the Bluetooth notification system prototype is illustrated as in Figure 4.4. There are three main components:

- 1. Bus-stop Bluetooth server attached to a USB, running the Java Bluetooth stack.
- Bluetooth mobile phone client capable of running Java (J2ME), Java Bluetooth API (JSR-82) and Multimedia API (JSR-135).
- 3. Bluetooth-bus as simulated by the Ericsson Bluetooth Radio Kit.

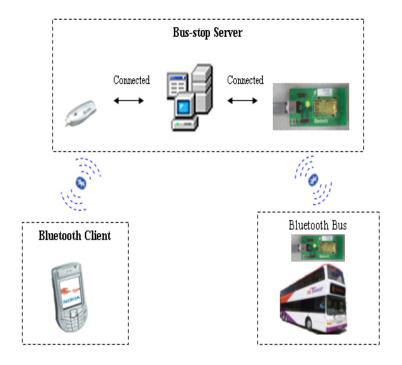


Figure 4.4: Implementation of AT system prototype

As shown in Figure 4.4, the Bluetooth server is started and ready to accept a client connection. The bus commuter uses his mobile phone to initiate a Bluetooth connection with the bus-stop server as seen by the personal computer with the USB Bluetooth adaptor. Once a Bluetooth connection has been established, the bus commuter will key in the required bus number into his mobile phone using the preinstalled Java (J2ME) software. The number is then sent to the server.

Using the requested bus number, the server will initiate the Ericsson Bluetooth Radio kit to search for the requested bus number, based on its corresponding Bluetooth address.

Once the requested bus is within the range of the Ericsson kit, a match between the Bluetooth address and the bus number will be made, and if there is a match, a notification message will be sent to the Bluetooth client. This, in turn, activates the audio notification message of the Bluetooth client, which inform the bus commuter that his bus is approaching.

4.4 Bluetooth Server

As the Serial Port Profile is used, the server needs to define a service record to track the Serial Port Profile service. Thus, the particular service of the server is advertised to any client who wants to use the particular service.

After registering the service record in the Service Discovery Database, the server application initiates a blocking call on its incoming Bluetooth connection port. It waits for a client application to initiate contact with the server to access the service. In this

case, the mobile phone (client) of the bus commuter will try to initiate contact with the server to access the advertised service to send the requested bus number to the server.

The client application and the server application then establish a Bluetooth connection.

Once the data transmission completes (server receive requested bus number), the client disconnects from the server. The server then goes back to the state where it waits for the next client connection. The flowchart, depicted in Figure 4.5, illustrates the server application.

Note that the client will disconnect from the server once the requested bus number has been transmitted. This will free the server for other user connection. A more elegant method will be to perform master/slave switch and maintain the connection. This situation is similar to the LAN access point application; whereby the connecting devices join a Piconet by forming a small Piconet containing themselves as master and server as slave, after which a master/slave switch to allow the server to control the Piconet. This is to save resources as it is a waste of resources for the server to constantly poll devices for connection [57]. However, not all the mobile devices support the master/slave mechanism. Thus, a disconnection is implemented to allow other user connection.

Once the Bluetooth connection is established between the server and the client, the requested bus number is sent from the client to the server via stream connection using the RFComm layer of the Serial Port Profile. The requested bus number, in this case "96", is sent to the server. An acknowledgement message from the server is sent to the client. The client replies with a "Waiting" message and waits for the notification

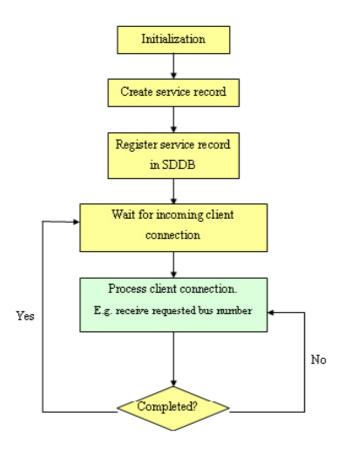


Figure 4.5: Flowchart for the initialization of server

message that the requested bus number is approaching the bus-stop.

In the meantime, the Ericsson Bluetooth radio kit application starts to poll for any available Bluetooth addresses in the vicinity of the server. The server contains a list of numbers of buses which will stop at the bus-stop and their corresponding Bluetooth addresses. For every Bluetooth address detected by the Ericsson radio software, the server will check if it corresponds to the requested bus "96" in the server database. Once a match has been found, a notification message is sent to the client and the client plays an audio notification to alert the bus commuter that his bus is approaching. The flowchart, shown in Figure 4.6, illustrates the concept.

Two Ericsson Bluetooth modules are used to detect the approaching bus number. A single-point and multi-point Bluetooth module is to be mounted on the bus and the server respectively. The module on the bus will periodically enter the inquiry scan state to listen for ID packet transmitted from the inquirer (Bus stop server). During the inquiry procedure, the bus's module will respond with a Frequency Hop Synchronisation packet to the inquirer. Based on the FHS packet received, the server will be able to determine the BD_ADDR (Bluetooth address) of the approaching bus module. The unique BD_ADDR is then translated to the bus number. In this method, no connection needs to be established to transmit the bus number information, thus reducing the overall response time of the system.

The proprietary Ericsson PC Reference Stack allows the control of the Bluetooth module at a different stack layer. In the initialization process, the application configures

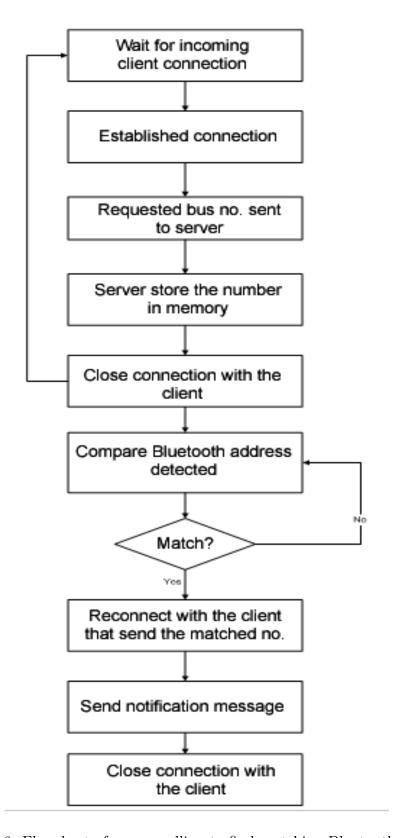


Figure 4.6: Flowchart of server polling to find matching Bluetooth address

the HCI layer and registers with the Stack Connection Manager component. Figure 4.7 show the address of the bus's Bluetooth module detected by the server.

The HCI_ReqPeriodicInquiry function is used to set the module into periodic inquiry stage. The HCI_ReqPeriodicInquiry will return the address of the bus Bluetooth module detected in the vicinity.



Figure 4.7: Address of the bus module detected by the server

4.5 Bluetooth Client (Mobile device)

For the client application, there are five main components:

- 1. Initialization of the client device
- 2. Device Discovery

3. Service Discovery

4. Connection and Notification

5. User Interface

The client (mobile devices) needs to be initialized first. This means that the service address (Serial Port Profile service), provided by the Bus-stop server, is incorporated in the client application. This is to enable the client to search for the correct Bluetooth server and the correct service to initiate connection.

Once the client device is initialized, it starts to carry out device inquiry for the available Bluetooth devices in the vicinity. For every Bluetooth device found, the corresponding name and Bluetooth address of each device is stored in a hash table in the client application. Once the process of device discovery has completed, the client application moves to the next stage - Service Discovery. The flowchart of this process and the screenshot of the client device are shown in Figure 4.8.

After the Device Discovery process has completed, the Service Discovery process kicks in. All the available Bluetooth devices data are now stored in the hash table. The client application will search for each device data: the Bluetooth address, the type of services and corresponding service address offered by each device. If the Bluetooth and service addresses correspond to that of bus-stop server, the client device will attempt to initiate a Bluetooth connection with the remote device (which is the bus-stop server in this case) as shown in Figure 4.9.

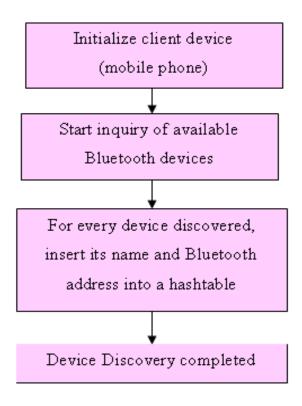


Figure 4.8: Flowchart of Device Discovery process of client device

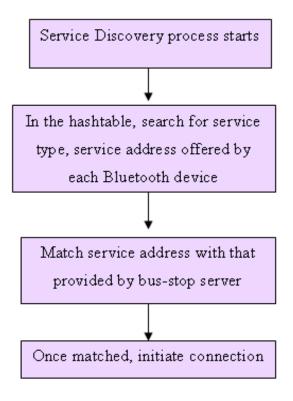


Figure 4.9: Flowchart of Service Discovery process of client device

Once a match between the service address of the remote device and that given by the bus-stop server is found, the client device will attempt to initiate connection to the bus-stop server. As soon as the connection is successful, the requested bus number keyed in by the bus commuter will be sent via Bluetooth stream connections to the server. The server will acknowledge by sending a "RECEIVED" message back to the client to inform it that it has received the requested bus number. The client then waits for the server to check if the requested bus is approaching the bus-stop.

