The University Of Sheffield.

Intelligent Remote Monitoring and Management System for Type 1 Diabetes

By

Bilal Ahmad Malik

Thesis Submitted for the Degree of Doctor of Philosophy

Department of Electronic and Electrical Engineering

July, 201

Abstract

The work presented in this thesis focuses on developing a telemedicine system for better management of type1 diabetes in children and teenagers. The research and development of the system is motivated by the inadequate communication in the current system of management of the disease, which results in non-compliance of patients following the regimen. This non-compliance generally results in uncontrolled blood glucose levels, which can result in hypoglycaemia, hyperglycaemia and later life health complications. This further results in an increase in health care costs. In this context, the thesis presents a novel end-to-end, low cost telemedicine system, WithCare+, developed in close collaboration between the University of Sheffield (Electronics & Electrical Engineering) and Sheffield Children's Hospital. The system was developed to address the challenges of implementing modern telemedicine in type 1 diabetic care with particular relevance to National Health Service children's clinics in the United Kingdom, by adopting a holistic care driven approach (involving all stakeholders) based on specific key enabler technologies such as low cost and reconfigurable design. However, one of the major issues with current telemedicine system is non-compliance of the patients due to invasive procedure of the glucose measurement which could be clearly addressed by non-invasive method of glucose measurement.

Hence, the thesis also makes a contribution towards non-invasive glucose measurement using Near Infrared spectroscopy in terms of addressing the calibration challenge; two methods are proposed to improve the calibration of the Near Infrared instrument. The first method combines locally weighted regression and partial least square regression and the second method combines digital band pass filtering with support vector regression. The efficacy of the proposed methods is validated in experiments carried out in a non-controlled environment and the results obtained demonstrate that the proposed methods improved the performance of the calibration model in comparison to traditional calibration techniques such as Principal Component Regression and Partial Least Squares regression.

Dedication

This work is dedicated to my late grandparents, Abdul Ahad Bhat and Sara Bano, whose sacrifice and unconditional love made this journey possible. I will miss them always and love them forever.

To My Parents and My Family

To My Children Yaeesh, Fatima and Ibrahim

Your patience and love made this long journey possible

Your support and encouragement made this journey possible

This work is also dedicated to millions of children and their families worldwide who suffer from diabetes. The thought of doing something helpful for these children motivated me to keep on going.

Acknowledgements

First and foremost, I wish to express my sincere gratitude to my supervisor Dr. Mohammed Benaissa for his guidance and patience at every stage of this research. Without his support and invaluable feedback, it would not have been possible to complete this work. I acknowledge his contribution in enhancing my knowledge on the subject.

I am also grateful to Dr. Neil Wright, Consultant at the Children's Hospital, Sheffield and his team for allowing me to attend diabetic clinics to understand the current system of type 1 diabetes management and sharing the information. I would also like to acknowledge their feedback from the inception of WithCare+ to the clinical trial.

I would like to express my gratitude for the studentship award to the Engineering and Physical Sciences Research Council (EPSRC) without which it would not have been possible for me to carry out this work.

I thank my fellow labmates: Anastasios Kanakis, Zia Khan, Krishna Chaitanya, Jee K Chong for acting as volunteers during the field testing of the WithCare+. My special thanks go to Krishna Chaitanya for proof reading this thesis. An extension of my gratitude goes to all my friends in Sheffield who have helped me since I arrived in Sheffield. My thanks also go to the administrative staff and technical staff at the University of Sheffield for their support and help during the period of my research.

Words cannot express my thanks and appreciation to my parents and late grandparents for their support and inspiration which encouraged me to return back to studies after working for 12 years. Although they did not get the chance to attend school, they inspired me to obtain the highest possible degree. Their love and affection kept me going during the difficult times of my life. Huge thanks to my uncles, brothers and sisters for their love and support throughout my life.

Last but not least, I need to express my thanks to my wife and children for their love and encouragement during last four years. One phrase "Dad! Smile please" from my 6 year old daughter had a profound impact on my personality and kept me going during hard times. Finally, I would like to thank my creator for blessing me with the lovely family and excellent opportunities throughout my life.

Contents

List of Figures

List of Tables

Chapter 1 **Introduction**

Diabetes mellitus, commonly known as “diabetes”, is a chronic disease which has been recognised for centuries. It is a metabolic disorder which affects the metabolism of carbohydrates and results in high blood sugar. The blood sugar is raised mainly for two reasons: either the pancreas does not produce enough insulin, which is known as type 1 diabetes, or because cells do not respond to insulin produced by the pancreas, which is referred to as type 2 diabetes. The other known form of diabetes is gestational diabetes, which occurs during the later stages of pregnancies in some women but normally disappears after birth [1,2].

Type 1 diabetes (also referred to as Insulin Dependent Diabetes Mellitus) is a form of diabetes that results from the autoimmune destruction of insulin-producing beta cells of the pancreas, which results in the body not producing insulin [1]. This form of diabetes can affect people of any age, but is more prevalent in children or young adults and the reason for its occurrence is not fully understood. There is no known cure for type 1 diabetes, however the disease can be managed. Since this form of diabetes generally occurs in children, more challenges are faced in managing the disease.

1.1 Current Management of Type 1 Diabetes

The management of type 1 diabetes involves the measurement of blood glucose level (BGL) and insulin intake 3-4 times a day by the patient or caregiver. In addition, patients have to take care of their diet and exercise to maintain their BGL at near normal levels. The patient normally records their BGL before meals, and if it is within the normal range, the patient injects insulin in proportion to the carbohydrates intake. If the glucose level is higher than expected, the patient injects extra insulin, which is known

as bolus insulin, to avoid any occurrence of hyperglycaemia. On the other hand, if the glucose level is low, which is normally the case when the patient wakes up in the morning, they eat something to bring the glucose level back within the normal range and avoid the occurrence of hypoglycaemia.

Patients generally need to prick their finger with a needle and use a glucometer to record BGL. They are expected to record their glucose readings and insulin intake in a diary provided by the health-care centre. This process is repeated 3-4 times daily by the patient from the onset of the disease for the rest of their life.

The patient has to visit the healthcare centre usually every 3 months for a routine examination. During this visit, the glycosylated haemoglobin or glycated haemoglobin, (also known as haemoglobin A1c or HbA1c) of the patient is measured in the laboratory, and this identifies the average plasma glucose concentration. This measure provides the clinician with a much better indication of long term glycaemic control of the patient, as it reflects the mean glucose level over the previous period, which is usually 3 months. A team of health-care professionals, i.e. the nurse, doctor and dietician, is involved during the visit. This team advises the patient based on glucose levels, the insulin intake from the patient diary records as well as the HbA1c laboratory report. Higher amounts of HbA1c indicate poor control of blood glucose levels, which have been associated with cardiovascular disease, nephropathy, and retinopathy. The clinician can also advise the patient to visit the ophthalmologist, cardiologist, nephrologist and psychiatrist from time to time if needed [1, 2].

The healthcare providers face a number of challenges in the current system of management of the disease, especially in the context of teenagers. The issue with

adolescents is that they are in a transition period where they take over the management of BGL from their parents. As they grow older they may exert more independence and may not adhere as strictly to the desired regime as they did as children. Often teens do not take measurements regularly, and their parents do not have this information. This leads to uncontrolled BGL, which if ignored can lead to episodes of hyperglycaemia [1].

The process discussed in the section above involves communication on the part of the patient in terms of the daily glucose measurements and insulin intake, and feedback from the diabetic team to the patient in terms of glucose intake, exercise and dietary advice. Inadequate communication and working with uncertainty is generally a serious problem for all fields, but is particularly serious in the field of healthcare. To improve the quality of everyday life of the patient, there needs to be a low cost communication channel between members of the team on a daily basis. The team can consist of the patient, doctor, nurse, clinician, dietician as well as friends and parents. The results of the team effort are dependent on effective communication, which can result in better control of glucose levels. The Diabetes Control and Complications Trial (DCCT) concluded that maintaining controlled blood glucose levels can reduce the risk of long-term diabetic complications [3].The current management of the disease clearly lacks the communication among the members of the diabetic team on a daily basis, even in technologically advanced places of the world.

A well-engineered, low-cost, reliable telemedicine system with two-way communication is an ideal candidate for addressing the management issues of type 1 diabetes. The next section explains how such a system can help to improve the management of type 1 diabetes.

1.2 Telemedicine for Effective Management of Type 1 Diabetes

Researchers have evaluated and are evaluating the use of communications technology in the implementation and performance of telemedicine activities, and have examined the impact of telemedicine on medical care in terms of cost, quality and access [4-7]. Telemedicine has become a growing new interdisciplinary field, which will eventually contribute to improving the quality of healthcare for everyone. However, successful implementation of this vision depends not only on innovative telemedicine applications but also on computing and networking technical readiness. Furthermore, many ethical, social and political problems arising in telemedicine need technical solutions.

Most of these systems, however, suggest solutions based on the assumption that patients are motivated to take their blood glucose measurements as per the advice of health-care providers. The challenges of type 1 diabetic care for health-care providers, especially among teenagers, are much more complicated from those of type 2 patients [8,9]. Thus it is pertinent that more studies are conducted to demonstrate that telemedicine could improve clinical outcomes for type 1 diabetes in children and young adults, before it can be adopted widely in healthcare.

The recent advances in communications technologies, information technology and miniaturisation of electronic components offer considerable scope to refine and develop the concept of telemedicine into a clinically useful tool where the technology is used more pervasively to monitor and manage chronic conditions such as type1 diabetes. The system must be care driven, low cost and holistic in its approach. It must also be ubiquitous, secure and require minimal effort by the patient and clinician in order to be effective but at the same time should overcome the perception of loss of privacy. Another important feature of the technology is that it must be easily reconfigurable to

support a range of sensors and communication platforms. It is also essential to understand what technological innovations may be acceptable by both patients and health service providers and how widely they may or may not be embraced. These issues are crucial to the translation of telemedicine into clinical care in particular for chronic conditions such as type 1 diabetes, specifically in teenage patients where engagement with treatment strategies is not always forthcoming.

In this context, this thesis presents a novel end-to-end telemedicine system, WithCare+, developed in close collaboration between the University of Sheffield (Electronics & Electrical Engineering and Psychology departments) and Sheffield Children's Hospital and targeted at type 1 diabetes management in children and young adults. The system was developed to address the challenges of implementing modern telemedicine in type 1 diabetic care with particular relevance to NHS children's clinics in the UK, by adopting a holistic care driven approach (involving all stakeholders) based on specific key enabler technologies such as privacy aware security, low power, low cost and reconfigurable design. Although WithCare+ can resolve many difficulties in monitoring, feedback and management of patient care, there still remains one challenge which cannot be addressed by the current telemedicine system. This is the issue of the finger prick method involved in the measurement of glucose 3-4 times on a daily basis from the onset of disease for the rest of life. This method is clearly inconvenient and results in many patients not adhering to the required daily measurements.

This challenge can be addressed by non-invasive glucose measurement. Although the instrumentation for the non-invasive glucose measurement is available, the calibration of the instrument for the determination of glucose certainly remains one of the challenges for clinical translation of the technology. Near infrared spectroscopy (NIR),

due to its advantages such as low water absorbance, little or no sample preparation, and high signal to noise ratio (SNR) has been recognised as one of the most promising technologies for non-invasive glucose monitoring [10-13].

The aim of non-invasive glucose monitoring based on NIR spectroscopy is to transmit a limited band of radiation to a selected region of the body. The light intensity of transmitted or reflected radiation is changed as a function of the tissues and is correlated with the variation of the concentration of glucose. However, determination of the concentration of glucose from NIR spectra is not easy, as the collected spectra are affected by instrument noise, baseline variations and temperature variations. The collected spectra are also weak in intensity, broad and overlapped. The challenge of quantitative data analysis is addressed by multivariate calibration methods augmented by signal pre-processing techniques.

1.3 Research Questions

1) How to design a telemedicine system which will be adopted by the healthcare professionals at Sheffield Children's Hospital for management of type 1 diabetes.
2) How to improve the communication between the members of a diabetic team i.e. the patients, clinician, dietician and nurses involved in management of type 1 diabetes.
3) How to improve the multivariate calibration techniques for determination of glucose concentration from NIR spectra. The aim of this preliminary study is to contribute towards clinical translation of the non-invasive glucose measurement.

1.4 Thesis Aims and Scope

The aim of this thesis is to design, develop and implement "WithCare+", for remote monitoring and better management of type 1 diabetic patients by exploiting the current advances in telecommunication, information technology and mobile technology. The further aim of the thesis is the preliminary study towards the vision of future WithCare+, where the patient can measure their BGL non-invasively without pricking the skin. Researchers are working on different technologies to achieve the goal of non-invasive glucose measurement. However, the scope of the thesis is limited to development of the robust calibration models for Fourier transform near infrared (FT-NIR) spectrometer. The efficacy of the proposed models is validated for the determination of glucose concentration from NIR spectra of an aqueous solution composed of glucose, urea and triacetin in a phosphate buffer solution. The concentrations of the components are selected to be within their physiological range in blood and the experiments were carried out without controlling the environmental conditions such as temperature and humidity, to show that the proposed models can effectively suppress most of the experimental variation.

1.5 Key Contributions

The motivation for this study was to look into the various aspects where technology can be exploited in an innovative way to help children suffering from diabetes. The main contribution of this thesis is the design, development and testing of a telemedicine system. The new system was designed based on the requirements gathered from the diabetic team at Sheffield Children's Hospital. Initially the system was developed on Windows operating system using C# for developing business logic, asp.net for developing the user interface and SQLServer to develop the database. This version was developed with the view that the information technology (IT) department of the NHS

may prefer to integrate the Windows based system. Later on, the system was developed using open source technology such as PHP, PERL and MySQL on a Linux operating system (Ubuntu) which hosts an Apache Web server keeping in view the advantages offered by the open source technology.

The author has also contributed in designing and developing the proof of concept for Bluetooth communication over the radio frequency communication (RFCOMM) channel profile of Android based smart phones, which are the backbone of WithCare+ communication. The need for this approach arose due to the fact that the native serial port profile (SPP) does not exist in smart phones. The success of this proof of concept led to the implementation of Bluetooth communication between WithCare+ and Android. The same concept is extended to other platforms such as Blackberry.

This thesis also made a contribution in the area of non-invasive glucose measurement by developing novel calibration models by exploiting the advances in machine learning and signal processing, which can not only benefit the type 1 diabetic people but would directly help patients with any other type of diabetes as well. In this study, two novel regression techniques were proposed. The first method combines the locally weighted regression (LWR) with partial least squares regression (PLSR) to improve the performance of PLSR model. The second method couples the digital band pass filter with the support vector regression to improve the quality of collected spectra and develop the calibration model with better generalisation. These models were validated for the determination and prediction of glucose concentration from NIR spectra of a mixture composed of urea, triacetin, and glucose. The results when compared with traditional calibration techniques such as principal component regression (PCR) and PLSR demonstrate that the proposed techniques improved the performance of

calibration models and can be validated for determination of glucose concentration from blood serum in future.

1.6 Thesis Structure

Chapter 2 presents an introduction to diabetes, its economic importance and the need for a telemedicine system for remote monitoring of BGL of type 1 diabetic patients. The chapter further discusses state-of-the-art technology in the domain of telemedicine systems.

Chapter 3 explains the system architecture, design, and implementation of the proposed telemedicine system "WithCare+". Here the technologies selected for development and testing are also presented. The difficulties faced during the clinical trial of the system are mentioned.

Chapter 4 presents an overview of non-invasive glucose measurement and describes the near infrared spectroscopy as the choice for the non-invasive glucose measurement. The chapter further explains the theory of pre-processing and calibration methods for NIR data. In addition, the experimental setup for data collection is explained. The model development with results of proposed novel calibration technique is explained as well.