When the requested bus is approaching the bus-stop, the server will send an "AR-RIVED" notification message to the client via Bluetooth to indicate that the requested bus is arriving. The client will read in the message and will play an audio file to alert the bus commuter that his bus is approaching.

User Interface

An assistive feature of the Bluetooth bus notification system is that it incorporates the element of audio for the client application (in the mobile phone). This is especially important for the visually-disabled commuter. Therefore, to aid the visually-disabled commuter to use the notification system proficiently, a series of audio (mp3) files are stored in the client application. They are played at various stages of the client application.

As seen in the figure 4.10, the first audio file will sound to prompt the bus commuter to enter the bus number that he wants to be notified of when it approaches the bus-stop.

Once that is done, the second audio file will instruct the commuter to wait for the client

application to connect. The final audio file will play once the requested bus number is approaching the bus-stop.

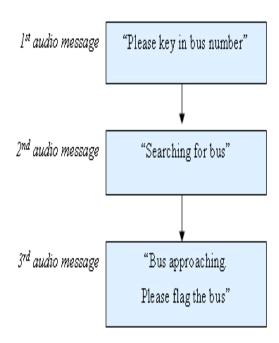


Figure 4.10: Flowchart demonstrating the sequence of audio files for the client application

4.6 Cost Issues

According to statistics given by the Singapore Association of the Visually Handicapped [58], there are 2439 visually-impaired personnel registered under it as of March 2003. The proposed application "Bluetooth Notification System for the Disabled Bus Commuter" solution may not be economically viable if it will only cater to for the disabled. To make the solution more economically feasible, the overall system framework can be enhanced to be beneficial to the general commuters and bus companies as a whole.

The proposed system requires all the buses and bus stop to be equipped with a Bluetooth module, this indirectly increases the normal operating cost of the public bus transport. If the aim of the Bluetooth module is solely meant to facilitate the minority visually-impaired commuter, the bus operator will find it hard to justify the cost. Therefore, there is a need to expand the use of Bluetooth module for other usage. A possible use of Bluetooth module will be to provide wireless diagnostics and conditions-based maintenance support for the bus fleet [59]. When the bus returns to the terminal after daily service, the maintenance data are wirelessly transmitted via Bluetooth to the server in the depot.

The newer generations of mobile phones are now all equipped with Bluetooth technology, so that no additional cost is incurred as well on the user's side in the near future.

In other words, the functions highlighted for the disabled are derivable from the overall framework at no significant incremental cost.

4.7 Conclusions

In this chapter, the use of wireless technologies to fulfill an AT application has been examined. Among the short range wireless technologies, Bluetooth has the highest market penetration and support. Through the specific implementation of the "Bluetooth Notification System for the Disabled Bus Commuter" project, the principles and mechanisms of Bluetooth is instilled. "Bluetooth Notification System for the Disabled Bus Commuter" which aim to assist the visually-impaired, physically-disabled and elderly

commuters to take the public bus service independently provides a good example of how wireless technologies can be used in AT applications.

Chapter 5

Development of Bluewave: A Wireless Protocol for Industrial Automation

5.1 Introduction

The routing algorithms for the wireless ad hoc networks can be broadly classified into two main categories [60]: table-driven (proactive) and on-demand (reactive). In recent years, different approaches towards routing for ad hoc network have been proposed: location-based and hybrids.

In table-driven routing protocols, every node constructs one or more tables that contain global routing information. The routing information is pre-computed in the background and is updated periodically. Examples of protocols belonging to this category are DSDV (Destination-Sequence Distance-Vector) [61] and WRP (Wireless Routing Protocol) [62]. Due to the overheads in processing and the large routing table that it needs to store even for a static ad hoc network, table-driven protocols are not scalable. Bandwidth is also wasted in the form of indiscriminate periodic flooding of update tables to the

nodes in the unused route. Hierarchical routing [63] aggregate nodes into subnet and limit frequent information exchange within the cluster. As compared to flat routing in which all the nodes maintain knowledge of all other nodes in the network, hierarchical routing techniques help to reduce the size of the table storage significantly. However, hierarchical routing introduces complexity in the design of the wireless routing protocols.

On-demand routing protocol reduces redundancy in routing information storage by creating routes only when communication is required. When communication between two nodes is needed, a query message will be sent by the source node to the surrounding neighboring nodes. The query message will propagate from node to node until it reaches the destination node or it is dropped by the network. During propagation, the intermediate node will record itself in the forwarding table. In this manner, the route path to source node is established. When the destination node receives the query message, a reply message is sent via the route path embedded in the query message. The route established will be maintained till the communication ends. On-demand based routing protocols include AODV (Ad-hoc on-demand Distance Vector) [64] and DSR (Dynamic Source Routing) [65]. The high route setup latency makes on-demand routing protocol unsuitable for most real-time monitoring applications, in which there is a hard deadline time constraint (typically in seconds or milliseconds) for the data to reach the server [66].

Recent proposals of a hybrid routing protocol aim to combine advantages of tabledriven (low route discovery latency) and on-demand (low processing overhead) approaches have been made. ZRP (Zone routing protocol) [67] limit the table-driven activities to the node's predefined neighborhood. A search outside the neighborhood is performed using on-demand method. Boundary of the neighborhood may be hard to define and is application dependent. This increases the difficulty in the implementation of ZRP.

The advantages of using the geographical location of the nodes in a routing decision are explored in location-based routing. Location-based routing is able to mitigate some of the problems faced by the conventional routing protocol. In location-based routing protocol, routing information need not be stored and transmission of route setup messages can be reduced, as route path can be derived from the node geographical location. Reduction in cost of a locator module (e.g., Ubisensor [68], GPS (Global Positioning System)) has enhanced the feasibility of location-based protocols like LAR (Location-Aided Routing)[69]. The additional cost and size of the locator module, however, is a major drawback.

5.1.1 Bluetooth for M2M

There are many protocols and standards for the wireless communication network. Currently, the more prominent wireless technologies are the 802.11, Bluetooth and Zigbee for industrial automation. In terms of support and availability, the 802.11 and Bluetooth are much more extensive than Zigbee. This section gives a comparison between 802.11, Bluetooth and Zigbee.

Power consumption for Bluetooth and Zigbee is far less than 802.11. The power sav-

ing for using Bluetooth or Zigbee compared to 802.11 may range from 10 to 100 times depending on the application. Power source is a critical factor for industrial wireless application. Usually the power source is from small chemical batteries rather than DC power line, so as not to limit the deployment freedom of the wireless application. Different low-power operation modes of Bluetooth enables the batteries life to be prolonged.

Bluetooth and Zigbee provide better connection interface than 802.11. For Bluetooth and Zigbee, the microcontroller can connect to the Bluetooth module via the serial or UART (Universal asynchronous receiver-transmitter) interface. For 802.11, complex parallel bus connection will be needed [70].

Bluetooth uses FHSS (frequency-hopping spread spectrum) while 802.11 and Zigbee uses DSSS (Direct spread spectrum). DSSS is more susceptical to multipath fading than FHSS, therefore Bluetooth is more robust than 802.11 and Zigbee in an indoor setting. The maximum throughput for Bluetooth and Zigbee is in the order of kbits/s, far lower than 802.11 which has a throughput in the order of Mbits/s. However, this may not matter much as most of the transmitted data are diagnostic or monitoring information which will not take up a lot of bandwidth.

The average transmission radius for class 3 type Bluetooth and Zigbee is 10m, much shorter than the 100m transmission radius of 802.11. With a shorter transmission radius, the Bluetooth nodes are able to form a multi-hop link to the sink instead of the point-to-point link usually formed by 802.11 applications. Multi-hop links have the advantages of self-healing; scalability and better link reliability than point-to-point link. Reduced

transmission radius also provides a form of spatial partition that reduces interference [71].

From the comparison, it is noted that Bluetooth and Zigbee have an edge over 802.11 in industrial wireless communication. The critical differences between Bluetooth and Zigbee lie in the throughput and power consumption. Bluetooth has a data throughput of 725kbits/s, while Zigbee has a data throughput of 125kbits/s. On the other hand, Zigbee power comsumption is lower than Bluetooth. Whether to use Bluetooth or Zigbee depends on the need of the specific applications.

5.1.2 Bluewave

In the near future, wireless PLCs will be a common phenomenon in most modern factories. There are two main methods to collect data from the wireless components. Depending on the application, data can be gathered either by a technician who moves around the machines with a wireless receiver on a regular basis or by linking all the machines wirelessly to a central server. The former data collection method is easier to implement while the latter will have greater cost savings in the long run. It is believed that data collection via a central server will be the dominant method. The scenarios whereby many PLCs link wirelessly to a central server can be viewed as a multiple-source single-sink static ad hoc network.

As discussed in the previous section, Bluetooth is an excellent technology to establish wireless communication in a factory setting. However, Bluetooth faces some constraints and limitations due to its specifications. In Bluetooth, a master node can only connect

seven active slave nodes at a time, and there is no broadcast mechanism. This severely affects the message flooding mechanism which is common in many wireless protocols.

The mechanisms that work well for wireless mobile ad hoc network to cater to mobility prove to be redundant, and instead add to processing overheads in a static wireless factory setting. Bluewave is an on-demand based routing protocol that utilizes the machinery static positioning in a rectangular factory floor setting. The prior zone discovery phase of Bluewave determines the relative geographical position of the machinery. This aids the route setup without the use of expensive locator module (e.g., Ubisensor). The alternate active and dormant states of Bluewave facilitate information dissemination through the network by reducing the interference, and increasing the chance of a successful connection. The following sections describes the algorithm of Bluewave in detail.

5.2 Bluewave Algorithm

As mentioned, the proposed algorithm is devised to be used in a single-sink and multiple-source scenario. This scenario, as shown in Figure 5.1, is typical in a PLC-controlled setting in a factory.

As an example, it is assumed that the machinery is distributed uniformly across a square factory floor. The proposed routing algorithm consists of four phases, namely:

1) Zone discovery, 2) Route setup, 3) Route reconfiguration and 4) Zone maintenance.

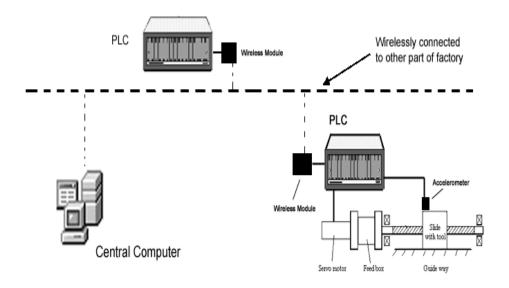


Figure 5.1: Multiple PLCs connected to central computer

5.2.1 Zone Discovery

In the discovery phase, the sink node will send a join message to the potential source nodes around it. Using the PLC example, the sink node will be the central computer, while the source nodes will be the PLCs. Those PLCs that receive the message from the central computer will form the first layer zone and record in its routing table that it belongs to central computer X Zone 1. After which, the first layer PLCs will in turn send the join message to the other PLCs around it. Those PLCs that have already joined the network and have an equal or lower than the sender zone number + 1 will not change. Those PLCs that have not joined the network or have a higher zone number will update its routing table as sender zone + 1. Figure 5.2 depicts this zone discovery formation in a simplified form. In this manner, Bluewave acquires a routing table storing the zone information (C X: 4) which translate to central computer X Zone 4. However,

this routing information will not grow with the number of PLCs node in the system, thus the space complexity is O(1). Figure 5.3 shows the actual zone formed. The three circles illustrated in Figure 5.3 represent the transmission radius of node A, B and C. Zone 1 will be the area covered by circle A. While Zone 2 corresponds to $(B \cup C)/(B \cup A) \cup (C \cup A)$.

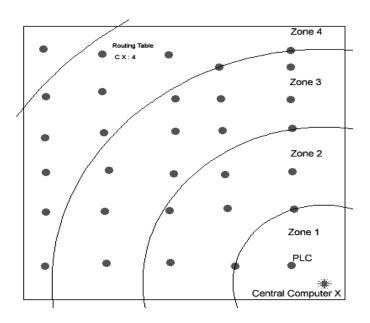


Figure 5.2: Simplified zone formations

The above mentioned mechanism can be considered as a form of broadcast in Multi-hop radio network, whereby the goal is to transmit a message from one node in the network to the all other nodes, under the assumption that only nodes that already have received the message can transmit. However, unlike the broadcast mechanism, Bluewave is subject to more stringent constraints. Transmission of message is considered successful only if the message received by a node in geographic Zone x is transmitted by the node in Zone x-1. This reduces the probability of getting a successful transmission at each inquiry-inquiry scan pair up to (number of surrounding node in x-1 zone) / (the total

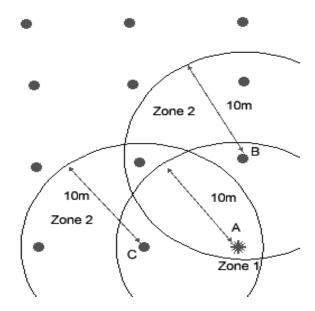


Figure 5.3: Detail zone formations

number of surrounding node).

To shorten the number of steps to inform all the surrounding nodes of their relative positions to the sink node, a modification to the above algorithm is needed. The following mechanism has been devised to shorten the discovery phase.

The entire network is synchronized using a global timer. The global timer is tagged to the zone id message sent out by the sink node. It will continue to decrease according to the native clock in the intermediate node while waiting to be retransmitted to other nodes. The value of the global timer will trigger off one of the following; inquiry scan, inquiry or dormant state in the node. The inquiry procedure in Bluetooth requires neighboring nodes to be in the opposite inquiry states before setting up a connection.

Initially, all the nodes except the sink node will be in the inquiry scan state. At time t = 0, the sink node is in an inquiry state calling out to its neighbor for connection by transmitting a packet containing an inquiry access code. After a time length of T_{window} ,

the sink node will change from the inquiry state to the dormant state. The nodes in Zone 1 which have just received the join message will change from the inquiry scan to inquiry state after the first T_{window} has lapsed. At time 3 T_{window} , nodes in Zone 2 will have received the message and are now in an inquiry state to spread the message to potential nodes in Zone 3, and nodes in Zone 2 will be in dormant state. The cycle will repeat with nodes in Zone x in inquiry state trying to established connection to nodes in Zone x + 1, with nodes in Zone x - 1 in a dormant state. Since the nodes in the same Zone x are in inquiry state and nodes in the Zone x - 1 are in dormant states at the same T_{window} period, any inquiry-inquiry scan pair up will be with node in Zone x + 1. This significantly raises the probability of getting a successful transmission at each inquiry-inquiry scan pair up. The T_{window} period is the time needed for nodes in Zone x to pass the message to all its neighboring nodes in Zone x + 1. The length of T_{window} can be calculated mathematically and verified by experiment, to be shown in the later section. Table 5.1 below illustrates the Inquiry \rightarrow Dormant \rightarrow Inquiry_scan stage undergoes by the zone nodes at different T_{window} period in a six zone level network.

Inquiry - I

Inquiry_scan - S

Dormant - D

 T_{window} - T_{wind}

Figure 5.4 depicted the finite state machine of a node in zone x in the discovery phase. When a node is in Zone x (where $x \neq 0$), the node begins in the "Inquiry scan" state. Table 5.1: States Table

Time	Zone 0	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
1st Twind	I	S	S	S	S	S	S
2nd Twind	D	I	S	S	S	S	S
3rd Twind	D	D	I	S	S	S	S
4th Twind	I	S	D	Ι	S	S	S
5th Twind	D	I	S	D	I	S	S
6th Twind	S	D	I	S	D	I	S

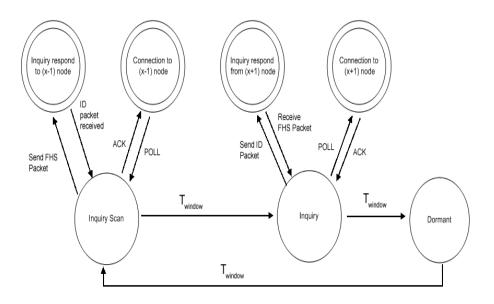


Figure 5.4: Finite state machine for Zone x node

The node will change state between the "Inquiry respond to x-1" and "Connection to x-1" states, depending on the transition conditions (e.g., send FHS (Frequency Hopping Synchronization)), to determine its zone id. After the node has obtained its id, it will enter the "Inquiry" state at T_{window} . In the "Inquiry" state, the node will broadcast its zone id to other surrounding nodes via the "Inquiry respond to x+1" and "Connection to x+1" states. After the second T_{window} , the node will be in "Dormant" state, after which the "Inquiry scan", "Inquiry" and "Dormant" states will repeat throughout the rest of the discovery phase.

5.2.2 Route Setup

An example depicted in Figure 5.5 is used to demonstrate the route setup procedure. Assuming that PLC A has data to send to the central computer. It will try to establish a connection with any PLC that has a lower zone number than itself. The connected node B will in turn search for another PLC in the lower zone number. This process will continue until the central computer X is connected. The concept here is similar to the location-based algorithm whereby the node will forward packet to another node that is closer to the destination. The Bluewave algorithm developed in this paper, unlike the Location-Based algorithm, does not require an external location service to indicate the node geographical position. It will also have a more accurate next hop transmission as the zone is formed based on the transmission signal rather than the geographical position. The set up is also robust as any node in the lower zone is able to act as the intermediate node. It also satisfies the loop free requirement for the mobile ad hoc

network. Like the on-demand algorithm, the path will only be setup when needed, thus no bandwidth is wasted for an unused path. The route setup time is expected to be fast as prior information of zone position is stored in the routing table.