Chapter 5 gives an overview of support vector regression and digital band pass filtering as a pre-processing calibration technique. In addition, the model development with results of proposed novel calibration technique is explained.

Chapter 6 concludes this report. The recommendations are presented here and areas for further work are identified.

1.7 Publications

1.7.1 Book Chapters in Edited book

1. BILAL AHMAD MALIK AND MOHAMMED BENAISSA, "Glucose Chemistry" in Preedy, Victor R., ed. *Dietary Sugars: Chemistry, Analysis, Function and Effects*. Vol. 3. Royal Society of Chemistry, 2012. ISSN 2045-1695 DOI:10.1039/9781849734929-00077.
2. MOHAMMED BENAISSA, AMNEH MBAIDEEN AND BILAL AHMAD MALIK "Assay of Glucose Using Near Infrared (NIR) Spectroscopy" in Preedy, Victor R., ed. *Dietary Sugars: Chemistry, Analysis, Function and Effects*. Vol. 3. Royal Society of Chemistry, 2012. . ISSN 2045-1695, DOI: 10.1039/9781849734929-00286.

1.7.2 Paper Publications

1. B.Malik and M.Benaissa, "Determination of glucose concentration from near-infrared spectra using locally weighted partial least square regression," in *Engineering in Medicine and Biology Society (EMBS), 2012 Annual International Conference of the IEEE*, 2012, pp. 6169-6171. Conf Proc IEEE Eng Med Biol Soc; 2012;2012:6169-71 PMID: 23367337.
2. Bilal Malik and M.Benaissa, "Determination of glucose concentration from near-infrared spectra using support vector regression" presented at International Conference on New Trends in Chemometrics and Applications at Ankara University, Turkey.
3. Bilal Malik, Krishna and M.Benaissa, "Determination of glucose concentration from near-infrared spectra using support vector regression coupled with digital band pass fitler," *submitted to Journal of Chemometrics*.
4. M.Benaissa, B.Malik, A.Kanakis, and N.P.Wright, "Tele-healthcare for diabetes management: A low cost automatic approach," in *Engineering in Medicine and*

Biology Society (EMBS), 2012 Annual International Conference of the IEEE, 2012, pp. 1290-1293.

5. A.Kanakis, B.Malik, and M.Benaissa, "Low cost universal remote patient monitoring system," in *High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESS), 2012 IEEE 14th International Conference on*, 2012, pp. 1587-1591.
6. Mohammed Benaissa, Neil Wright, Anastasios Kanakis, Bilal Malik and Barbara Johnson, "WithCare+ System: A holistic, care-driven approach to telemedicine for type-1 diabetes in children and young adults" *3rd International Conference on Wireless Mobile Communication and Healthcare November* 21–23, 2012 Paris, France.
7. Krishna Chaitanya, Bilal Malik, Mohammed Benaissa, "Prediction of glucose concentration from near infrared spectra using local linear embedded regression combined with a digital band pass filter". Conferentia Chemometrica 2013 Sopron, Hungary.

References:

1. D. R. Hanas, *Type 1 Diabetes, second ed.*: Class Publishing,Barb House,Barb Mews,London W6 7PA,Uk, 2004.
2. B. Higgins, *Type 1 Diabetes in Adults. London*: Royal College of physicians, 2004.
3. "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group," *N Engl J Med*, vol. 329, pp. 977-86, Sep 30 1993.
4. G. T. McMahon, H. E. Gomes, S. H. Hohne, T. M.-J. Hu, B. A. Levine, and P. R. Conlin, "Web-based care management in patients with poorly controlled diabetes," *Diabetes Care*, vol. 28, pp. 1624-1629, 2005.
5. Farmer, Andrew, et al. "A real-time, mobile phone-based telemedicine system to support young adults with type 1 diabetes." *Informatics in Primary Care* 13.3 (2005): 171-178.
6. Farmer, Andrew J., et al. "A randomized controlled trial of the effect of real-time telemedicine support on glycemic control in young adults with type 1 diabetes (ISRCTN 46889446)." *Diabetes Care* 28.11 (2005): 2697-2702.
7. Rami, Birgit, et al. "Telemedical support to improve glycemic control in adolescents with type 1 diabetes mellitus." *European Journal of Pediatrics* 65.10 (2006): 701-705.
8. J. Silverstein, G. Klingensmith, K. Copeland, L. Plotnick, F. Kaufman, L. Laffel, et al., "Care of children and adolescents with type 1 diabetes," *Diabetes Care*, vol. 28, pp. 186-212, January 1, 2005.

9. M. Graue, T. Wentzel-Larsen, E. Bru, B. R. Hanestad, and O. Søvik, "The coping styles of adolescents with type 1 diabetes are associated with degree of metabolic control," *Diabetes Care*, vol. 27, pp. 1313-1317, June 1, 2004.
10. Poddar, Raju, et al. "Non-invasive glucose monitoring techniques: A review and current trends." *arXiv preprint arXiv*:0810.5755 (2008).
11. Khalil, Omar S. "Spectroscopic and clinical aspects of noninvasive glucose measurements." *Clinical Chemistry* 45.2 (1999): 165-177.
12. Khalil, Omar S. "Non-invasive glucose measurement technologies: an update from 1999 to the dawn of the new millennium." *Diabetes Technology & Therapeutics* 6.5 (2004): 660-697.
13. http://www.researchgate.net/publication/215519631_The_Pursuit_of_Noninvasive_Glucose_Hunting_the_Deceitful_Turkey/file/747f0180f0cf2751e6eee0840d5b527b.pdf. [Accessed: 20 Jan 2012]

Chapter 2 **Background**

This chapter explores the literature that focuses on diabetes mellitus, its harmful effects, and associated economic costs, and examines state-of-the-art technology in the area of telemedicine systems specifically designed to improve the diabetic care. Although diabetes insipidus has common symptoms with diabetes mellitus, they are entirely two different conditions with unrelated mechanisms.

2.1 Overview of Diabetes Mellitus

2.1.1 Diabetes Mellitus Concepts

Diabetes mellitus, commonly known as Diabetes, is one of the major chronic diseases associated with abnormally high levels of glucose (sugar) in blood, due to the reason that either the pancreas does not produce enough insulin or cells do not respond to the produced insulin [1,2]. Insulin is a naturally occurring hormone needed to transform sugar into energy required for daily life. The symptoms associated with high blood sugar are polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger). The complications of diabetes constitute a major health problem in modern societies worldwide [1, 2].

Diabetes is a serious chronic disease which causes a high rate of morbidity and mortality all over the world. Diabetes is ranked seventh in the list of diseases for cause of death [3]. In 2012, it has been reported that this disease affects about 371 million people across the globe and it is expected by the year 2030 that disease will increase by 51% [4]. The greatest increase in occurrence is, however, expected to occur in Asia,

where it is predicted that most patients will probably be found by 2030. According to the International Diabetes Foundation (IDF) in 2011, China with the diabetic population of 90 million people topped the list of countries, whereas India with 60 million diabetic population ranks second [4]. The average age of onset is 42.5 years. The high incidence is attributed to a combination of genetic susceptibility plus the adoption of a high-calorie, low-activity lifestyle by India's growing middle class [5]. The increase in incidence of diabetes in developing countries follows the tendency of urbanization and changes in lifestyle [6].

2.1.2 Types of Diabetes

There are three major types of diabetes:

- Type-1 Diabetes also known as insulin dependent diabetes mellitus-IDDM
- Type 2 Diabetes also known as non-insulin dependent diabetes mellitus-NIDDM
- Gestational Diabetes Mellitus

Other rare types of diabetes include those caused by genetic conditions, surgery, drug use, malnutrition, infections, and other illnesses.

2.1.3 Type-1 Diabetes (Insulin Dependent Diabetes Mellitus-IDDM)

Type 1 diabetes is an auto-immune disease where the body's immune system destroys the insulin-producing beta cells in the pancreas, leading to insulin deficiency. This type of diabetes, also known as juvenile-onset diabetes, accounts for about 10% of all people with the disease. There is no known preventive measure against this type of disease. It can appear at any age, although common under 30, and is triggered by environmental factors such as viruses, diet or chemicals as well as genetic factors. At the onset of the

disease most patients are otherwise healthy and of normal weight. People with type 1 diabetes must inject themselves with insulin several times a day and follow a careful diet and exercise plan. Patients normally respond well to insulin, particularly at the early stages [1,2].

Type 1 diabetes is accompanied by irregular and unpredictable hyperglycemia, ketosis, sometimes serious hypoglycemia, eye infections, endocrinopathies and erratic absorption of dietary carbohydrates due to many different reasons[1,2].

2.1.4 Type 2 Diabetes (Non-Insulin Dependent Diabetes Mellitus-NIDDM)

This type is more common usually accounting for about 90% of cases. It is diagnosed in adults but young people can also suffer from this disease. In this type there is a reduced sensitivity to Insulin on a cellular level in addition to inadequate levels of insulin production by the pancreas. It is the most serious type of diabetes; often people do not know they are suffering with it. Three major causes of type 2 diabetes are lifelong bad diet, an inactive or sedentary lifestyle, and being overweight. Type 2 diabetes is becoming more common due to risk factors like older age, obesity, lack of exercise, family history of diabetes [8]. Some changes in lifestyle, importantly increase in physical activities and modest weight reduction can reduce the chances of suffering from type 2 diabetes by as much as 58% [9]. Use of metformin (a drug used as a treatment for type 2 diabetes) along with a good lifestyle and a healthy diet, reduces the risk of development of type 2 diabetes by 31%, but metformin is not effective in patients above 60 years of age [10]. Several former terms like childhood-onset diabetes, juvenile diabetes, and insulin-dependent diabetes mellitus (IDDM) have been replaced by the term ‘type 1 diabetes’; ‘type 2 diabetes’ has also replaced former terms like obesity-related diabetes, noninsulin-dependent diabetes mellitus (NIDDM) and adult-

onset diabetes. Table 2.1 compares some of the characteristics of type 1 and type 2 diabetes.

Table 2.1 General Characteristics of Types 1 And 2 Diabetes Mellitus

Characteristics	Type 1	Type 2
Age at onset	Most commonly < 30 yr	Most commonly > 30 yr
Associated obesity	No	Very common
Ketoacidosis	Yes	No
Twin concordance	<=50%	>90%
Associated with specific HLA-D Antigens	Yes	No
Islet cell antibodies at diagnosis	Yes	No
Islet pathology	Insulitis, selective loss of most β cells	Smaller, normal-appearing islets; amyloid (amylin) deposition is common
Prone to develop diabetes Complications	Yes	Yes
Hyperglycemia responds to oral antihyperglycemic drugs	No	Yes, initially in many patients

2.1.5 Gestational Diabetes Mellitus

Gestational Diabetes Mellitus (GDM) is high blood glucose that develops in 25% of women during pregnancy. But it disappears once the baby is born. Gestational diabetes is fully treatable, but requires careful medical supervision throughout the pregnancy. About 20%–50% of affected women develop type 2 diabetes later in life. Treatment of gestational diabetes reduces serious perinatal morbidity and may also improve the woman's health-related quality of life. On the other hand, untreated gestational diabetes can damage the health of the foetus or mother. Risks to the baby include macrosomia (high birth weight), congenital cardiac and central nervous system anomalies, and skeletal muscle malformations. Increased foetal insulin may inhibit foetal surfactant production and cause respiratory distress syndrome. Hyperbilirubinemia may result from red blood cell destruction. In severe cases, perinatal death may occur, most commonly as a result of poor placental perfusion due to vascular impairment. Labour induction may be indicated with decreased placental function. A Caesarean section may be performed if there is marked foetal distress or an increased risk of injury associated with macrosomia, such as shoulder dystocia [11-13].

In 2008, a study reported that the number of American women entering pregnancy with pre-existing diabetes is increasing at an alarming rate [14]. This is particularly problematic as diabetes raises the risk of complications during pregnancy, as well as increasing the potential for the children of diabetic mothers to become diabetic in the future [11-13]. Other forms of diabetes mellitus include congenital diabetes, which is due to genetic defects of insulin secretion, cystic-fibrosis related diabetes, steroid diabetes induced by high doses of glucocorticoids, and several forms of monogenic diabetes.

2.1.6 Metabolic Mechanisms

All forms of diabetes have been treatable since insulin became available in 1921, and type 2 diabetes may be controlled with medications or diet alone. Insulin is the principal hormone that regulates uptake of glucose from the blood into most cells (primarily muscle and fat cells, but not central nervous system cells). Therefore, deficiency of insulin or the insensitivity of its receptors plays a central role in all forms of diabetes mellitus.

Both type 1 and 2 are chronic conditions that usually cannot be cured. Pancreas transplants have been tried with limited success in type 1 DM; gastric bypass surgery has been successful in many with morbid obesity and type 2 DM. Gestational diabetes usually cures itself after delivery. Diabetes without proper treatments can cause many complications. Acute complications include hypoglycemia, diabetic ketoacidosis, or nonketotic hyperosmolar coma. Serious long-term complications include cardiovascular disease, chronic renal failure, and diabetic retinopathy (retinal damage). Thus it is important that diabetes must be managed well.

Humans are capable of digesting some carbohydrates, in particular those most common in food. Some disaccharides such as sucrose are converted within a few hours to simpler forms, most notably the monosaccharide glucose, the principal carbohydrate energy source used by the body. The rest is passed on for processing by gut flora largely in the colon. Insulin is released into the blood by beta cells (β-cells), found in the islets of Langerhans in the pancreas, in response to rising levels of blood glucose, typically after eating. Insulin is used by about two-thirds of the body's cells to absorb glucose from the blood for use as fuel, for conversion to other needed molecules, or for storage.

Insulin is also the principal control signal for conversion of glucose to glycogen for internal storage in liver and muscle cells. Lowered glucose levels result both in the reduced release of insulin from the β-cells and in the reverse conversion of glycogen to glucose when glucose levels fall. This is mainly controlled by the hormone glucagon, which acts in the opposite manner to insulin. Glucose thus forcibly produced from internal liver cell stores (as glycogen) re-enters the bloodstream; muscle cells lack the necessary export mechanism. Normally, liver cells do this when the level of insulin is low (which normally correlates with low levels of blood glucose).

Higher insulin levels increase some anabolic (building up) processes, such as cell growth and duplication, protein synthesis, and fat storage. Insulin (or its lack) is the principal signal in converting many of the bidirectional processes of metabolism from a catabolic to an anabolic direction, and vice versa. In particular, a low insulin level is the trigger for entering or leaving ketosis (the fat-burning metabolic phase).

If the amount of insulin available is insufficient, or cells respond poorly to the effects of insulin (insulin insensitivity or resistance), or the insulin itself is defective, then glucose will not have its usual effect. Hence, it will not be absorbed properly by those body cells that require it, nor will it be stored appropriately in the liver and muscles. The net effect is persistent high levels of blood glucose, poor protein synthesis, and other metabolic derangements, such as acidosis.

When the glucose concentration in the blood is raised beyond its renal threshold (about 10 mmol/L, although this may be altered in certain conditions, such as pregnancy), reabsorption of glucose in the proximal renal tubuli is incomplete, and part of the glucose remains in the urine (glycosuria). This increases the osmotic pressure of the

urine and inhibits reabsorption of water by the kidney, resulting in increased urine production (polyuria) and increased fluid loss. Lost blood volume will be replaced osmotically from water held in body cells and other body compartments, causing dehydration and increased thirst.

2.1.7 Diabetes in the UK

It has been reported in 2011 that there are 2.9 million people in the UK who have been diagnosed with diabetes and it is estimated that this figure will increase to five million by 2025 [7]. Type 1 diabetes is more prevalent in children and the UK ranks fifth highest in the world for the rate of children diagnosed with type 1 diabetes. Finland, Sweden, Saudi Arabia and Norway have higher rates as compared to the UK as per the available data [7]. UK charities such as Diabetes UK and the Juvenile Diabetes Research foundation (JDRF) warn that it is important that people are aware of symptoms of type 1 diabetes because if left undiagnosed and untreated, it can lead to serious illness and even death. There are about 23,000 children under the age of 17 with diabetes in England and 97% of them have type 1 diabetes. The peak age for diagnosis is between 10 to 14 years of age [7].

2.1.8 Related Complications

Diabetes-related complications include:

- Microvascular diseases (e.g. retinopathy, blindness, nephropathy, and kidney failure).
- Macrovascular diseases (coronary heart disease, stroke, peripheral vascular disease and lower-extremity amputation).
- Influenza and pneumococcal infections.
- Eye complications.

2.1.9 Economic Importance of Diabetes

Diabetes mellitus is a public health concern of considerable economic importance due to its chronic nature, increasing global prevalence, and the serious complications associated with long-term disease duration. Diabetes imposes large economic burden on national healthcare systems and affects both national economies, individuals and their families. Indirect costs of disease include lost productivity caused by morbidity, disability, and premature mortality. The intangible costs of diabetes are the reduced quality of life for people with diabetes due to stress, pain, and anxiety. It is estimated that National Health Service (NHS) spends £10 billion on diabetes, which is about 10% of the NHS budget and currently the total cost (direct costs and indirect costs) is £23.7 billion [15-18]. Data from Europe indicate that the healthcare expenditure for patients with diabetes mellitus is significantly higher for comparable patients without this disease. These and other studies also suggest that the majority of such 'excess' costs can be attributed to diabetes-related complications [15-18].

2.2 Monitoring and Management of diabetic patients

Diabetes mellitus is a chronic disease which cannot be cured except in very specific situations. So, the disease needs to be managed by the patients in order to avoid complications. Generally, the goal of good management is to keep blood sugar levels close to normal ("euglycemia"), without causing low levels of blood sugar known as hypoglycaemia or high levels of blood sugar known as hyperglycaemia. Proper diabetes management involves close monitoring of glucose levels in the blood, insulin intake (for insulin dependent patients), diet and exercise. Blood glucose monitoring refers to the on-going measurement of blood glucose levels and is typically done using a portable device called a glucometer.

2.2.1 Operation of a Glucometer

A glucometer is a battery operated device used by diabetic patients for measurement of their blood glucose concentration from home. Most of the glucometers in the market use an electrochemical method for measurement of glucose concentration in the whole blood. In these meters a test strip acts as a glucose sensor; a small amount of blood drawn from the finger of a patient is sucked by the capillary on the test strip. An electrochemical reaction takes place between the glucose in the blood with an enzyme electrode containing glucose oxidase (or dehydrogenase). The glucose in the blood sample reacts with the glucose oxidase to form gluconic acid, which then reacts with ferricynaide to form ferrocynaide. The electrode oxidizes the ferrocyanide, and this generates a current directly proportional to glucose concentration. Most of the glucometers display results between 20 to 600 mg/dL (1.1-33.3) mmol/L in a small window and it generally takes 3 to 60 seconds for measurement depending on the model of the glucometer.

Monitoring vital parameters from a patient with a chronic condition enables the general practitioner (GP) or healthcare provider to intervene before the patient's state of health becomes critical thereby saving lives while avoiding unnecessary extra costs of emergency treatment. Telemedicine is an ideal candidate to address the challenges of type 1 diabetes management.

2.3 Telemedicine

Recent advances in information and communication technologies (ICTs) such as computers, the internet, and cell phones are changing how people communicate with each other, seek and exchange information. There is great potential of adapting these

advances in ICT to address contemporary global health problems such as access, equality, quality and cost-effectiveness.

Telemedicine literally means “healing at a distance” [19], which implies the utilization of ICT to support healthcare by overcoming the classical barriers of time and distance. It generally includes, but is not limited to, transmission of patient data to healthcare providers at a distance and after the analysis of data by the server software, immediate feedback is sent to the patient by the software or the healthcare professional via the different channels of communication available in the system. It is recognized that there is no definitive definition of telemedicine: Sood (2007) found 104 peer-reviewed definitions of the word [20]. However, the World Health Organization has adopted the following definition:

> “The delivery of healthcare services, where distance is a critical factor, by all healthcare professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of healthcare providers, all in the interests of advancing the health of individuals and their communities.” [19]

Some authors differentiate between terms ‘telehealth’ and ‘telemedicine’, with the latter limited to service provided by clinicians only, and the former service delivered by health professionals in general. However, within the scope of this thesis, these terms are used interchangeably and are synonymous.

The multiple descriptions of telemedicine highlight that it is an evolving technology, as it assimilates both changing health needs and rapid advances in ICT. The main purpose

of telemedicine is to improve interaction between patients and their healthcare provider in order to achieve desired clinical outcomes and reduce monitoring and management costs. The following components are integral to telemedicine:

- A process or a device for accurate patient data collection in digital format.
- Secure transmission of the collected patient data securely to a remote server.
- A set of protocols to ensure system security between client (patient unit) and server.
- Different channels of communication to enhance present level of communication among the stakeholders.
- A facility for immediate alert to healthcare professionals in case of abnormal behaviour.
- Facilities that enable healthcare professionals to send therapeutic advice to the patients in different formats, such as text message, email or video link.

Although there has been renewed research and commercial interest in the area of telemedicine over the last decade, its inception can be traced back to middle of the 19th century [21]. Later developments include the transmission of electrocardiograph data over telephone wires in the early 20th century [22] and developments in the military and space sectors in 1960s that started telemedicine in its modern form [21,23]. Recent advances in ICT and its usage by the general population has been the main driving force for emerging trends in telemedicine in both developed and developing nations [24].

A telemedicine application generally involves the transmission of data from a distant location and interactions between healthcare professionals and other health professionals or between healthcare professionals and patients. A telemedicine

application is generally classified into two basic types, asynchronous (store-and-forward) and synchronous. The former involves exchange of pre-recorded data between stakeholders at different times as in case of email, whereas, in the latter communication takes place in real time with an immediate exchange of information between individuals, such as in videoconferencing. Both of these approaches can transmit data using different media such as text messages, audio, video or still images and can be applied to a wide array of services in diverse settings such as teleradiology, telepathology and teledermatology care [21,24].

2.3.1 Telemedicine for Diabetic Management

Telemedicine can be utilized as a support tool for better management of diabetes by improving communication among the stakeholders of diabetes management team. Telemedicine has an impact on different aspects of patient care such as the informational, clinical, behavioural, structural, and economic. The informational impact is an improvement in the quality of information due to the digital transmission, as compared to handwritten records in a diary, which may be forgotten at home on appointment days or inadvertently left incomplete. Frequent communication of information, instructions and feedback can have a clinical impact in terms of lowered glycated haemoglobin (HbA1C) levels. The behavioural impact involves reminders and frequent therapy adjustments, leading to greater patient empowerment and education. Patients may need to come in to the physician's office for fewer visits, which will have a structural impact. The economic impact of telemedicine is poorly understood at this time and needs more studies to draw any conclusions. Whitten et al. [25] from their study on the cost effectiveness of telemedicine concluded that there is presently no persuasive evidence of cost benefits of telemedicine in comparison to normal care.

2.3.2 Goals of Telemedicine

The main goals of using a telemedicine system are to collect and transmit medical data securely from a patient at a distant location to a server managed by a healthcare provider. The data received at the server is analysed after passing security checks and appropriate feedback is sent to the patient or caregiver. This process of two way communication enhances the interaction among different stakeholders involved with the management of the disease, which can lead to an improvement in clinical outcomes [26].

2.3.3 Benefits of Telemedicine

A successful telemedicine program is robust, effective, user-friendly, and cost-effective. A robust telemedicine system helps the transmission of patient data from remote site to the server of the healthcare provider by maintaining data integrity. A robust system will include time stamping of patient data to avoid data duplication and a protocol between the patient unit and server to manage security. A telemedicine system can be classified as effective if it is adopted by the healthcare providers and improves clinical outcomes. The clinical outcomes for diabetes patients can be measured by the variation in glycaemic value, number of hypoglycaemic events, emergency room visits for diabetes-related events and A1C levels. The effectiveness of the system can be measured by patient and healthcare provider satisfaction levels. Telemedicine management of the disease could be cost-effective in the long run but initial costs may be high. Thus analysis of the cost effectiveness must not be based on short term goals and must instead take into account long-term benefits to compare it with the conventional cost of care [27].

A user-friendly telemedicine system with the above benefits will overcome the problems that have hindered the adoption of these systems [26,28].

2.3.4 Barriers to the Adoption of Telemedicine

Considerable barriers from both healthcare providers and patients currently limit the penetration of telemedicine to deliver routine services, both in industrialized and developing countries. Both technical and systemic obstacles exist, such as fear of technology or fear of change, privacy concerns or a lack of training in ICT. Furthermore, in the case of diabetic management, telemedicine systems may fail to accommodate the brand of glucometer preferred by the patient. These barriers on the part of health stakeholders of diabetes deter the adoption of current telemedicine systems [23,26,27,29].

The lack of studies demonstrating persuasive evidence for the cost benefits of telemedicine in comparison to conventional care is also a barrier to convincing policy-makers to invest in telemedicine [27]. The other major obstacle is the absence of international legal framework and the risk of medical liability for the healthcare providers. The complexity of the system with the potential of hardware or software failure is also a hurdle [30].

It is important that telemedicine must be implemented to the highest ethical standards to ensure that differences in physical and mental ability, age, gender and geographical location will not lead to marginalisation. Telemedicine must be regulated by unambiguous and exhaustive guidelines, applicable worldwide, in order to overcome the challenges for greater acceptability of telemedicine as a routine practice [30].

2.3.5 Research Opportunities in Diabetes Telemedicine

Recent advances in ICT have generated great interest in developing telemedicine care for diabetes management worldwide, even though the evidence of its benefits is limited. There is a strong suggestion that, with the current economic situation, governments are

looking to telemedicine to improve outcomes and lower costs. However, studies of telemedicine to date suffer from a combination of one or more design flaws: small sample size, poor design and user interface, or failure to demonstrate the long-term benefits. There are several reasons for fostering research in the area of diabetes management and future research in ICT for supporting diabetes management should ensure the acceptability of technology by the diabetic team [26,28,31].

Due to small amount of published research on the economic benefits of telemedicine interventions for diabetes, it is difficult to influence medical policymakers who are currently reluctant to invest more in telemedicine. This type of information is needed before telemedicine will be able to find wide acceptability and support from all stakeholders. Thus it is important that more studies are conducted in order to make the telemedicine management of not only diabetes but other chronic diseases as well [25,31].

2.4 Literature Review of Telemedicine Systems for Diabetes

Remote monitoring through telemedicine offers an efficient means of overcoming some of the limitations of the current system of monitoring and management of diabetic patients. This fact has stimulated research into the effectiveness of remote monitoring and the development of prototype systems for diabetes management. These developments were tailored to specific requirements. The limitations of the technologies used (such as cost, complexity and support) have restricted their deployment to a few niche locations.

There is a significant body of published literature on the benefits of remote management of patients with diabetes mellitus. Early incarnations of such remote monitoring systems relied on the telephone as the underlying technology supporting communication. In

recent times however, new approaches that use the internet and smart phones for communications have been developed and this is a common feature of existing commercial ventures.

Currently the diabetic team at Sheffield Children's Hospital use Diasend telemedicine system [32] for management of type 1 diabetic patients. The advantage of this system is that diabetic nurse does not need to install the software provided by different meter companies. The hardware unit of Diasend, a small box, reads most of the current blood glucose meters and data is transmitted to the remote server which is running at the headquarters of Aidera AB, the manufactures of Diasend. The diabetic team have access to the patient's data. The disadvantage of this system is that until the patient visits the clinic for the routine or emergency check-up, the diabetic team does not have access to the BGL of the patient.

McMahon et al. [33] have evaluated a Web-based care management system for patients with poorly controlled diabetes. The aim of this system was to enhance diabetes patient education and to encourage patients' self-empowerment and management abilities. This web based care management system seems to support these objectives. The strengths of the study are the careful design and the novel use of technology, together with education, remote health monitoring, and 2-way communication with a practitioner. This system has the ability to post professionally vetted material on secure websites, has 24 hour accessibility, and is available to individuals in their homes regardless of the distance from their sites of healthcare. Patients with poorly controlled diabetes who adopt such a system and regularly exchange information with their healthcare providers are likely to derive important clinical benefits. With the increase in web connectivity

and worldwide deployment of internet connections, this infrastructure can support many current and emerging healthcare applications.

However, the sample was mainly composed of all men with a mean age of 63 years and had above average level of education. The feasibility of implementation of such a system is restricted by the availability of computer equipment, software, internet access, training, and support. In addition, healthcare agencies may not have the staff required to mount this labour intensive approach. An advanced practice nurse made treatment recommendations, communicated with clinicians, and responded to patients within one working day. In addition, if patients did not log on to the website for two weeks, a coordinator prompted them to do so. Considering these feasibility challenges, the implementation of this system in clinical practice at present is unlikely.

While investigating modem transmission of glucose values for an adolescent population with type 1 diabetes, Chase et al [34] concluded that fortnightly electronic transmission of blood glucose levels and other data in place of clinic visits is a possible alternative to clinic visits every 3 months. The glycaemic control and the incidence of acute diabetes-related complications were comparable between the modem group and control group, even though the modem group did not have a 3-month clinic visit. The cost analysis of the study also revealed that using the modem for transmission of data was cost-effective in comparison to standard care. The other advantage of the modem group was that the children missed fewer days at school and parents missed fewer days at work as a result of omitting the 3-month visit. Care providers also benefitted in terms of saving the clinic space and the clinician's time. The degree of patient satisfaction was reported to be similar in both groups. One important point determined by this study was that the clinicians found it hard to keep the patients motivated to communicate their blood

glucose levels to the clinicians in the standard care after a few years of the onset of disease.

This study concluded that modem technology is suitable for use in an adolescent and young-adult population with type 1 diabetes and would be most appropriate for those who travel long distances to reach the clinic. However, a challenge for healthcare providers will be convincing health insurance companies to pay for implementation of this technology.

In order to improve diabetic care, Roudsani et al [35] evaluated a web-based diabetes management system called DiabNet. The system described an integrated portable glucometer, handheld computer, mobile phone and internet access in the diabetes management system. The salient features of this system are security steps which have been incorporated for the safety of the patient data and advanced graphics for patient data visualization by the diabetic team. DiabNet required the patient to input measurements into handheld computers which were connected to the DiabNet website through their mobile phones. The system has a client-agent-server architecture and tele-consultation features. However, the use of the system may be restricted to the patients who have the computer skills and are highly motivated to look after their health. The drawbacks of this system are:

- A handheld computer is required, hence extra luggage and high cost.
- This system relies on manual data entry and hence is prone to data entry error.
- A phone with an active internet connection is required which introduces extra costs of subscriptions to the telecom network data service.

In examining the security requirement for communication of patient records to remote management systems, Nigrin and Kohane [36] describe a prototype called Glucoweb, which is a system that allows patients with diabetes to transmit self-monitored blood glucose data from their glucometer device directly to their care provider over the internet. The key authentication and security measures to be addressed in the transmission of medical data to remote locations are highlighted in this study.

Farmer et al [37] have described a mobile phone-based telemedicine system for management of type 1 diabetes. This system was implemented on a Motorola T720i phone and a One Touch Ultra® blood glucose meter. The main advantage of the system is that it uses the existing mobile technology for transmission of data and feedback among the members of the diabetic team. However, this system uses a particular brand of mobile phone and glucose meter. It is very unlikely that all patients could have the same brand of the mobile device and glucometer. The other disadvantage of the system is that the patient must have access to the internet in order to transmit the data to the central server.

Rudi et al [39] evaluated a telemedicine system which requires patients to enter the records of their glucose, insulin and food intake using a computer. The system has a rule based expert system based on Prolog to suggest to the patient or clinican the optimal amount of insulin intake required by the patient.