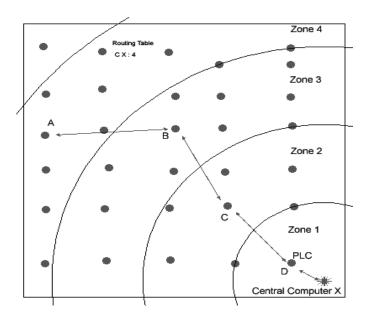


Figure 5.5: Route setup

5.2.3 Route Reconfiguration

Route reconfiguration is needed to recover from the effects of undesirable events, such as host or link failures of various kinds. Although, the factory setting is a static ad hoc network, the probability of a node failure due to machinery breakdown or temporary link failure to adverse transmission environment is still existent. Therefore, a route reconfiguration process is proposed below. Considering the route set up from nodes A to X. Initially the path is set from $A \Rightarrow B \Rightarrow C \Rightarrow D \Rightarrow X$. When there is a link failure between $B \Rightarrow C$, the node B will fail to receive acknowledge message from node C. After several attempts B will initiate contact with another node from Zone 2 and a new

continuation path will form from $E \Rightarrow F \Rightarrow X$. The following pseudocode describe the reconfiguration process:

Reconfiguration(ACK is not received){

```
1
           While(True){
2
               Inquiry();
3
               If (response from lower zone){
                  ObtainFHS(); //Frequency Hopping Synchronization
4
5
                  break;
               }else{ //continue searching for lower zone node }
6
           }
7
8
         Establish_connection();
9
         Continue_sending();
10 }
```

Due to the fact that the form of traffic being handled is file transfer traffic, it is only logical to assume that it is none bursty and always "on". Thus after a timeout period where node C does not receive any packet from node B, node C will send a message down to dissolve the path $C \Rightarrow D \Rightarrow X$. It is important that the shortest possible timeout period is used. This will help to save bandwidth and reduce undue interference.

An example of route reconfiguration is illustrated in Figure 5.6.

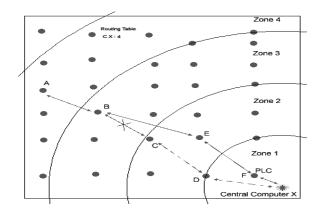


Figure 5.6: Route Reconfiguration

5.2.4 Zone Maintenance

New PLCs will be added to the factory. PLCs will also be switched off for maintenance or repair. In view of this, a method has to be devised to enable new nodes to join the network and for updating the central computer view of the network. When a new PLC is added to the network, it will scan for the surrounding nodes and update itself as $(\min(\text{found node's zone number}) + 1)$. The new node will also send a message to the central computer to inform the central computer of its presence. The central computer will also conduct a periodic update procedure similar to zone discovery to update itself the PLCs that has shut down.

5.2.5 Comparison Studies

This section provides a comparison of Bluewave with routing protocols from various categories; WRP (Table-driven), AODV (On-demand) and ZHLS (location-based). The discussion here is based on Table 5.2.

The main advantage of Bluewave over WRP is in the memory requirement. In WRP

each node requires four routing tables which will severely reduce the scalability of WRP. In network with a large number of nodes, the WRP memory requirement will be substantial. A Bluewave node uses route cache to store the next-hop node address; therefore its memory requirement is independent of the size of the network. WRP periodically transmit "hello" message to the surrounding nodes to update the connectivity. Bluewave does not have any beaconing measure, thus Bluewave takes up less communication bandwidth than WRP during the maintenance phase. Both Bluewave and WRP have the same communication complexity during link failure and addition, as only the neighboring nodes are affected. WRP, through verification of predecessor node information, ensures the route formed is loop free, while Bluewave intrinsic zone segregation allows loop free routing.

AODV expands on DSDV algorithm, it improves on the DSDV by reducing the number of required broadcast. As the name suggests, AODV creates route on demand. When setting up the route, the source node will broadcast the RREQ (Route request) packet to the surrounding nodes, which will forward to it their neighbors until the destination node is found. A reverse path is then set up from the destination node to the source. Both AODV and Bluewave do not require global advertisement, and do not need global network routing information. Therefore, in term of memory and bandwidth requirement, ADOV and Bluewave performed better than many table-driven routing protocols. The route setup time for AODV is longer than Bluewave. ADOV discovers the route on demand by flooding RREQ packet to its neighboring nodes and a reverse path messaging

performed to establish the route. In Bluewave, the route discovering message flows down the hierarchical zone to the destination node. When an intermediate node on the route fails, in the worst case scenario with no other valid route in the cache, ADOV will reinitiate the route discovery process. Bluewave, on the other hand, reconfigures the route path from the point of broken link; therefore it has faster recovery than ADOV.

ZHLS is a (GPS)-based routing protocol that extend on the hybrid ZRP concept.

Bluewave is similar to ZHLS in the use of non-overlapping zone to create a hierarchical structure for routing. However, unlike ZHLS, Bluewave does not require a locator module to determine the zone segregation. This is a major advantage of Bluewave over ZHLS, as the indoor locator module is difficult to implement. Multiple route possibility is possible for ZHLS and Bluewave, since transmission across the zone can be performed by any gateway nodes. The multiple route property avoids traffic bottlenecks and single node failure. In term of bandwidth consumption, ZHLS performs periodic beaconing within the intra-zone node.

Table 5.2 gives a comparison of Bluewave with other popular existing wireless protocols.

- h Height of the hierarchy
- d Diameter of the network
- x Number of nodes affected by the topology change

Table 5.2: Comparison of the characteristic of bluewave with existing protocol

Parameters	WRP(Table-driven)	AODV (On demand)	ZHLS	Bluewave	
Time complexity (route discovery)	O(h)	O(2d)	O(h)	O(h)	
Communication complexity (link failure)	O(x)	O(2d)	O(x)	O(x)	
Loop-free	Yes, after a set- tling down pe- riod	Yes	Yes	Yes	
Number of Table required	Four	One	e Two		
Route maintained in	Routing table	Routing table	Routing Table	Routing cache	
Beaconing	Yes	No	Yes (Partial Beaconing within intra-zone)	No	
Multiple route possibilities	No	No	Yes	Yes	
Auxiliary mod- ule	No	No	Yes (locator module)	No	
Routing metric	Shortest path	Shortest path	Shortest path	Shortest path	

5.3 Evaluation of Bluewave

In this section, a case study is conducted to examine the performance of the proposed algorithm. The subject of the case study is a CNC (computer numerical control) milling machine as shown in Figure 5.7.



Figure 5.7: CNC milling machine

CNC machines are commonly used in manufacturing processes to shape metal workpieces. It is desirable to implement an online condition-based monitoring system that is
able to determine the cutting tool's wear state, Figure 5.8 illustrate the communication
pathway for the system. Timely replacement of the cutting tool will result in substantial
cost saving [72]. Various methods have been proposed to detect the tool wear condition
[73]. Figure 5.9 depicts the wireless monitoring that is used to estimate the tool's wear
state.

In this experiment, the system consists of three jobs which jointly implements the fault-monitoring task. The preliminary on-site test is J_0 , the wireless data transmission

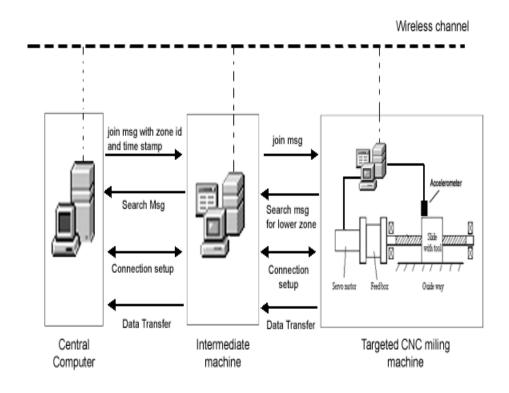


Figure 5.8: Communication pathway

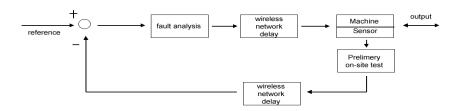


Figure 5.9: Fault-monitoring network

is J_1 , and the computation of nueral analysis is J_2 .

 J_0 is an acyclic event that detects adnomalities in the vibration signature of the machinery. When the vibration of the machinery exceeds a certain threhold, J_0 will set off an alarm that triggers J_1 and J_2 , in which 50kB of vibrations histories data [74] is transmitted to the central computer for processing. Therefore, the release time of J_1 and J_2 is dependent on J_0 . The average cutting feed rate of the CNC machine is 200mm/min and the average workpiece size is 210mm. Therefore, a cycle of to and fro cutting action take approximately 120s. Thus the deadline of J_1 and J_2 is 120s. Since, J_2 , the processing time for the neural analysis is approximately 5s, the desirable response time for J_1 is less than 115s.