Kim et al [40] presented a system which utilizes SMS for providing the feedback to the patients but glucose measurement data is still transmitted through an internet connection.

Despite of growing interest and research in telemedicine, not many studies have been reported for management of type 1 diabetes of children and young adults. Currently, the NHS in the UK still uses conventional care for diabetes management. Hence, more research is required to design a system which will be acceptable to all stakeholders of a diabetic team.

2.5 Conclusions

In this chapter, an overview and background of diabetes is presented and state-of-the-art technology in the area of telemedicine is discussed and analysed. Diabetes is a chronic disease and the number of people with diabetes is rising at an alarming rate worldwide. The people who suffer from diabetes are not receiving enough support to manage their diabetes due to the current system of management. Type 1 diabetes in particular is difficult to manage as the patient population suffering from the disease are generally children and young adults. One of the problems with the existing system of type 1 diabetes management is the time delay in communication between the patient and healthcare providers. The patients see the diabetic team only after 3 months, which results in poorly managed diabetes. The results of poorly managed diabetes are catastrophic for the patient and incur massive costs to the NHS.

Telemedicine is particularly well suited to manage diabetes as it requires interpretation and predetermined responses to many types of data that can be measured in the home by the patient. Recent advances in ICT can be exploited to develop the systems which are care driven to increase their acceptability among the various stakeholders of disease management. Further, as regulatory bodies become comfortable with legal and liability issues inherent to this approach to home care, then it is expected that the adoptability of telemedicine for monitoring and management of type1 diabetes will increase. Hence

more studies are needed before telemedicine becomes a novel 21st-century tool for the monitoring and management of diabetes. During the analysis it was found that the existing telemedicine has been designed for highly motivated people and most systems need internet access for the patient to send the glucose reading to healthcare providers.

Hence, the author of this thesis proposes a telemedicine system which has been designed and optimized specifically for management of type 1 diabetes in close coordination with Sheffield Children's Hospital diabetic team.

In the next chapter the architecture, design and implementation of the proposed system is explained. The chapter further explains the various levels of testing which are performed for validation and verification of the system.

References:

1. D. R. Hanas, Type 1 Diabetes, second ed.: Class Publishing, Barb House, Barb Mews, London W6 7PA,Uk, 2004.
2. B. Higgins, Type 1 Diabetes in Adults. London: Royal College of Physicians, 2004.
3. E. I. Portuese, L. Kuller, D. Becker, D. Ellis, C. E. Lloyd, and T. J. Orchard, "High mortality from unidentified CVD in IDDM: time to start screening?," Diabetes Research and Clinical Practice, vol. 30, pp. 223-231, 1995.
4. International Diabetes Federation, Diabetes Atlas, Fifth edition: Nov 2012 [online]. Available: http://www.idf.org/diabetesatlas [Accessed: 10 Jan 2013]
5. Yajnik, C. S. "Early life origins of insulin resistance and type 2 diabetes in India and other Asian countries." The Journal of Nutrition 134.1 (2004): 205-210.
6. Hossain, Parvez, Bisher Kawar, and Meguid El Nahas. "Obesity and diabetes in the developing world—a growing challenge." New England Journal of Medicine 356.3 (2007): 213-215.
7. Diabetes UK, Diabetes in the UK 2011/2012: Key statistics on diabetes, Diabetes UK, Nov 2012 [online]. Available: http://www.diabetes.org.uk/documents/reports/diabetes-in-the-uk-2011-12.pdf [Accessed: 10 Jan 2013]
8. G. Panzram, "Mortality and survival in type 2 (non-insulin-dependent) diabetes mellitus," Diabetologia, vol. 30, pp. 123-131, 1987.
9. D. E. Laaksonen, J. Lindström, T. A. Lakka, J. G. Eriksson, L. Niskanen, K. Wikström, S. Aunola, S. Keinänen-Kiukaanniemi, M. Laakso, T. T. Valle, P. Ilanne-Parikka, A. Louheranta, H. Hämäläinen, M. Rastas, V. Salminen, Z. Cepaitis, M. Hakumäki, H. Kaikkonen, P. Härkönen, J. Sundvall, J. Tuomilehto,

and M. Uusitupa, "Physical activity in the prevention of type 2 diabetes," Diabetes, vol. 54, pp. 158-165, January 1, 2005

10. S. M. Setter, J. L. Iltz, J. Thams, and R. K. Campbell, "Metformin hydrochloride in the treatment of type 2 diabetes mellitus: A clinical review with a focus on dual therapy," Clinical Therapeutics, vol. 25, pp. 2991-3026, 2003.

11. "Gestational Diabetes Mellitus," Diabetes Care, vol. 26, pp. s103-s105, January 1, 2003

12. M. F. Greene and C. G. Solomon, "Gestational diabetes mellitus — Time to treat," New England Journal of Medicine, vol. 352, pp. 2544-2546, 2005.

13. S. L. Kjos and T. A. Buchanan, "Gestational diabetes mellitus," New England Journal of Medicine, vol. 341, pp. 1749-1756, 1999.

14. Lawrence JM, Contreras R, Chen W, Sacks DA (May 2008). "Trends in the prevalence of preexisting diabetes and gestational diabetes mellitus among a racially/ethnically diverse population of pregnant women, 1999–2005". Diabetes Care 31 (5): 899–904. doi:10.2337/dc07-2345. PMID 1822303.

15. B. Leese, "The costs of diabetes and its complications," Social Science & Medicine, vol. 35, pp. 1303-1310, 1992.

16. I. Köster, L. von Ferber, P. Ihle, I. Schubert, and H. Hauner, "The cost burden of diabetes mellitus: the evidence from Germany—the CoDiM Study," Diabetologia, vol. 49, pp. 1498-1504, 2006.

17. Department of Health, Turning the corner: improving diabetes care, Department of Health, June 2006 [Online]. Available: http://www.bipsolutions.com/docstore/pdf/13587.pdf [Accessed: 20 Jan. 2013].

18. Jönsson, "Revealing the cost of type II diabetes in Europe," Diabetologia, vol. 45, pp. S5-S12, 2002.

19. WHO. A health telematics policy in support of WHO's Health-For-All strategy for global health development: Report of the WHO group consultation on health telematics, 11–16 December, Geneva, 1997. Geneva, World Health Organization, 1998.

20. Sood SP, et al. Differences in public and private sector adoption of telemedicine: Indian case study for sectoral adoption. Studies in Health Technology and Informatics, 2007, 130:257–268.

21. Craig J, Patterson V. Introduction to the practice of telemedicine. Journal of Telemedicine and Telecare, 2005,11(1):3–9.

22. Hailey, David, Risto Roine, and Arto Ohinmaa. "Systematic review of evidence for the benefits of telemedicine." *Journal of Telemedicine and Telecare* 8.suppl 1 (2002): 1-7.

23. Currell R et al. Telemedicine versus face to face patient care: effects on professional practice and healthcare outcomes. Cochrane Database of Systematic Reviews, 2000, Issue 2. Art. No.: CD002098.

24. Rao B, Lombardi A II. Telemedicine: current status in developed and developing countries. Journal of Drugs in Dermatology, 2009, 8(4):371–375.

25. Whitten, Pamela S., et al. "Systematic review of cost effectiveness studies of telemedicine interventions." BMJ: British Medical Journal 324.7351 (2002): 1434.

26. Klonoff, David C. "Using telemedicine to improve outcomes in diabetes—an emerging technology." Journal of Diabetes Science and Technology (Online) 3.4 (2009): 624.

27. Craig, John, and Victor Patterson. "Introduction to the practice of telemedicine." Journal of Telemedicine and Telecare 11.1 (2005): 3-9.

28. Bellazzi, R., et al. "A telemedicine support for diabetes management: the T-IDDM project." Computer Methods and Programs in Biomedicine 69.2 (2002): 147-161.

29. Rao B, Lombardi A II. Telemedicine: current status in developed and developing countries. Journal of Drugs in Dermatology, 2009, 8(4):371–375.

30. Stanberry B. Legal and ethical aspects of telemedicine. Journal of Telemedicine and Telecare, 2006, 12(4):166–175.

31. Paré, Guy, Mirou Jaana, and Claude Sicotte. "Systematic review of home telemonitoring for chronic diseases: the evidence base." Journal of the American Medical Informatics Association 14.3 (2007): 269-277.

32. Lamb, W. "DIASEND®: one device to read them all and in the clinic guide them." Endocrine Abstracts. Vol. 17, pp. 48, 2008.

33. G. T. McMahon, H. E. Gomes, S. Hickson Hohne, T. M.-J. Hu, B. A. Levine, and P. R. Conlin, "Web-based care management in patients with poorly controlled diabetes," Diabetes Care, vol. 28, pp. 1624-1629, July 1, 2005.

34. H. P. Chase, J. A. Pearson, C. Wightman, M. D. Roberts, A. D. Oderberg, and S. K. Garg, "Modem transmission of glucose values reduces the costs and need for clinic visits," Diabetes Care, vol. 26, pp. 1475-1479, May 1, 2003.

35. N. Zhao, A. Roudsari, and E. Carson, "MED31/437: A web-based diabetes management system: DiabNet," Journal of Medical Internet Research, vol. 1, p. e68, September 19 1999.

36. Nigrin, Daniel J., and Isaac S. Kohane. "Glucoweb: a case study of secure, remote biomonitoring and communication." Proceedings of the AMIA Symposium. American Medical Informatics Association, 2000.

37. A. G. Farmer, Oliver; Hayton, Paul; Bryden, Kathryn1; Dudley, Christina; Neil, Andrew; Tarassenko, Lionel, "A real-time, mobile phone-based telemedicine system to support young adults with type 1 diabetes," Informatics in Primary Care, vol. 13, pp. 171-178, 2005.

38. A. Roudsari, N. Zhao, and E. Carson, "A web-based diabetes management system," Transactions of the Institute of Measurement and Control, vol. 26, pp. 201-222, August 1, 2004 .

39. R. Rudi and B. G. Celler, "Design and implementation of expert-telemedicine system for diabetes management at home," in Biomedical and Pharmaceutical Engineering, 2006. ICBPE 2006. International Conference on, 2006, pp. 595-599.

40. C. Kim, H. Kim, J. Nam, M. Cho, J. Park, E. Kang, C. Ahn, B. Cha, E. Lee, S. Lim, K. Kim, and H. Lee, "Internet diabetic patient management using a short messaging service automatically produced by a knowledge matrix system," Diabetes Care, vol. 30, pp. 2857-2858, November 2007.

3.1 INTRODUCTION

This chapter explains the architecture, design and development of "WithCare+" an innovative closed loop remote blood glucose monitoring system.

The motivation for designing WithCare+ was to exploit recent technological advances in telecommunication and information technology [1,2] for the benefit of patients with chronic diseases in general. However, the scope of this work is to validate WithCare+ in the context of the remote monitoring of type 1 diabetic patients. The system has been evaluated in a field test, in collaboration with the Children's Hospital, Sheffield and has been accepted by healthcare professionals of the hospital for a clinical trial. The clinical trial has started and will last for 18 months.

In order to develop a telehealthcare solution which can improve the patients' quality of life, it is important that engineering teams understand the requirements of both care-providers and care-receivers. Even if an engineering team designs a telehealthcare solution with state-of-the-art technology, it is bound to fail if developed in isolation without taking into consideration the requirements of all stakeholders [3,4]. These ideas were given due consideration while designing the WithCare+ solution. A thorough requirement analysis was performed by taking into account inputs from all stakeholders. The interviews were conducted with different members of the diabetic team i.e. the clinicians, nurses, dieticians and patients.

During the requirement gathering phase of the system, it was necessary for the author of this thesis to attend a couple of diabetic clinics to understand the current system of diabetic monitoring. Currently, the diabetic team at Sheffield Children's Hospital use

"Diasend", a commercial diabetes management system for glucose monitoring of patients [5]. However, there is no patient unit in Diasend and during a routine appointment the clinician uploads the glucose readings from the patient's meter after plugging it to the Diasend hardware. The hardware sends the data to a remote server and the clinician is able to view the records of the patient. However, there is no facility for communication between patient and care provider on a daily basis. Based on this study and the literature review, a new system was envisaged which can address the problems faced by the members of the diabetic team. The analysis of the requirement gathering phase [6-8] established that WithCare+ has to accomplish the following specific objectives in order to be accepted by stakeholders of the diabetic team:

- The WithCare+ system will need two main units: patient unit and server.
- The WithCare+ client (Patient unit) will read the glucose readings from the glucometer via a universal asynchronous receiver/transmitter (UART) and transmit them to the patient's phone in a specific format via Bluetooth.
- The data received on the patient's phone should be sent as an SMS on the GSM (Global System for Mobile Communications) /GPRS (General packet radio service) network to the GSM modem connected to the serial port of the WithCare+ server. The patient should receive a sent message confirmation on their mobile phone.
- The SMS should be analysed at the server and rejected if it does not match an expected format, but any such event should be logged to a log file.
- The SMS with correct format should be unpacked and the glucose readings should be saved into the database.

- A browser based web application should provide the healthcare providers with 24/7 access to the patient data in a textual and graphical format. The user interface should match “Diasend”, a commercial telehealthcare solution which the diabetic team at Sheffield Children’s Hospital are currently operating, in order to cause minimum change to the system usability.
- The user interface should be simple and the clinicians should be able to login to the system using a browser enabled device such as a desktop, laptop, tablet computer or smart phone.
- The user interface should run on commonly used browsers such as Internet Explorer, Google Chrome and Mozilla Firefox.
- With administrative rights the clinician should be able to add and edit the patient details.
- The system should generate alerts in response to any undesired glucose levels such as hypoglycaemia or hyperglycaemia. The alarms generated should be visible to all members of the diabetic team on logging in to the system.
- The system should also generate alerts if the glucose readings are not received by the server for a specific time which should be customizable as per the requirements of the patient.
- The system should have the ability to send feedback to the patients using SMS or email.
- To motivate patients to record their glucose measurement by improving the communication between all stakeholders of the diabetes team, the system should be integrated with social networking sites such as Facebook to improve the communication across the members of the diabetic team.

- The system should be integrated with Skype, which will reduce the cost of communication.
- The system administrator should be able to diagnose any problems with the modem via the user interface.
- A script should automatically run every day at 8 a.m. for system diagnostics and should be sent via email to the system administrator.

If achieved, these objectives can encourage diabetic patients to monitor their blood glucose levels on a daily basis, which could in turn improve their HbA1C [9] and eventually reduce later-life complications. The system will be a friendly tool for the diabetic team to ease the pressure on their time and improve communication between all stakeholders of the team. Over time, these improvements could also reduce the cost of care for the National Health Service (NHS). However, the evaluation of the system for improving the HbA1C and reducing costs is beyond the scope of this study. The study about these benefits will be published by Dr. Neil Wright and his team (from Sheffield Children's Hospital) after running the full trial.

3.2 Architecture of WithCare+

Software architecture refers to high level abstraction of a software system that consists of high-level components and connectors, as opposed to implementation details. It gives an overview of how the system will behave and serves as the blueprint for low level design to produce a robust system. Performance, modifiability, and security of the system cannot be achieved without a unifying architectural vision. Design risks can be identified and mitigated during the initial design phase by an effective architecture [10,11].