As the tardiness of J_1 increases, the usefulness of the transmitted vibration signature data will decrease. However, the usefulness of data does not fall abruptly to zero or a negative value when the tardiness is greater than zero. Delay in the J_1 and J_2 will only result in the deterioration in cutting performance of the CNC machine. Therefore, the system has a soft deadline.

From the experimental data collected from the CNC machine case study, the benefits of Bluewave can be verified. The alternate zone idle and active stage changes through a synchronized timer, reducing the time needed for initialization drastically. The results will show that the use of Bluewave in data transfer meets the reponse time requirement of monitoring system for CNC machines. Through the case study, better performance of Bluewave as compared to protocol using flooding mechanism is demonstrated.

Experiment Setup

It is assumed that the machinery is distributed uniformly across a square factory floor with an area of 900 m². Figure 5.10 illustrates the setting. Where the nodes represent the location of the machines. Let d be the distance between two machines, where d is greater than 4m (safety distance between the machines) and less than 30m, i.e., 4m < d < 30m. Let N denotes the total number of node (i.e. machinery).

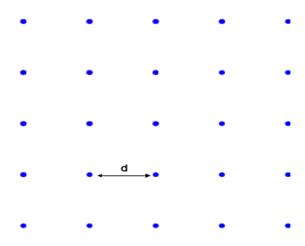


Figure 5.10: Machinery position

Six PCs (Personal Computer) are setup acting as the CNC machines as shown in Figure 5.11 to investigate whether Bluewave performance satisfy the above constraints imposed by the CNC application requirements. The PC was installed with the Linux OS (operating system) and Bluez distribution from [75]. Bluez is a software implementation of the Bluetooth wireless standards specifications for Linux. Using Bluez, a program was developed to control DBT-120 USB Bluetooth Adapter [76] attached to the PC via USB. The program executes the Bluewave algorithm and collects the performance data. The collected data is then extrapolated using mathematical formulation to different scenarios

involving up to 100 machines.

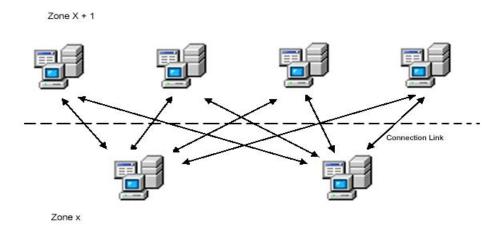


Figure 5.11: PC acting as CNC machine

5.4 Discovery Phase

This section illustrates the improvement of the developed algorithm as compared to conventional flooding technique in term of initialization time.

5.4.1 Connectivity

Table 5.3 shows the average connectivity of each zone to the next immediate higher zone. Such connectivity hereinafter will termed as "higher mean connectivity" of Zone x. The higher mean connectivity indicates how many connections need to be made before all nodes in the higher zone will receive the message.

The higher mean connectivity is calculated as follows:

1. There are n nodes. Node a is position at (x_a, y_a) and node b is position at (x_b, y_b)

Num nodes zone 0 zone 1 zone 2 zone 3 zone 5 zone 6 zone 4 2.0N.A N.A N.A N.A N.A 4 1.0 N.A 16 2.02.0 2.0 1.5 1.33 1.0 N.A 253.03.0 3.0 3.0 N.A N.A 36 5.0 4.23.892.631.86 N.A N.A 497.0 N.A N.A N.A 6.435.931.7664 7.0 5.93 N.A N.A 6.434.711.1781 10.08.7 7.234.470.59N.A N.A

Table 5.3: Average Connectivity of zone x to zone (x+1)

2. Let $a \in Z_i$ where i is the zone id and Z_i is the set of nodes with the same zone id

5.1

10.19

2.0

N.A

N.A

3. $b \in 1,...,$ n and $b \neq x$ and Euclidean distance $D_{ab} = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}$; transmission radius

The higher connectivity $a_b = 1$ if $b_i > a_i$ otherwise 0

11.25

Therefore, the mean higher connectivity of Zone i is:

$$\frac{\sum_{b=1}^{n} a_b}{|z_i|}$$

5.4.2 Inquiry Window Length and Discovery Time

A trial is an attempt by the node to establish a connection and to pass a message to another neighboring node. The time taken for this process is as follow:

$$T_{trial} = T_{inquiry} + T_{connection},$$

100

12.0

where $T_{inquiry}$ is the time taken for discovering surrounding Bluetooth devices using

inquiry, $T_{connection}$ is the time taken for setting up the a connection and send a packet. Thus,

 $T_{inquiry} \gg T_{connection}$

Therefore, the duration of a trial can be shortened by minimizing the length of inquiry phase. However, a shorter inquiry phase may result in a failure to detect the surrounding nodes. Below is the result of the experiment of varying the time of inquiry phase and its effect on the probability of a trial. In Bluez, the inquiry phase is conducted in steps and each step takes approximately 1.28s.

Table 5.4: Effect of a inquiry length on the success of a trial

The inquiry length	Percentage of successful trial			
1	0			
2	0			
3	63			
4	98			
5	100			
6	100			
7	100			
8	100			

From Table 5.4, it can be concluded that an inquiry window length of 4 is the lowest inquiry length that will give the highest probability of successful trial. The time taken for discovery phase $T_{discovery}$ is the sum of N_{Zi} , the trial taken for Zone x nodes to pass the id message to Zone x+1 nodes multiplied to the minimum inquiry window $T_{inquiry}$ length duration of 4 * 1.28s, i.e.,

$$T_{discovery} = \sum_{i=0}^{I} N_{Zi} * T_{inquiry}$$
 where I is the total number of zones

5.4.3 Conventional Flooding Method

Using the conventional flooding method, the number of trials needed for Zone i to pass message to Zone i+1 can be modelled by extending the classical "Balls and bins" problem also known as the coupon's collector problem. With reference to Section 7.1, the average connectivity of Zone x to Zone (x+1) is calculated. The "higher mean connectivity" is also the ratio of the "Balls" and "Bins". For example, with 16 nodes at Zone 4, the "higher mean connectivity" is 1.33, this indicates that the ratio of balls and bins is 3:4. The number of balls is denoted as α .

Let $i \in 1,...,$ n in the set Z_{x+1}

Pr(i is not inquired in the first m trial)

$$= (1 - \frac{1}{n})^{\alpha m} \approx e^{-\frac{\alpha m}{n}}$$

Thus, the expected number of vertices contacted in \mathbb{Z}_{x+1} after m trial is

$$n(1 - e^{-\frac{\alpha m}{n}}),$$

Let E_i^m denote the event that the ith node is not collected after m trials. Then

$$Pr(E_i^m) \approx e^{-\frac{\alpha m}{n}},$$

and

$$Pr(\neg E_i^m) \approx 1 - e^{-\frac{\alpha m}{n}}$$

Thus,

Pr (All the Z_{x+1} are inquired in m trials)

$$= Pr(\neg(\vee E_i^m))$$

$$= Pr(\wedge \neg E_i^m)$$

$$= \prod Pr(\neg E_i^m)$$

$$\approx n(1 - e^{-\frac{\alpha m}{n}})^n$$

$$\approx e^{-ne^{-\frac{\alpha m}{n}}}$$

Let $m = \frac{1}{\alpha}(n \ln n + nc)$

Pr (All the Z_{x+1} are inquired in m trial) $\approx e^{-e^{-c}}$

When c = 3

Pr (All the Z_{x+1} are inquired in m trial) ≈ 0.95

Using $m = \frac{1}{\alpha}(n \ln n + nc)$, the number of trials needed for the message to propagate from zone x to zone x+1 can be determined.

The above equation is verified with the following experiment. In the experiment a node sends a message to other three neighboring nodes, the number of trial was varied and the probability that all the three neighboring nodes receive the message will be examined.

From the Figure 5.12, it can be deduced that $m = \frac{1}{\alpha}(n \ln n + nc)$ gives a good approximation to the number of trials needed to propagate message from Zone x to Zone x+1.

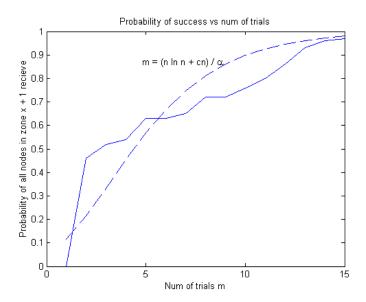


Figure 5.12: Probability of success vs trial

Therefore, it is reasonable to conclude $\mathbf{N}_{zi} \approx \mathbf{m}_i$.