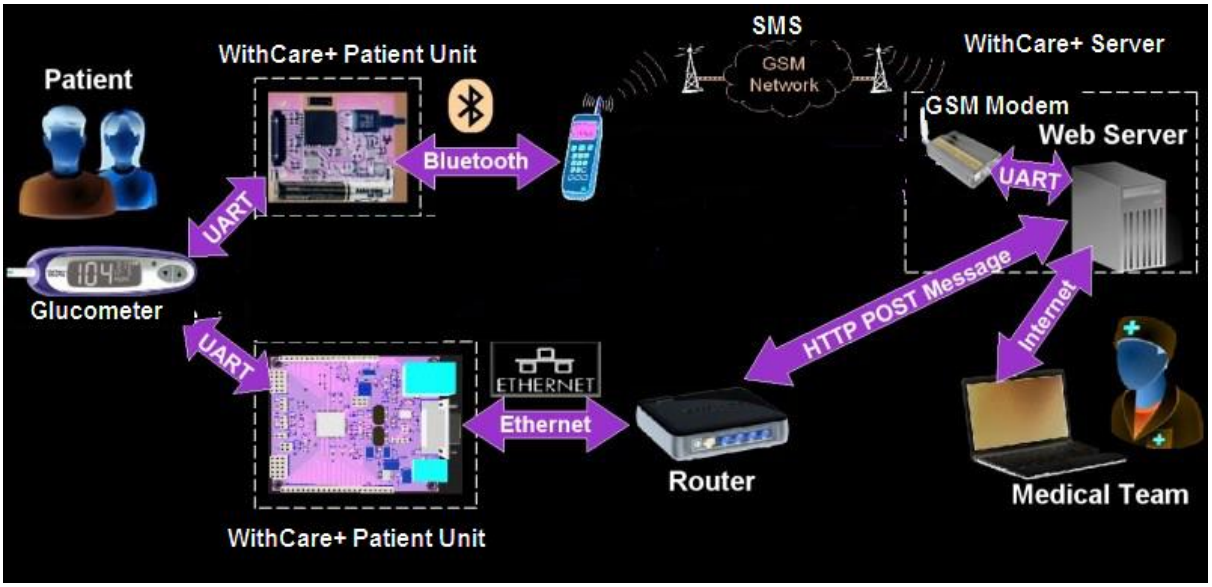


Figure 3.1 Architecture of WithCare+

The proposed architecture of WithCare+ is illustrated in figure 3.1. As shown in the figure, the two major building blocks of the WithCare+ system are the patient unit and the server. The main aim of the patient unit is to transfer glucose readings from the glucometer to the server. However, the method of transferring the data to the server is flexible in this architecture, which is an innovative approach. In this work, two approaches of the data transfer have been tested.

In the first approach, the patient unit is a wireless module, capable of reading glucose measurements from the glucometer via UART. The glucose readings can be transferred by the press of a button from the wireless module to a patient's mobile phone via Bluetooth and further sent as an SMS via the GSM [12]. The SMS is received by the modem connected to the serial port of the WithCare+ server. The incoming SMSs are received on the SIM card of the phone and are read by the daemon process, which runs in the background. The daemon process continuously monitors the GSM modem for the incoming SMS and after the analysis of the SMS, it is logged into the database for

further processing. In future, this approach will be referred to as an SMS based-approach.

In the second approach, the patient unit uses an Ethernet connection [13] and sends the data to the webserver as a Hypertext transfer protocol (HTTP) message on a TCP/IP protocol [14-16], which after analysis is stored in the database for further processing. In future, this approach will be referred to as an ethernet-based approach.

As is evident from figure 1, the WithCare+ server architecture is innovative in the sense that it can receive the data more than one way depending on the preference and circumstances of the patient. Both approaches have their pros and cons.

The main advantage of an SMS approach is that the patient does not need an internet connection. The initial design was that the glucose data can be transferred from the patient unit to the mobile phone of the patient by exploiting the Bluetooth [17] serial port profile (SPP) of a patient's phone. This approach worked with the low-end mobile phones which allowed communication over Bluetooth SPP. However, this design fails with smart phones, as Bluetooth SPP is not activated in these cell phones by the manufacturers, which was discovered after many hours of debugging time. This challenge was overcome by developing an application which established communication over the RFCOMM [18] protocol instead of SPP and successfully sent data as an SMS. This finding could save other researchers a lot of their debugging time. RFCOMM is a connection-oriented, streaming transport over Bluetooth and emulates serial ports. The section 3.3 will explain how in detail how this communication was achieved.

An SMS approach could be very useful for developing countries, where the internet is still not available in rural areas but mobile phones are. In these areas, there is often no specialized diabetic healthcare centre and patients have to travel long distances to attend appointments in the main hospitals. This results in patients spending a full day travelling and in certain cases, patients may need to stay overnight to receive the advice from clinicians. In this case the patients also need to arrange for the overnight accommodation. An SMS-based approach can be very helpful for the patients in these areas as the glucose readings can be sent by the press of a button and devices can use the device with minimal training. The healthcare providers can advise the patients by an SMS. However, at this stage the plan is to perform the pilot trial of the system for the type 1 diabetic population of Sheffield and in future WithCare+ could be evaluated in similar pilot trials in rural areas in different countries.

The ethernet approach can be useful only if an internet connection is available to the patient. However, this is not always the case, which is a drawback. This approach has been tested and verified to prove that the WithCare+ server can receive the data either as an SMS or as an HTTP message. The ethernet-based approach will not be used in this pilot trial as the diabetic team at Sheffield Children's Hospital prefer the SMS based-approach due to its advantage that patients need neither a computer nor an internet connection. The work related to the ethernet-based approach has been presented at IEEE conference on embedded software and systems [19].

In order to handle messages from the SMS based approach, a GSM modem is connected to the serial port of the WithCare+ server machine. Text messages from the patient unit are received on the subscriber identity module (SIM) card of the modem and a daemon process continuously monitors the SIM card for the incoming SMS messages by

sending Hayes modem commands, known as AT commands, to the GSM modem. A list of AT commands is included in Appendix A. As soon as a new message is received, it is read by the process and analysed further. The daemon process either logs the message to the WithCare+ database if it matches a specific format or rejects the message if it does not match the format to ensure security. One more daemon thread reads the received SMS from the database and extracts the patient glucose readings with a timestamp. Then this extracted data is stored into the readings table of the database.

The web application of WithCare+ is designed and developed using an n-tier web client architecture based on the model-view-controller design pattern. The Web-based user interface helps the health-care provider with data analysis, data visualization and decision support. It allows them to send feedback and therapeutic advice to patients from anywhere using the browser enabled device. The feedback can be sent via an SMS or email depending on the preference of the patient or caregiver and on the circumstances. WithCare+ is integrated with Skype and to other social networking sites in order to improve communication between stakeholders of the diabetic team, reduce costs and time of the communication. The idea to integrate the social networking sites was implemented in order to engage young adults. However, no clinical related data will be available on social networking sites in order to ensure the privacy of the patients.

The end-to-end system provides information to the health-care provider, helping them make an informed decision for the management of the patient's disease. All the relevant information about the patient including subjective data such as physiological and sociological data can be stored in the database. Health-care providers can view the data in tabular form or in the form of charts. It will alert the health-care providers when the

system detects the out-of-bounds glucose levels. The system can perform statistical analysis of the data, which can be useful in finding any correlation in the data. The system can provide a 24-hour open channel of communication between the various stakeholders of the diabetic team and can reduce some of their burden by letting them focus on problematic cases where more attention is required and the system can take care of routine communication.

The proposed low-cost system, WithCare+, is expected to improve communication among the members of the diabetic team. This approach of using the existing communication channels and social networking sites for improvement of communication between patients and clinicians could lead more people to follow this strategy and could be useful to improve the clinical outcome. The work related to this approach has been presented at the Engineering in Medicine and Biology Society (EMBS), IEEE Conference [20].

3.2.1 Description of the Patient Unit

The patient unit (electronic module) is connected to the glucometer and is operated by the patient. The module is physically small, requires minimal energy with lightweight packaging. The operation is intended to be almost transparent to the user with the minimum of configuration.

There are four parts to the module, a Bluetooth IC (including host processor supporting a virtual machine for client applications), power supply, programmable serial interface and user interface. The user interface is a single button (to address issues of consent) and a tri-coloured LED. A switched mode power supply was designed around a low-power boost converter. The power supply unit at rest is disabled. The module is woken by pressing the button which enables the boost converter, and the Bluetooth IC then

asserts a hold-power signal which maintains the enable to the regulator until the module wishes to power off. The application software is written in embedded-C using the virtual machine (VM) and divided into a number of separate interacting sub-systems performing various required tasks. Message passing is used between the different threads. The operation of the applications functions are summarized in figure 3.2 and the module is shown in figure 3.3. A key requirement of the technology is minimal user interaction and therefore the front end was designed (based on firmware) to be compatible with most glucometers and most mobile phones including Android Smartphones. The device has been successfully tested with a range of mobile phones and glucometers (including the 5 most popular meters used by Sheffield Children's Hospital patients). The device has received the required UK regulatory (MHRA) clearances to be used by the NHS patients in the clinical trials. The work related to development and testing of the patient unit was done by another Ph.D. student namely Anastasios Kanakis.

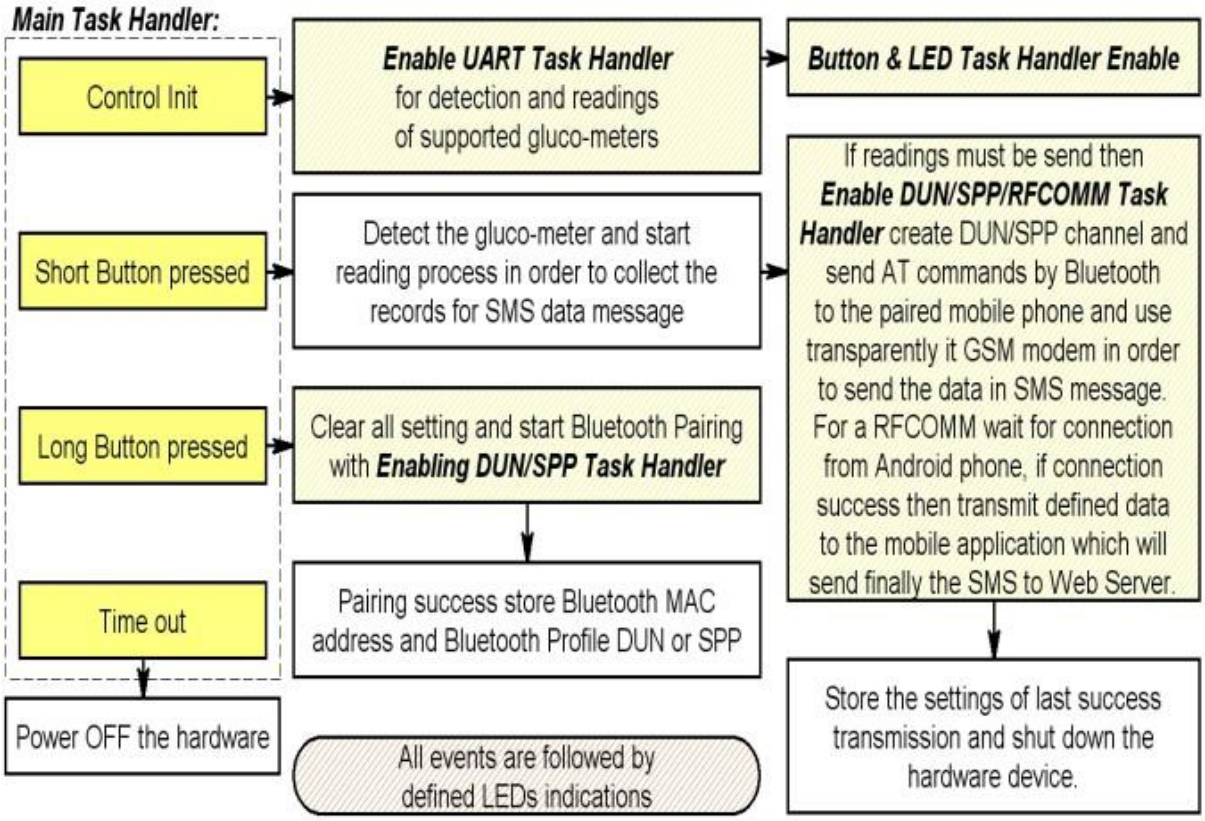


Figure 3.2 Summarized functions for Bluetooth applications

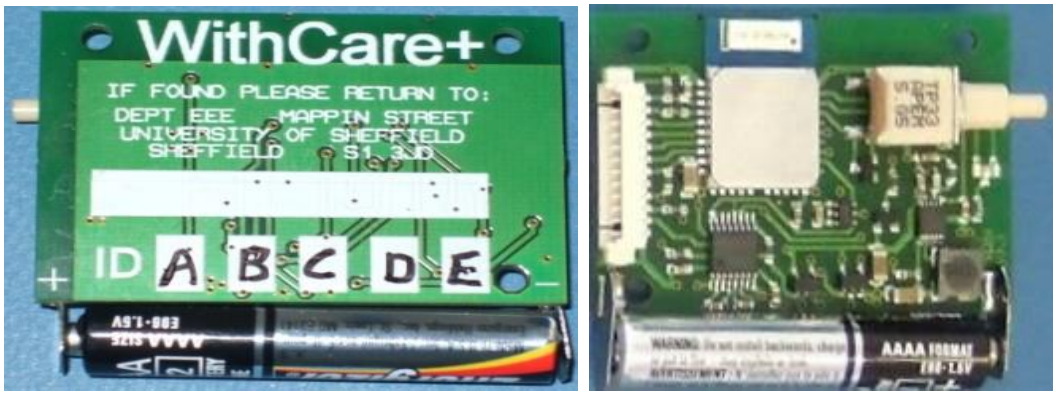


Figure 3.3 WithCare+ Patient Unit

3.3 Design of Bluetooth Communication using RFCOMM Protocol

The aim of this work was to design and develop a proof of concept for establishing communication between an Android handset and a Bluetooth enabled device using Android's Bluetooth application programming interface (API). Although the developed

software can run on any Android platform supporting the Bluetooth network stack and any Bluetooth enabled device, this concept has been tested on Android version 2.2.1 (HTC wildfire mobile) and PSoC development Kit from Cypress Semiconductor Corporation [21].

The APIs provided by the Android platform allows an application connecting to a Bluetooth enabled remote device. An Android application can use the Bluetooth APIs to perform the following actions:

- Scan for other Bluetooth devices
- Query the local Bluetooth adapter for paired Bluetooth devices
- Establish RFCOMM channels
- Connect to other devices through service discovery
- Transfer data to and from other devices
- Manage multiple connections

The main steps involved for establishing communication and software module developed for the purpose are explained in the appendix D.

3.4 Design of WithCare+

In order to allow flexibility and accommodate changing requirements, the planning and design of the remote monitoring system was done using a combination of agile processes [22,23] and structured iterative development.

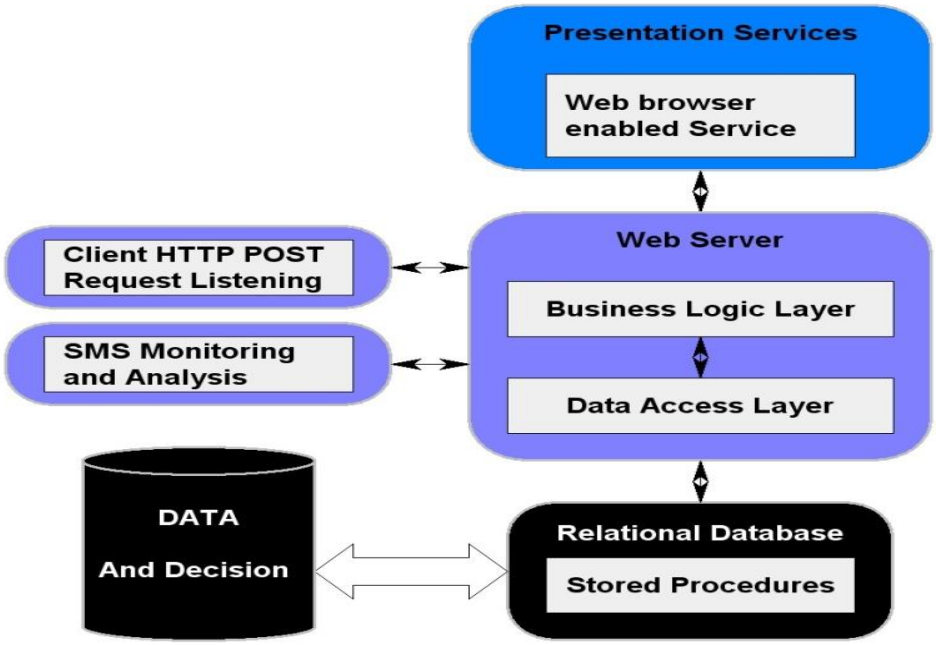


Figure 3.4 High Level Design of the WithCare+ Server

The high level design of WithCare+ is illustrated in figure 3.4. The application layers of the system are shown and the interconnecting arrows indicate the data flow. The system is divided into the following components for development.