5.4.4 Bluewave Method

In the Bluewave method, when nodes in zone x+1 receive a message, they will go into dormant state until the next stage. Nodes in dormant stage will not respond to inquiries, thus preventing repeated sending of message to nodes in zone x+1 that has already received the id message from other node in zone x. The synchronized alternate dormant and active states change mechanisms in Bluewave greatly reduces the time needed for the discovery phase.

$$N_{zi} = \lceil \frac{\theta}{|Z_i|} \rceil$$
, where θ is the total higher connectivity

Given that $T_{discovery} = \sum_{i=0}^{I} N_{Zi} * 4 * 1.28$, the total time needed for the discovery phase for different number of nodes is calculated and is presented in Table 5.5 and Figure 5.13. The discovery time growth gradient of Bluewave is much smaller than the

Table 5.5: Comparison of initialization time of bluewave with conventional flooding

Number of nodes	Bluewave discovery time (min)	Flooding discovery time (min)		
4	0.2560	0.9390		
16	0.9390	3.4100		
25	1.0240	4.4370		
36	1.6210	8.8740		
49	1.8770	11.5200		
64	2.3000	13.6500		
81	2.8160	18.3460		
100	3.6700	23.98.00		

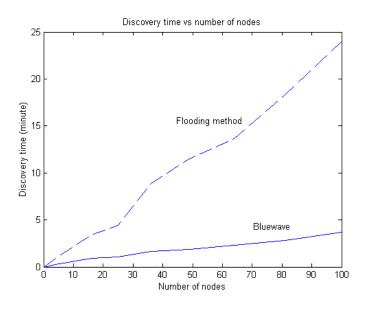


Figure 5.13: Discovery time vs number of nodes

conventional flooding method. This ensures the scalability of the Bluewave. Machines in factory automation are subject to periodical shutdown for maintenance and repair. Fast initialization stage will be a valuable advantage of Bluewave when used in a factory setting.

5.5 Data Transfer Phase

The CNC machine monitoring is a real-time application with a hard deadline for its data transfer. As illustrated in the previous section, the 50kB of data has to be delivered to the central computer for processing within 115s. In this section, the data transfer latency for different number of nodes will be discussed. The data transfer latency comprises of two parts, the data transfer time and the route setup time.

5.5.1 Average Hop

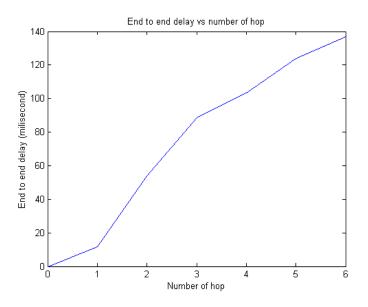


Figure 5.14: Packet end-to-end delay vs number hop

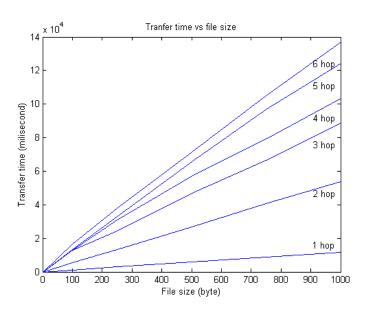


Figure 5.15: Transfer time vs file size

Table 5.6: Percentage of nodes in zone x

Num nodes	Zone 0	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
4	25	50	25	N.A	N.A	N.A	N.A
16	6.25	12.5	18.75	25.0	18.75	12.5	6.25
25	4.0	12.0	20.0	28.0	36.0	N.A	N.A
36	2.77	13.8	25.0	30.5	19.4	8.33	N.A
49	2.0	14.2	28.57	42.85	12.24	N.A	N.A
64	1.56	10.9	21.8	32.8	28.1	4.68	N.A
81	1.23	12.3	27.16	37.03	20.98	1.23	N.A
100	1.0	12.0	26.0	39.0	19.0	3.0	N.A

From Figure 5.14 and 5.15, it is noted that the data transfer time increases monotonically with the number of hops. Therefore, it is crucial to keep the hop distance from the sender node to the sink short. From Table 5.6, it can be observed that using the Bluewave algorithm, most of the nodes will be three hops away from the sink. This is accounted by the circular geometrical shape of the zone. Assuming that most factory floor setting is rectangular and the center of the zone is at a corner, the middle zones will occupy the most area. Therefore, with the machinery uniformly distributed in the factory, most of the machines are found to be in zone 3. In the worst case scenario, as indicated by Figure 5.15, the time taken for sending 50kb data through six hops is 8.3s.

5.5.2 Route Setup

When a node wants to send a message to the sink, it will attempt to contact a node in the lower zone and pass the message to it. Upon receiving a message from a higher zone node, the node will relay the message to a lower zone node. This process will continue until the message reaches the sink. Two methods to contact the nodes in the lower zone are 1) memory inquiry 2) memoryless inquiry.

In the memory method, each node during the discovery phase will store the address of the node that has contact it. Therefore, during the route setup, the node will successfully find a lower zone node in the first attempt.

In the memoryless method case, each inquiry may return address in any of the following case 1) higher zone, 2) equal zone, 3) lower zone. Each inquiry attempt is defined as a trial. A trial is successful only when it returns the address of a node in the lower zone.

When the returned address is a node of a higher or equal zone, the trial is considered a failure.

In the memoryless case, the trials are in a sequence of independent and identically distributed Bernoulli, therefore the number of trials required to achieve success has a geometric distribution.

$$P(X = n) = (1 - p)^{n-1}p$$
 where $p = \frac{Connectivity_lower_node}{Total_connectivity}$

It is estimated that the average number of trials a node has to conduct to obtain the first successful trial is $\frac{1}{p}$.

Thus, the average number of trials need to transverse through a zone is as follows:

$$E(z_x) = \frac{\sum_{i} \frac{1}{p_i}}{|z_x|},$$

where $i \in 1...n$ in the set Z_x

The mean time of a route setup in the topology:

$$E(route_setup) = \sum_{x=i} \frac{|z_x|}{N} E(z_x).$$

Figure 5.16 depicts the average time needed for route setup in both the memory and memoryless case. As observed, the memoryless case will give the best average setup time of approximately 3s. However, the memoryless route setup may not be possible all the time due to unexpected node failure of the known lower zone nodes. In such events, memoryless route setup has to be performed. Assuming the worst case whereby all known lower zone nodes have failed, the average time needed for memoryless search

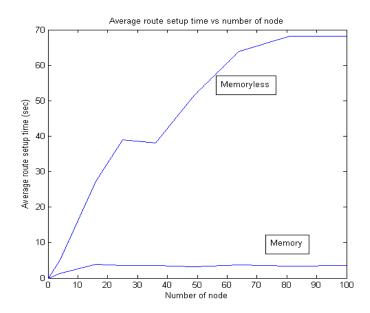


Figure 5.16: Average route setup time

through all lower zone nodes is found to be 68.27s as illustrated in Figure 5.16. The time taken for data transfer phase for worst case scenarios is the sum of file transfer time for 50kB and memoryless route setup for machine in Zone 6, which is 76.57s. This value is lower than the required hard deadline time of 115s.

5.6 Conclusions

Ad-hoc wireless networks are making significant impact in many arenas from smart home to factory automation. Till now, there is no dominant ad hoc wireless protocol that can cater to all applications. Currently, most of the current wireless protocols focus on the mobility issues and are mainly suitable for mobile narrowband radio devices. The mechanisms that worked well for wireless mobile ad hoc network proved to be redundant, and instead add to processing overhead in a static wireless factory setting.

This paper presented the development and implementation of a new wireless rout-

ing protocol, Bluewave, cater specifically for wireless Conditional-Based Maintenance applications in a factory setting. The developed protocol utilized the features of Bluetooth technology when performing the route setup. The Bluewave displayed features that shorten initialization time and route setup time. The alternate zone idle and active state changes reduced the time needed for initialization drastically as compared to protocols using flooding mechanism.

Through the experimental data collected, Bluewave performance in reducing the response time of the system to meet the soft deadline of the data transmission job in a CNC fault-monitoring system is demonstrated. This reinforces the usefulness of Bluewave for industrial conditional-based monitoring systems which use Bluetooth as the wireless technology.

Chapter 6

Conclusions

6.1 Summary of Contributions

Wireless technologies have caused a profound impact in many aspects of our daily lives. It has revolutionalized the way we work, the way we play and the way we communicate with one another. Now, with new emerging wireless technologies, concurrent voice and data communication via wireless devices is now also possible. This allows short range wireless personal communication technologies to play a key role in enabling new applications. This thesis made use of different wireless technologies to develop useful applications in three different arenas; home automation, Education and Industry automation.

First, in home application, a comprehensive architecture that connects mobile devices to LAN PC is presented. The proposed architecture provides solution for overcoming problems such as dynamic IP of NAT, firewall and WAP gateway Timeout. The architecture employs distributive IP redirection method to minimize the resources use in connecting mobile devices to PC.

Then, a secure wake-on-ring hardware device was constructed to cater a power-off PC.

The device work on DFTM principle to allow authenticate user to boot up the PC via mobile devices.

Following that, in education, based on the architecture a mobile simulator that allows mobile devices to utilize PC resources to run spreadsheet simulation was developed. The resultant Mobile Spreadsheet-based PID Control Simulation is used to aid students' understanding in PID control system. The student's feedback obtained verified the usefulness of the Image processing techniques developed for the mobile simulator and also revealed the impact of mobile computing on tertiary education.

Subsequently, the thesis presents work in the use of short range wireless technology in AT application. The "Bus Notification System for the Disabled Bus Commuter" developed enable the visually-impaired, physically-disabled and elderly commuters to take the public bus service independently.

Finally, in Industry automation, a new wireless routing protocol that caters specially for sensor network in factory automation is proposed. The performance of the proposed protocol is compared with conventional routing method. Experimental data is furnished to support the merits of the proposed protocol.

The work performed on different wireless technologies for different fields can be unified to produce a grander architecture for Mobile-to-Machine communication. This will allow mobile online condition-based maintenance of machinery. The Bluewave algorithm will provide the diagnostic portion while the architecture proposed for mobile to remote PC

control will allow the user to remedies instruction to the machinery. Image processing techniques developed will provide good quality image capture by camera that monitor the machinery to user mobile device. The Mobile-to-Machine architecture will allows operation of industrial machinery using mobile devices. This is especially important in countries like China and Thailand, where the mobile devices take up rate is more than desktop computers.

6.2 Suggestions for Future Work

The thesis has presented the research works on the use of wireless technologies in various fields. Further research topics in this area are suggested as follows.

In Chapter 2, "UDP Bombarding" method is devised to resolve the NAT issue. It is realized that the method can only be used when the NAT is in full cone, restricted cone or port restricted cone mode. When the NAT is in a symmetric mode, only the external host that receives a packet can return a UDP packet to the internal host. Therefore, a gateway that maintains continuous TCP connection with the internal PC has to be used. However, an external gateway will increase the resources used in establishing mobile device to PC connection. In order for the architecture to be generic, in the ongoing and future work, a method to detect the status of the NAT and perform appropriate switch between the "UDP Bombarding" and gateway solution is to be devised.

In Chapter 3, the development of the Mobile EXCEL simulator is a pilot project to assess the viability for a greater role of mobile education in the engineering courses.

With the positive feedback obtained from the students, plans can be made to widen the scope for the use of mobile education in the course. Investigations can be carried out on the feasibility of Over-The-Air provisioning content delivery system to deliver lecture materials and applications to mobile devices. Improvement can also be made on the current system to further enhance the image quality and increase the response speed of the system.

In Chapter 5, assumptions that the nodes in Bluewave received continuous power source is made. To widen the scope of applications, it is advisable that Bluewave cater for nodes that are battery-powered. Further researches have to be conducted on power saving measure and node failure handling to improve on the Bluewave algorithm.

Bibliography

- [1] F. Hartung, F. Ramme, "Digital rights management and watermarking of multimedia content for m-commerce applications", *IEEE Communications Magazine*, vol. 38, no. 11, pp. 78 84, Nov. 2000.
- [2] U. Varshney, "Multicast support in mobile commerce applications", *IEEE Transactions on Computer*, vol. 35, no. 2, pp. 115 117, Feb. 2002.
- [3] S. Paurobally, P.J. Turner, "Automating negotiation for m-services", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 33, no. 6, pp. 709 724, Nov. 2003.
- $[4] \ http://www.research.philips.com/technologies/misc/homelab/$
- [5] K. Lunby, O. smordal, A. Larsen, A. Fjuk, "Networked PDAs in a Community of Learners", Proceedings of CSCL, pp. 548 - 549, 2002.
- [6] Roschelle, R. Pea, "A walk on the WILD Side: How wireless Hand held may change CSCL", *Proceedings of CSCL*, pp. 51-60, 2002.
- [7] Ufarooq, Wschafe, R. Rosson, Carroll, "M-education: Bridging the Gap of Mobile and Desktop Computing", Proceedings IEEE International Workshop on Wireless and Mobile Technologies in Education 2002, pp. 91- 94, 2002.

- [8] Y.M. Lee, Feng Cheng, Ying Tat Leung, "Exploring the impact of RFID on supply chain dynamics", Proceedings of the 2004 Winter Simulation Conference, vol. 2, pp. 1145- 1152, Dec. 2004.
- [9] Yilin Zhao, "Mobile phone location determination and its impact on intelligent transportation systems", *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 1, pp. 55- 64, Mar. 2000.
- [10] A. Willig, K. Matheus, A. Wolisz, "Wireless technology in industrial networks", Proceedings of the IEEE, vol. 93, no. 6, pp. 91- 94, June 2002.
- [11] Mark Beaulieu, Wireless Internet Application and Architecture, 1st ed. Addison-Wesley: London, 2002.
- [12] S. Thomson, T. Narten, "IPv6 stateless address autoconfiguration", IEFT RFCInternet Engineering Task Force, Request for Comments: 1971, Aug. 1996.
- [13] C. Bissell, D. Chapman, "Modelling applications of spreadsheets", IEE Review, vol. 35, no. 7, pp. 267 - 271, July 1989.
- [14] M.C. Liu, "Fault tree analysis-using spreadsheet", Reliability and Maintainability Symposium Proceedings, pp. 513 516, Jan. 1990.
- [15] C. Bissell, "Spreadsheets in the teaching of information engineering", Engineering Science and Education Journal, vol. 3, no. 2, pp.89 96, Apr. 1994.

- [16] James R. Evans, "Spreadsheets as a Tool for Teaching Simulation", INFORMS

 Transaction on Education, vol. 1, no. 1, Sept. 2000.
- [17] G. Lawton, "Machine-to-machine technology gears up for growth", *IEEE Computer*, vol. 37, no. 9, pp. 12 15, Sept. 2004.
- [18] F. Benzi, G.S. Buja, M. Felser, "Communication Architectures for Electrical Drives", IEEE Transactions on Industrial Informatics, vol. 1, no. 1, pp. 47 - 53, Feb. 2005.
- [19] http://ieee1451.nist.gov/
- [20] Bob Craft, "Appropriate Use of Wireless Technology in Condition Monitoring", SKF Revolutions, Volume 9, pp. 4 - 5.
- [21] A. Alheraish, "Design and implementation of home automation system", *IEEE Transactions on Consumer Electronics*, vol. 50, no. 4, pp. 1087 1092, Nov. 2004.
- [22] D.B. Strecker, Pounders, "Wireless coke oven control and data acquisition", *IEEE Industry Applications Society Annual Meeting*, vol. 3, pp. 2057 2060, Oct. 1994.
- [23] Christopher McLean, Dave Wolfe, "Intelligent Wireless Condition-Based Maintenance",
 - http://www.sensorsmag.com/articles/0602/14/main.shtml
- [24] John cox, "Bluetooth gets greasy and gritty",

 http://www.networkworld.com/news/2003/0728bluetooth.html

- [25] http://www.openshark.com/
- [26] http://www.intellisync.com/
- [27] Andrew Harbourne, Professional Java Servlets 2.3, 1st ed. Wrox Press: USA, 2002.
- [28] Wrox Multi Team, Professional Java Server Programming J2EE Edition, 1st ed. Wrox Press: USA, 2002.
- [29] Jason Weiss, Java Cryptography Extensions: Practical Guide for Programmers, Morgan Kaufmann, 2002.
- [30] http://jakarta.apache.org/commons/httpclient/
- [31] Jess Garms, Daniel Somerfield, Daniel Somerfield, Professional Java Security, 1st ed. Wrox Press: USA, 2001.
- [32] T. Socolofsky, C. Kale, "A TCP/IP Tutorial", Network Working Group, Request for Comments: 1180, Jan. 1991.
- [33] Bob Quinn, David Shute, Windows Sockets Network Programming, 1st ed. Addison-Wesley: London, 1995.
- [34] Y. Goland, E. Whitehead, A. Faizi, S. Carter, D. Jensen, "HTTP Extensions for Distributed Authoring – WEBDAV", Network Working Group, Request for Comments: 2518, Feb. 1999.