3.5 Components of the WithCare+ Server

The main components of WithCare+ server are:

1. Incoming SMS monitoring and analysis software module
2. Web based interface for user interaction and maintenance
3. Automatic scheduler

3.5.1 SMS Monitoring and Analysis Software Module

The patient unit module of WithCare+ reads glucose levels from the glucometer of the patient, packages them in a specific format as shown in Appendix B, and transmits data via Bluetooth to the patient's phone. This data is then sent as an SMS to the Server, which is received by the modem connected to the Server.

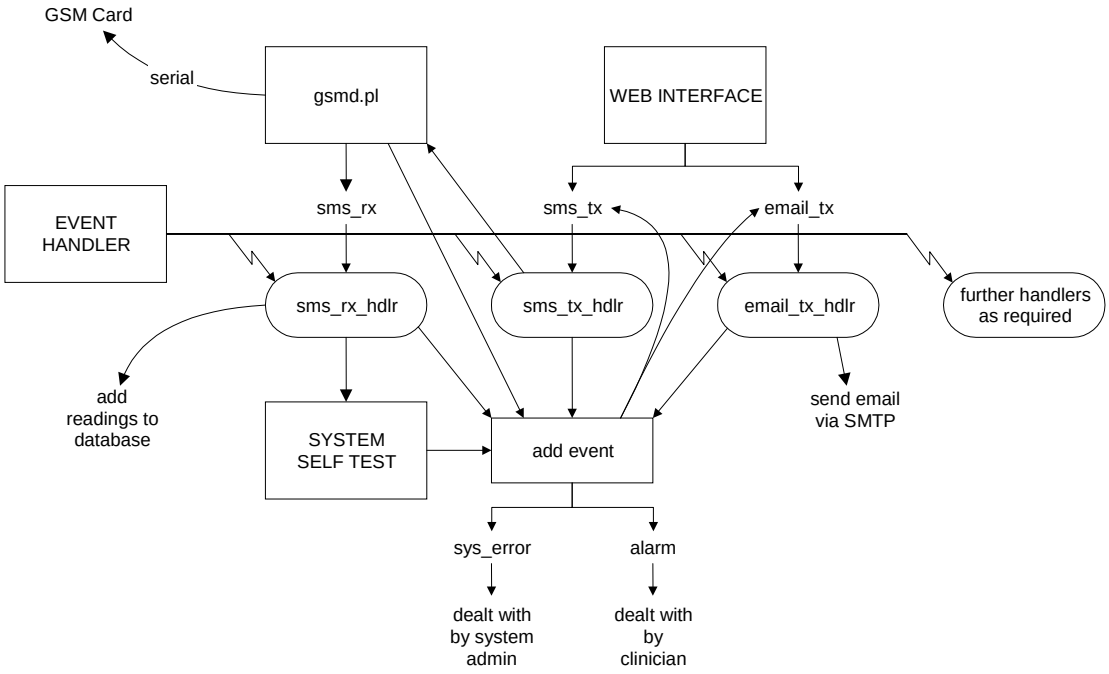


Figure 3.5 Event Interactions of SMS Monitoring and Analysis

Figure 3.5 shows the interaction of all the events in the SMS monitoring and analysis module. The two main components of the module are SMSHandler and EventHandler. These two server side components have been developed in Perl [24] and run as daemon processes on the Ubuntu Linux Server [25]. SMSHandler continuously polls the GSM modem for incoming messages by AT (Attention) commands [26]. This study uses a Wavecom modem but the commands are standard and applicable to all GSM modems.

When a new SMS message arrives on the modem, SMSHandler reads the message and analyses its format with the expected format as shown in Appendix B. The message is inserted into the database events table and is also logged to a log file. If it does not match the format, the event is logged to the log file only and not stored in the database. The flow chart of SMSHandler is depicted in figure 3.6.

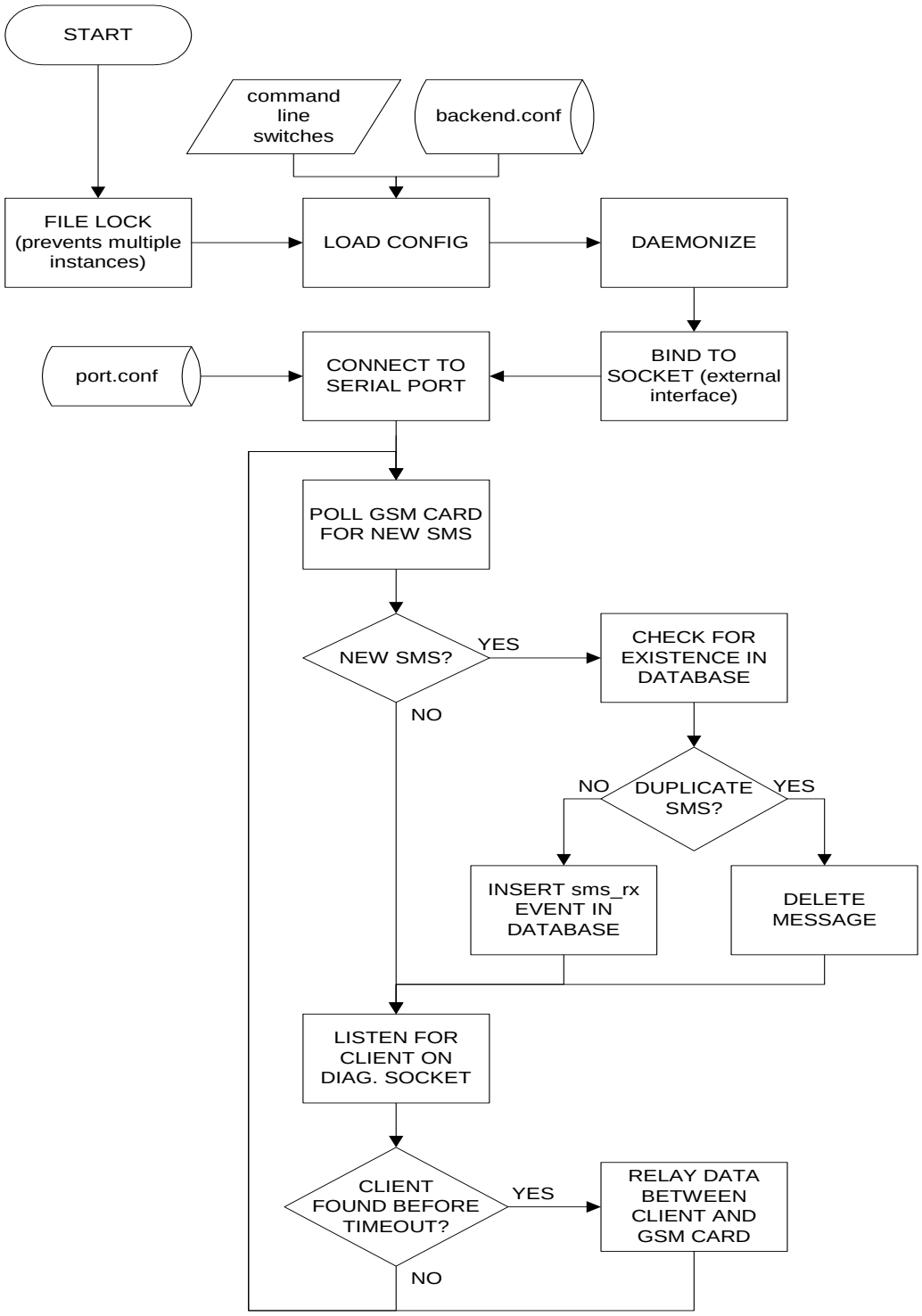


Figure 3.6 Flow Chart of SMSHandler

Another important software module is EventHandler which acts as the central nervous system of the WithCare+ server software, reading the messages from the events table and analysing these messages based on the designed rules. Then, the glucose readings of the patient are extracted and then stored in the readings table of the database. If there are any alarming conditions, such as low or high readings, the event handler generates an alarm in the event table. EventHandler also generates the alarm if the patient does not send the readings within the prescribed time, and it alerts the clinician. In addition, it keeps track of the health of the system using the automatic diagnostics software module, and raises an alert for the system administrator if any problem is detected with the modem. The system administrator can easily configure from the user interface how often the diagnostics should be run. This feature detects the problems early on and thus minimizes the chances of a system failure.

3.5.2 WithCare+ Web Based User Interface

The server software is designed using a model-view-controller design pattern, which is popular for designing robust systems [27, 28]. The main components of a web-based interface are:

1. Data Access Tier
2. Business Logic Tier
3. User Interface

3.5.2.1 Data Access Tier

The data access tier is the code that executes select, insert and update queries directly on the database. This code has direct access to the database, and knows all the database details, i.e., the schema of tables, the name of fields, and the stored procedures. The data provider design pattern was used in the development of the data access tier to provide

the flexibility of using any other relational database management systems (RDBMS) that may have to be supported in future.

The GlucoseReadingProviderBase and PatientProviderBase abstract classes implement the data provider pattern. These classes define Create, Read, Update and Delete (CRUD) methods and furthermore, include the implementation of helper methods for extracting data returned from the underlying database in an IDataReader object. The actual data access code is inside the secondary classes i.e. SqlGlucoseReadingProvider and SqlPatientProvider which inherit from the base class, and provide a concrete implementation for base class abstract methods. These are called providers, and are usually specific for one type of data store which in this implementation is MySQL [29]. The return types of provider class methods are entity objects or generic lists of entity objects. The provider classes also include a private helper method that extracts data from the data reader and creates the entity object for that table which is then returned to the business logic tier of the application. The methods of the provider classes follow a similar design pattern.

At runtime, the DataProvider static class factory reads the configuration setting to identify the concrete data provider type to instantiate it. The new provider object is created, cast as the base provider type and returned to the caller. Data from the database is passed to the business logic layer by the entity classes PatientDetails, GlucoseReadingDetails and FeedbackDetails. An entity object has a default constructor and also a parameterized constructor that assigns parameter values to the properties of the class. All the entity objects follow a similar design pattern. Figure 3.7 shows the entity-relationship diagram of the WithCare+ database [30, 31].

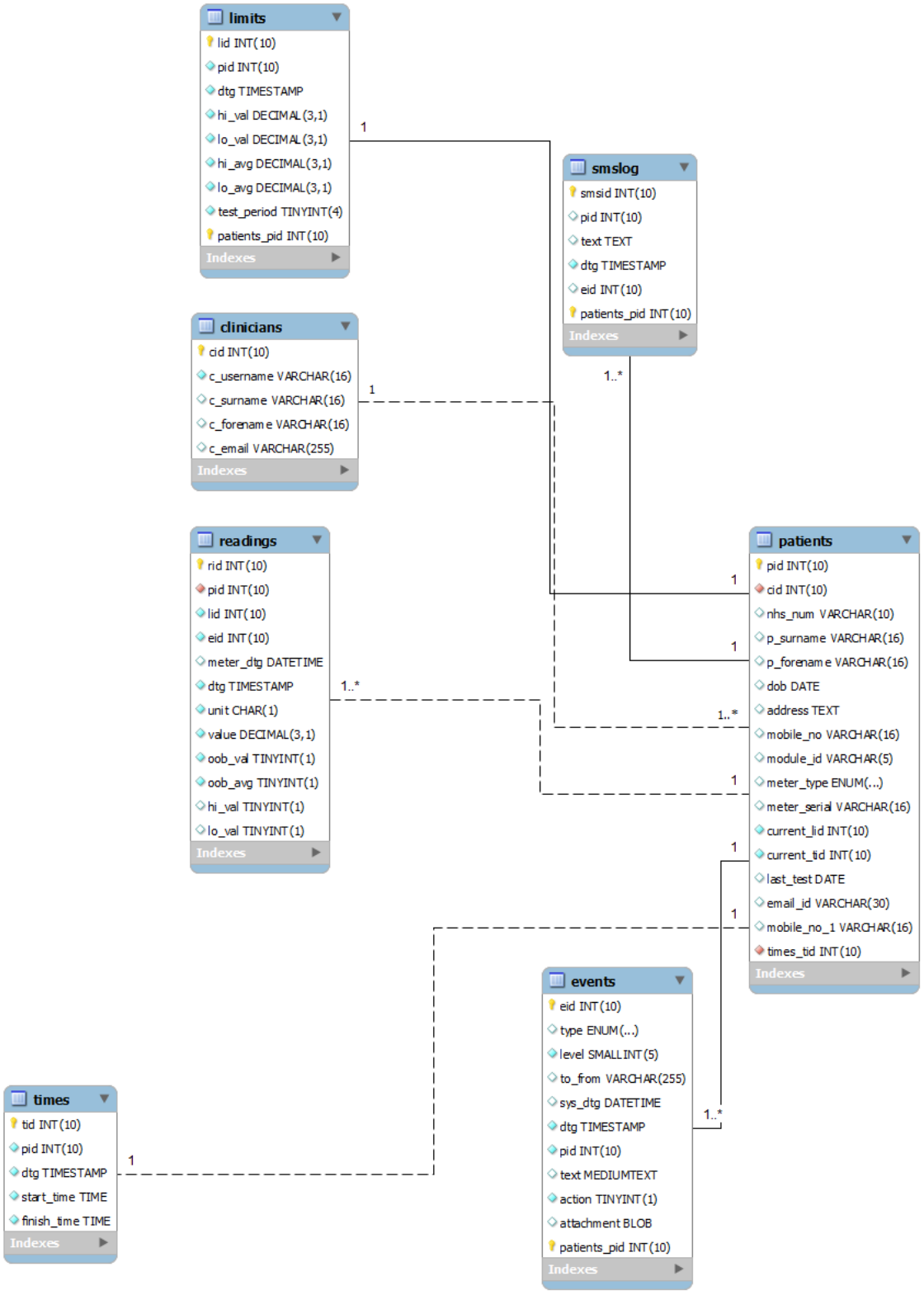


Figure 3.7 Entity relationship diagram

The entity-relationship diagram for the WithCare+ database was developed using MySQL Workbench 5.2. This diagram shows pictorial representation of physical tables in the database and relationship between the tables. In this database, the tables have one-to-one or one-to-many relationship with each other. The patients table is the most important table of the database as it contains the subjective information of the patients. It has one-to-many relationship with the events table, which means more than one events will be associated with a patient. The events table stores messages received at the GSM modem for further analysis. The EventHandler reads the messages from the events table and stores the glucose reading values in the readings table after extracting them from the SMS. There is one-to-many relationship between the patients table and the readings table as one patient is associated to one or more readings. The clinician table has one-to-many relationship with the patient table as one or more patients are associated with one clinician. The smslog table stores the manual messages sent by the patients in response to the feedback from the clinicians and has many to one relationship to the patients table as many SMSs are associated to one patient. The limits table stores the high, low and average values of a patient and has one-to-one relationship with the patient table as each patient has only one or high value. The times table stores the times between which the patients have given their consent for receiving the feedback. If the clinician tries to send the feedback outside of these hours, the system waits till the preferred time of the patient and then sends the feedback. The times table has one-to-one relationship with the patients table as every patient has only one preferred timings.