- [35] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, L. Stewart, "HTTP Authentication: Basic and Digest Access Authentication", Network Working Group, Request for Comments: 2617, Jun. 1999.
- [36] Laura Naismith, Peter Lonsdale, Giasemi Vavoula, Mike Sharples, "Literature Review in Mobile Technologies and Learning", University of Birmingham NESTA Futurelab, Report 11, 2005.
- [37] k. Lunby, O. smordal, A. Larsen, A. Fjuk, "Networked PDAs in a Community of Learners", Processdings of CSCL, pp. 548 - 549, 2002.
- [38] Stefan Berger, Rainer Mohr, Holger Nosekabel, Klaus J.S, "Mobile Collaboration Tool for University Education", Proceedings of the Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE03), 2003.
- [39] Yuh-Shyan Chen, Tai-Chien Kao, Gwo-Jong Yu, Jang-Ping Sheu, "A mobile butterfly-watching learning system for supporting independent learning", Wireless and Mobile Technologies in Education Proceedings (2nd IEEE International Workshop), pp. 11 18, 2004.
- [40] L. Mifsud, "Alternative learning arenas-pedagogical challenges to mobile learning technology in education", Wireless and Mobile Technologies in Education Proceedings (IEEE International Workshop), pp. 112 116, Aug. 2002.
- [41] http://www.bbc.co.uk/schools/revision/

- [42] O. Malley, Vavaula, Glew, JP, Talyor, Sharples, Lefrere, "Guidelines for Learning/Teaching/Tutoring in a Mobile Environment"
 http://www.mobilelearn.org
- [43] F. Nah, K. Siau, H. Sheng, "The Value of Mobile Applications: A Utility Company Study", Communications of the ACM, vol. 48, no. 2, pp. 85-90, Feb. 2005.
- [44] Raymond R. Hill, "Process Simulation in EXCEL for a Quantitative Management Course", INFORMS Transactions on Education, vol. 2, no. 3, May 2002,
- [45] K.J.Astrom, B.Wittenmark, computer-controlled Systems, 2nd ed. Prentice Hall: New Jersey, 1990.
- [46] Jan Krikke, "Thin Clients Get Second Chance in Emerging Markets", *Pervasive Computing IEEE*, vol. 3, no. 4, pp. 6 10, Oct-Dec 2004.
- [47] Starner. T, "Thick clients for personal wireless devices", IEEE Computer, vol 35, no 1, pp. 133 135, Jan. 2002.
- [48] Dargos A Manolescu, "Java Thin client Revisited: An architecture for responsive, live wireless Applications", 04 Nov. 2002, http://itpapers.zdnet.com
- [49] http://henson.newmail.ru/j2me/Float.htm
- [50] Tom Krazit, "Cell Phones Sales Skyrocket", 08 Dec. 2003, http://www.pcworld.com/news/article/0,aid,113788,pg,1,00.asp

- [51] K.K. Tan, H.L. Goh, "Complete Mobile Email Management", proceeding of International Symposium of Santa Caterina on Challenges in the Internet and Interdisciplinary Research, 2004.
- [52] Marco Cantu, Mastering Delphi 7, 1st ed. SYBEX: London, 2003.
- [53] Rafael C. Gonzalez and Richard E. Wood, Digital image processing, 2nd ed. Prentice Hall: New Jersey, 2002.
- [54] http://www.mathworks.com/
- [55] Shu-Ching Yang, "Information display interface in hypermedia design", *IEEE Transactions on Education*, vol 43, no. 3, pp. 296 299, Aug. 2000.
- [56] Palo Alto, "HP and Partners Build on HP's Pocket PC Platform to Bring Assistive Technology Advances to People with Disabilities", http://www.hp.com/hpinfo/newsroom/press/2003/031015a.html
- [57] Jennifer Bray, Charles F. Sturman, "Bluetooth 1.1: Connect Without Cables", 2nd ed. Prentice Hall: New Jersey, 2002,
- [58] "Singapore Association of the Visually Handicapped Annual Report", http://www.savh.org.sg/publications/annual_reports/2003/swd_report.pdf
- [59] "Bluetooth in Public Transportation Bus Fleet of Milan, Italy", http://whitepapers.zdnet.com/casestudy.aspx?dtid=3&scid=1121&docid=113218

- [60] Mohammad llyas, The handbook of Ad hoc wireless networks, CRC PRESS: London, 2003.
- [61] C.E. Perkins, P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers", Proc. ACM SIGCOMM, vol. 24, pp. 234-244, Oct. 1994.
- [62] S. Murthy, J.J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks", ACM Mobile Networks special Issue on Routing in Mobile Communication Networks, pp. 183-97, Oct. 1996.
- [63] K.K. Kasera, R. Ramanathan, "A location management protocol for hierarchically organized multihop mobile wireless networks", *IEEE ICUPC*, pp. 158-162, 1997.
- [64] C.E. Perkins, E. M. Royer, "Ad-hoc On-Demand Distance Vector Routing", Proc.
 2nd IEEE Wksp. Mobile Comp. Sys. and Apps., pp. 9C100, Feb. 1999.
- [65] D.B. Johnson, D.A. Maltz, "Dynamic Source Routing in Ad-Hoc Wireless Networks", Mobile Computing, pp. 153-81, 1996.
- [66] Yangsheng Xu, Ming Ge, Ruxu Du, "A real-time monitoring and diagnosis system for manufacturing automation", Proceedings ICRA '04. 2004 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1424 1429, Apr. 2004.
- [67] Z.J. Haas, "A New Routing Protocol for the Reconfigurable Wireless Networks", Proceedings of IEEE ICUPC, vol. 2, pp.562 - 566, Oct. 1997.

- [68] http://www.ubisense.net/Hardware/Ubisensor.htm
- [69] Y.B. KO, N.H. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks", Proceedings of ACMDEEE MOBICOM'98, Vol.6-4, 2000.
- [70] Bill Drake, "Bluetooth gets a filling", http://www.isa.org
- [71] Theodore S. Rappaport, Wireless Communications, Principles and Practices, 2nd Ed. Prentice Hall, 2002.
- [72] Sick, B., "Monitoring the wear of cutting tools in CNC-lathes with artificial neural networks", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 3381 3384, Apr. 1997.
- [73] Xiaoli Li, Han-Xiong Li, Xin-Ping Guan, Du, R., "Fuzzy estimation of feed-cutting force from current measurement-a case study on intelligent tool wear condition monitoring", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 34, no. 4, pp. 506 512, Nov. 2004.
- [74] Kok Kiong Tan, Sunan Huang, "Fault Detection and Diagnostic Using Neural Networks Work Design", International Symposium on Neural Networks, Jun. 2006.
- [75] http://www.bluez.org/
- $[76] \ \mathrm{http://www.dlink.com/}$

Author's Publications

Journal Papers:

- 1. H.L. Goh, K.K. Tan, "Complete Mobile Email Management", Computers and Electrical Engineering, Vol. 31, Issue 3, pp. 241-261, May 2005.
- 2. H.L. Goh, K.K. Tan, "Simulation Using Handheld Devices", *International Journal of Engineering Education*, Vol. 26, No. 1, pp. 1093-1101, 2005.
- 3. K.K. Tan, H.L. Goh, "Development of a Mobile Spreadsheet-based PID Control Simulation System", *IEEE Transaction on Education*, Vol. 49, Issue 2, pp. 99-207, May 2006.
- 4. H. L. Goh, K.K. Tan, S. Huang, C. W. De Silva, "Bluewave: Development of Wireless Protocol for Industrial Automation", *IEEE Transaction on Industrial Automation*, Vol. 2, Issue 4, pp. 221 230, Nov. 2006.

Conference Papers:

1. K.K. Tan, H.L. Goh, "Session Initiation Protocol, Industrial Technology", *IEEE ICIT*'02. 2002 IEEE International Conference, vol. 2, pp. 1310 - 1314, 11-14 Dec. 2002.

2. K.K. Tan, H.L. Goh, "Complete Mobile Email Management", proceeding of International Symposium of Santa Caterina on Challenges in the Internet and Interdisciplinary Research, 2004.

Chapters in Books:

"Assistive Technologies for the Visually-Disabled", Springer-Verlag Berlin Heidelberg, 2007.

Patent:

K.K. Tan, H.L. Goh, "Complete Mobile Email Management", US Patent Pending,
 U.S. serial no. 10/549,045.

Awards:

 "Remote Access to PC via Mobile Devices", Commendation Award, Open Award Competition Category 2004, Tan Kah Kee Young Inventors' Award by the Tan Kah Kee Foundation,

 $http://www.eng.nus.edu.sg/research/2004/2004C3_026.htm$

2. "Bluetooth Convoy", Merit Award Defense Science Award Competition Category 2005, Tan Kah Kee Young Inventors' Award by the Tan Kah Kee Foundation, http://www.ece.nus.edu.sg/showcase/StudAchiev/achievements/aca/TKK_DSC.pdf

3. STARTUP@Singapore competition: NUS INTRO AWARD (2005),

http://www.ece.nus.edu.sg/showcase/StudAchiev/achievements.htm

Press:

1. "Access your home PC on the go ... with your cell phone", Straits Times, 10 Sep

2003.

Others:

- K.Z. Tang, H.L. Goh, K.K. Tan, and T.H. Lee, "Knowledge-Based Control via the Internet", International Journal of Control, Automation and Systems, Vol. 2, No. 2, pp. 207-219, June 2004.
- 2. K.K. Tan, Y.L Lim, H.L. Goh, "Remote Adaptive Control and Monitoring", *Industrial Technology IEEE ICIT '02. 2002 IEEE International Conference*, Vol. 2, No. 2, pp. 892-896, 11-14 Dec. 2002.