3.5.2.2 Business Logic Tier

Business logic tier is the middle tier of the application and contains the business rules, validation checks and functions for data manipulation. The code in the business logic layer takes the data retrieved by the data access layer and exposes it to the presentation layer client in a more abstracted and intuitive way, hiding low-level details such as the database schema and adding validation logic that ensures that the input is valid. The class diagram of the business logic layer is shown in figure 3.8.

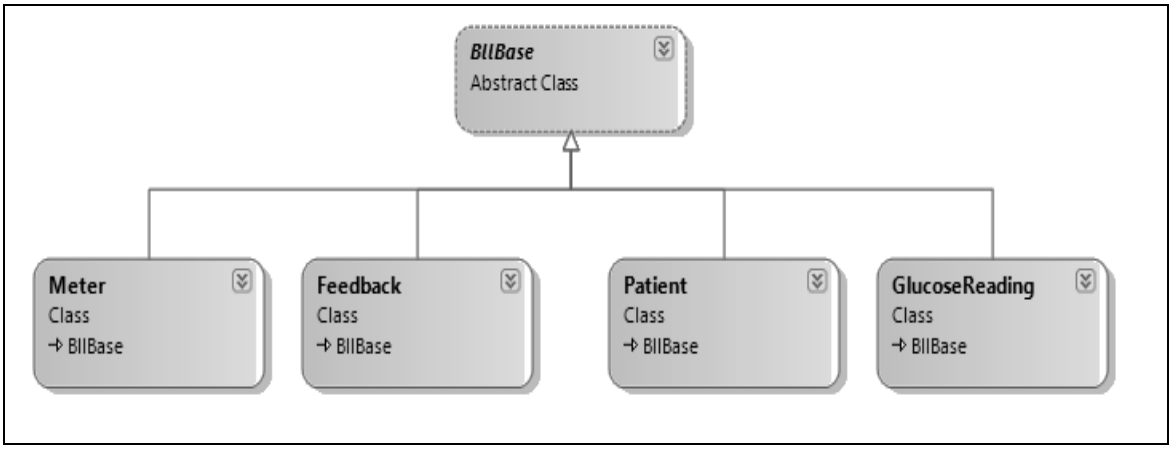


Figure 3.8 Business logic Tier classes

The business logic layer has four business classes which inherit from the BllBase class. Each class has shared methods for common functionality and instance methods to provide functionality unique to each object. Each property of the class is mapped to a column in the database and the GetReading helper method converts an entity object from the data access tier to the corresponding business object. Other methods in the class perform select, insert, update and delete operations on the database using the appropriate methods of the data access tier.

3.5.2.3 WithCare+ Web Interface

The user interface is very important in any application because it is the only part of a system that the user interacts with. While designing the presentation tier the key

consideration was the requirements of the diabetic team, chief of which was simplicity. The WithCare+ web application was hosted on a Linux operating system running an Apache webserver. Screenshot of the user interface (UI) pages created in the web application are shown in the following diagrams.

The web pages were developed using Hypertext Preprocessor (PHP) and PERL [32,33] and the layout of the pages was designed using a combination of XHTML mark-up, Cascading Style Sheet (CSS) and JavaScript. Only those features of JavaScript were used which are browser independent to make the application cross browser compatibility. The browsers used for testing are Internet Explorer, Mozilla Firefox and Google Chrome. The important pages of the WithCare+ UI are listed below:

1. Patient Registration and modification page
2. Inbox
3. Event Registration
4. Clinician Registration
5. Event Logger Page
6. Readings in tabular form and graphical form
7. Feedback
8. Admin for diagnostic purposes

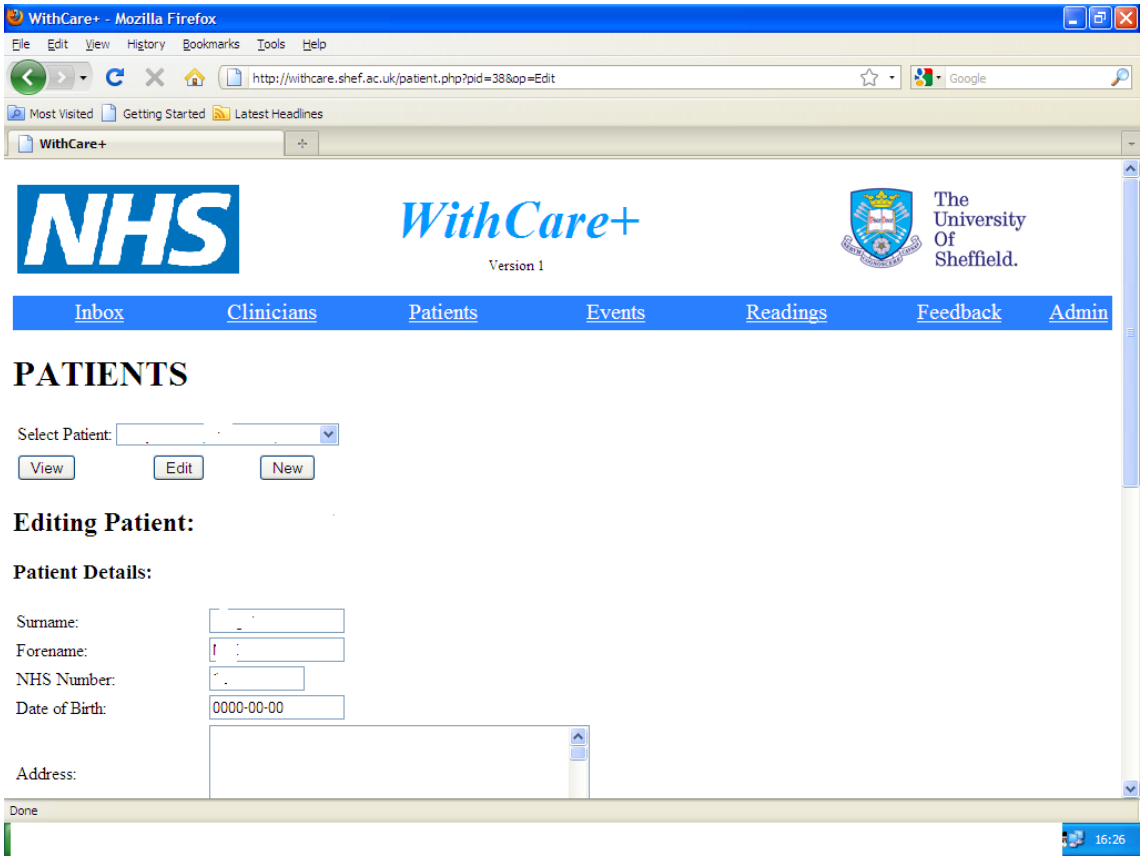


Figure 3.9 Patient Registration Page

Figure 3.9 shows the 'patient registration page' which allows a clinician to register a new patient, modify or view the details of an existing patient. The clinician can select this patient from the drop down box and the details of the selected patient are displayed. The rendering of the page was programmed to be context sensitive and in one mode it allowed existing patient records to be edited while the second mode enables clinicians to add a new patient record to the system. To ensure accurate data entry, existing fields are populated dynamically based on the existing records in the database. The page layout was improved by grouping the fields into sections so that the data input becomes more intuitive. Descriptive headers made it easier to distinguish the sections, for example the address section from the glucose limits section.

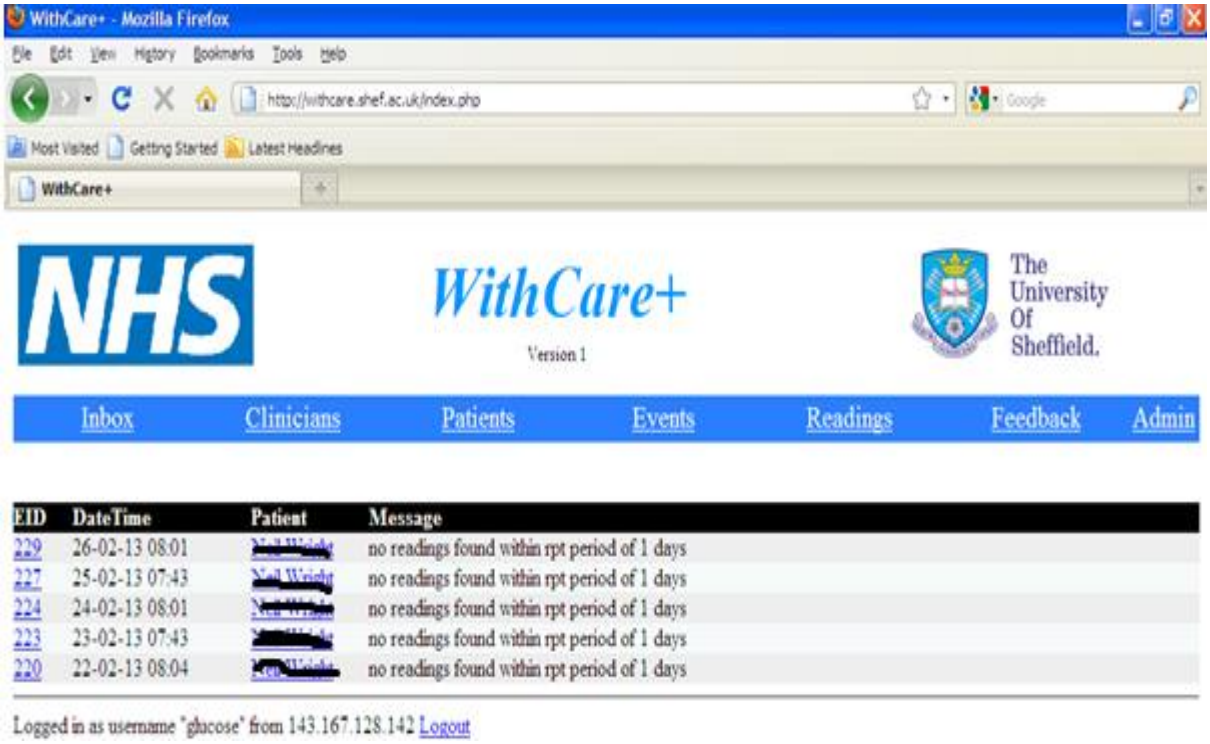


Figure 3.10 Inbox page gives an overview of Events

Figure 3.10 shows the 'inbox page' of the WithCare+ user interface, which displays alerts generated by the system 3 cases. Firstly, glucose readings outside the set normal range are displayed in the inbox for the immediate attention of the diabetic team. Secondly, if a patient does not send the glucose readings for more than the specified period, an alert is generated in the inbox for the attention of the diabetic team so that the patient will be contacted immediately. Finally, the alerts are also generated by the self-diagnostic software, which runs automatically every morning to check the health of the modem. In all these cases, the general format of the alert message is an event id (EID), date along with the time of the event, patient name and message about the event which caused the alert. EID acts as a hyperlink to the page where the action can be confirmed by the clinician as shown in figure 3.11. Similarly, the patient's name acts as the hyperlink to the 'feedback page' as shown in figure 3.12. A member of the diabetic team can contact the patient from this page by choosing one of the options for sending

the feedback i.e. text message, Email, Skype or Facebook. The feedback is sent using the patients preferred method.

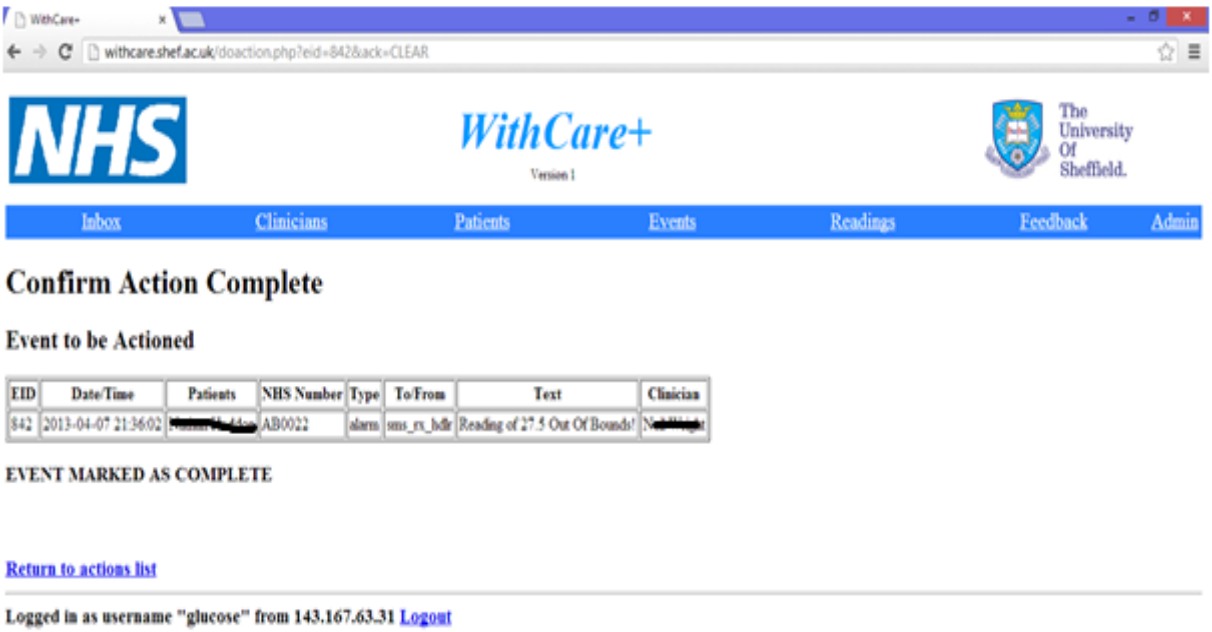


Figure 3.11 Inbox page gives an overview of Events

Figure 3.12 Feedback to Patient

The inbox is one of the most important pages the clinician monitors on a daily basis for any alerts. In the case of an alert, the clinician sends the feedback to the patients and clears the alert from the ‘action-confirmed page’, which automatically deletes it from the database and also clears from the inbox. The other important page for the clinician is the ‘readings page’ as shown in figure 3.12. The clinician can select a patient from the drop down menu which shows the patient name and their NHS number to uniquely identify them. Also, healthcare professionals have options for selecting the time period by a radio button as shown in figure 3.13. Finally, the clinician can click the ‘view As’ button for viewing the data as a table, Day-Graph, Week-Graph or Time-Graph.

Figure 3.13 Selection of viewing Reading

If the clinician clicks on the table button, the glucose readings will be displayed in tabular form as shown in figure 3.14. The value of the glucose reading has been colour coded. Green indicates a normal value, red signals a high value and blue implies a low reading. The weekends in the table are coloured in red. This helps the diabetic team to

look at the problem cases quicky and send immediate feedback, as cases of hyperglycemia are more common during weekends.

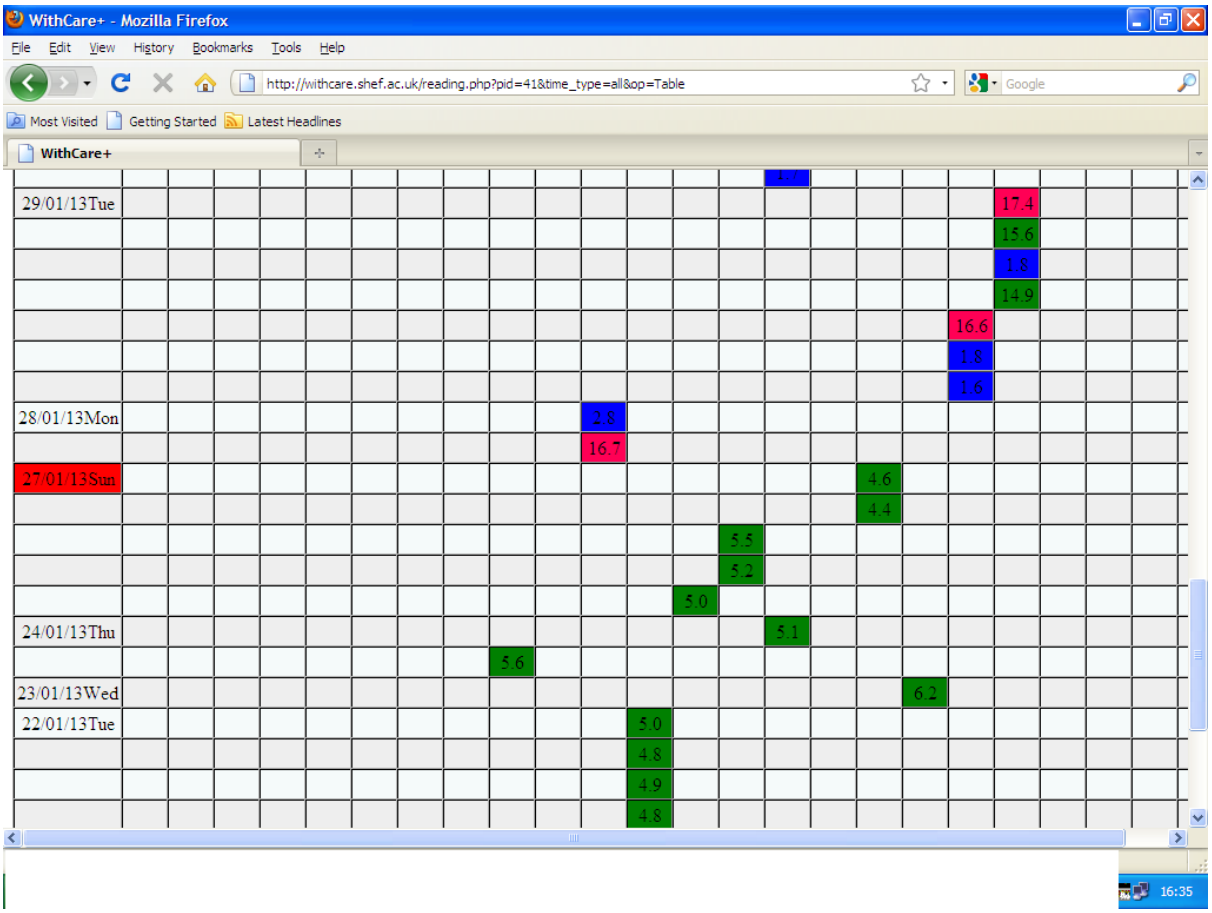
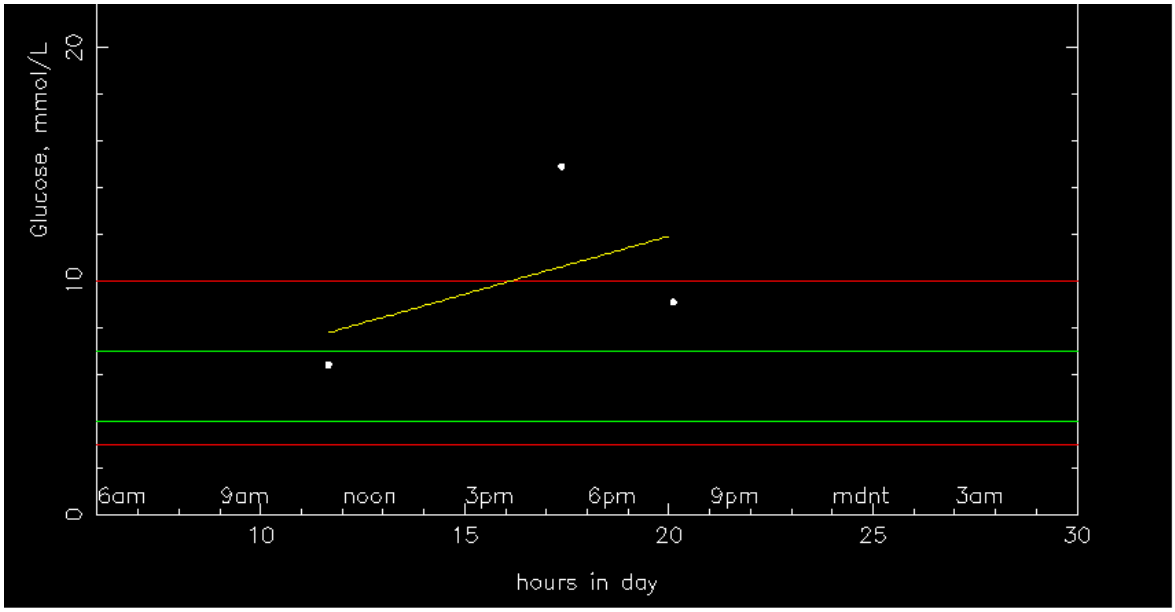


Figure 3.14 Readings in tabular form

Figure 3.15 is a scatter plot of a typical day during the pilot trial. The x-axis of the graph shows the hours of day and the y-axis shows the glucose readings in mmol/dL. The range of the glucose level between the green lines (middle lines) is the optimal or desired level for a patient and the glucose reading value outside the red lines (top and bottom lines) indicates an abnormal BGL. The values above the top (red) line indicate a high level of glucose whereas the value below the bottom (red)line signals a low level of glucose.



Statistics min = 6.40, max = 14.90, median = 9.10, mean = 10.13

Figure 3.15 Day graph of a patient with descriptive statistics

Descriptive statistics such as minimum (min), maximum (max), median and mean values are calculated, which are helpful for the diabetic team when they choose to look at the data over a long duration e.g. three months which is generally the period after which HbA1C is repeated. The graph also shows the trend line or regression line of the glucose levels.

The layout of the pages shown above was designed to improve usability and productivity. System testing and validation, as explained in the next section, were performed to ensure the web interface is functional and operates in line with the requirements.

3.5.3 Automatic Scheduler

This software has been developed to run certain jobs automatically, which is very helpful for system diagnostics. The jobs can be scheduled to run every minute or every day at a specific time or every week. Currently, the modem self-diagnostic test is run

every morning to check the health of the modem and the report is automatically sent to the system administrator. This utility software is also used to back up the log files.

3.6 Testing and Validation

In order to test the system, a helper program was written to generate test data to populate the tables in the database. With the database loaded with test data the application was run and extensive manual testing was performed.

Application testing was performed using advanced testing and debugging tools. The main test procedures carried out during the testing process were:

- Integration tests to ensure that that the assembled system operated according to the design.
- Functionality tests out to ensure that the system functions according to the specification.
- Regression tests after any changes in the software, to ensure that existing functionality was not affected by new changes.
- Performance tests to ensure that the codes were executed optimally with minimum overheads.

MySQL Workbench 5.2 was used to analyse the stored procedures and the WithCare+ database was tuned and optimised to ensure a high performance. Subsequent performance tests were carried out which showed a significant improvement in query execution times. Finally, verification was done to ensure that the final system met all the design objectives and operation was in line with the functional specification. The tests performed focussed on functionality and security in the deployment environment.

3.6.1 Usability Testing

The usability testing [34,35] of WithCare+ was performed in two stages. Five people volunteered to use the patient unit and performed the glucose tests with a calibration solution. One of the volunteers acted as a doctor who registered the other volunteers as patients. The volunteers used their personal phones for transmitting the data to the server and performed the glucose tests 3 times a day. The volunteers were able to use the system with minimal training. This testing was undertaken by the volunteers for one week and the feedback received was positive.

In the second stage of usability testing, the patient units were handed over to the diabetic team at the Children's Hospital, Sheffield. The diabetic team at the hospital performed the usability testing for two weeks. During this phase of testing, some members of the diabetic team acted as patients. The phones used by the members of the diabetic team were Blackberry and Android phones as most of the patient population have these two types of phones. After this testing period, it was agreed by the project management to start the clinical trial of the system in a phased manner, which will be explained in the next section.

3.7 The Clinical Trial

In the first phase of a pilot trial, 6 patients were registered by the Children's Hospital, Sheffield. All the patients have smart phones with 4 patients having Android phones and 2 patients having Blackberry phones. All of these patients were given training to use the WithCare+ patient unit. The two steps which are involved in setting up the WithCare+ device is the pairing between phone and the device for establishing Bluetooth communication. This step has to be performed only once by the clinician at

the time of the patient registration. The training manual for this step can be found in Appendix C.

The next step to be performed by the patient is to send the glucose readings to the server. This step is very simple, patients will perform the glucose readings as usual, but they need to connect the WithCare+ device to their glucometer and ensure their mobile phone is within the Bluetooth range to establish communication with the WithCare+ device. The glucose readings can be sent to the server on the press of a button. The training manual can be found in Appendix C.

3.7.1 Problems Faced During the Pilot Trial

One of the six patients who owned a Blackberry had many applications running on it and as a result, application although installed successfully did not function properly. One of the installed applications was blocking the Bluetooth communication between the phone and the device. This problem was resolved by resetting the phone (with consent of the patient) to the factory setting thereby deleting all the existing applications and reinstalling the applications. This work around seems to work well as the last resort in the problematic cases. However, this approach will not work if the patient does not agree to reset his phone to factory settings. This has been classed as known issue for the application.

One more patient who attended the training and was able to send data initially did not continue from the next day. Even though the diabetic team tried to contact her, there was no response. This is the typical case of non-compliance from the patient and will be quite common in practical situation. One of the reasons for integrating social media with WithCare+ was to engage these types of patients. However, the NHS at the

moment does not allow running Facebook or Skype as a policy matter. An application has been filed by the diabetic team to their information technology department to seek permission for running Facebook and Skype within the context of WithCare+. Once this permission is granted it may be possible to understand the impact of social media sites to engage the unmotivated patients.

One other patient stopped sending readings after one week. After analysing it was found that patient did not know that his phone had run out of credit. It was suggested that £10 voucher may be provided to overcome this kind of problem for the patient population without a contract phone. The contract phones generally provide unlimited free SMS.

The other three patients who started well, one patient sent readings after each measurement. The other two preferred to sent all three measurement at the end of the day. After two weeks, however, no readings were being received by the Server. On investigation, it was found that the resistor of the cable, which connects glucometer to the WithCare+ patient unit, was not able to withstand the rough use of the patients. To solve this problem, this resistor was replaced by a more robust surface mount device (SMD) type resistor in all the patient units. The pilot will run for 18 months and clinical outcomes, cost analysis will be published by the diabetic team.

3.8 Conclusions

The aim of this research was to design and develop a 24/7 telemedicine system for remote monitoring of type 1 diabetic patients, which will improve the communication between the patients and members of the diabetic team and give to healthcare professionals the ability to look at the glucose readings on a daily basis using any browser-enabled system. A number of challenges were faced during the design,

development and testing of the system in order to ensure acceptability of the system by all stakeholders.

The first major design challenge was to ensure possibility of widespread deployment in areas where internet access is limited and this was overcome by relying on the SMS messaging through components developed to support SMS communication. There is no telemedicine system, to our knowledge, which does not need patient internet connection for sending the glucose readings, which is a bottleneck for most of the systems. One distinct advantage of WithCare+ patient unit is that is does not need an internet connection as it operates on the GSM network.

The second major design challenge was that of scalability, performance and security, which is addressed by adopting model-view-controller (MVC) architecture. The third design challenge was to provide a web standard compliant, user friendly browser based interface to the application for use by healthcare providers in managing their patients and this was achieved by designing an aesthetically pleasing layout without using any web application framework to keep the interface simple and easily manageable. The fourth design challenge was to develop a system which will be low cost and this objective was fulfilled by using freely available open-source software.

Withcare+ has been accepted by Sheffield Children's Hospital for a clinical trial, which will run for 18 months. The system developed is recommended for all deployment scenarios and will be especially useful in rural areas of developing countries where mobile phone usage is common but internet access is limited.

Remote monitoring with continuous feedback from the diabetic team could motivate patients to take glucose measurements 3 to 4 times a day as needed. However, this

‘invasive’ procedure for glucose level monitoring on a daily basis for many years is clearly inconvenient and contributes to many patients not doing enough blood tests which in turn results in less control of the blood glucose level. This can lead to hypoglycaemia, hyperglycaemia or later life health complications. To address this issue researchers have tried to come up with non-invasive techniques for glucose measurement. The next chapter explains basic theory of non-invasive glucose measurement and the preliminary work done in this area.

References:

1. E. Coiera, "Medical Informatics, the Internet, and Telemedicine," *London: Arnold*, 1997.
2. Shortliffe, Edward Hance, and James J. Cimino, eds. *Medical Informatics: Computer Applications in Healthcare and Biomedicine*. Springer Verlag, New York, 2006.
3. Young, Ralph R. "Recommended Requirements gathering practices." *CrossTalk: The Journal of Defense Software Engineering* 15.4 (2002): 9-12.
4. K. E. Wiegers, *Software requirements*: Microsoft press, 2009.
5. Lamb, W. "DIASEND [reg]: One device to read them all and in the clinic guide them." *Endocrine Abstracts*. Vol. 17. 2008.
6. Pressman, Roger S., and Darrel Ince. *Software engineering: A Practitioner's Approach. Vol*. 5. New York: McGraw-Hill, 1992.
7. Jones, Capers. "Software project management practices: Failure versus success." *CrossTalk: The Journal of Defense Software Engineering* (2004): 5-9.
8. Kendall, Kenneth E., et al. *Systems Analysis and Design*. Englewood Cliffs: Prentice Hall, 1992.
9. Holland, Philip, et al. "Using self-monitoring of blood glucose to improve understanding and self-management of diabetes in children and young people with type 1 diabetes in a routine clinical setting." *Endocrine Abstracts*. Vol. 27. 2011.
10. Mary, Shaw, and Garlan David. "Software Architecture: Perspectives on an Emerging Discipline." *Prentice-Hall* (1996).
11. Bosch, Jan. "*Software Architecture: The Next Step*." Software architecture. Springer Berlin Heidelberg, 2004. 194-199.

12. Fisher, Michelle. "METHOD AND SYSTEM FOR TRANSMITTING DATA BETWEEN A SERVER AND A MOBILE COMMUNICATION DEVICE USING SHORT MESSAGE SERVICE (SMS)." U.S. Patent Application 11/933,337.
13. Kaplan, Gadi. "Ethernet's winning ways." *Spectrum, IEEE* 38.1 (2001): 113-115.
14. Fielding, Roy, et al. "Hypertext transfer protocol–HTTP/1.1." (1999): 12.
15. Whitehead Jr, E. James, and Meredith Wiggins. "WebDAV: IEFT standard for collaborative authoring on the Web." *Internet Computing, IEEE* 2.5 (1998): 34-40.
16. Forouzan, Behrouz A. *TCP/IP Protocol Suite*. McGraw-Hill, Inc., 2002.
17. Miller, Brent A., Chatschik Bisdikian, and Tom Foreword By-Siep. *Bluetooth Revealed*. Prentice Hall PTR, 2001.
18. MA, Jinxing, and Qijun CHEN. "Analysis and Implementation of RFCOMM Layer in Bluetooth Protocol Stack [J]." *Computer Engineering* 8 (2004): 042.
19. Kanakis, Anastasios, Bilal Malik, and Mohammed Benaissa. "Low cost universal remote patient monitoring system." *High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESS), 2012 IEEE 14th International Conference on*. IEEE, 2012.
20. M. Benaissa, B. Malik, A. Kanakis, and N. Wright, "Tele-healthcare for diabetes management: A low cost automatic approach," *in Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 1290-1293.
21. Fei-yu, Lin, et al. "System on programmable chip development system."*Education Technology and Computer Science (ETCS), 2010 Second International Workshop on*. Vol. 1. IEEE, 2010.

22. B. Boehm and R. Turner, "Management challenges to implementing agile processes in traditional development organizations," *Software, IEEE*, vol. 22, pp. 30-39, 2005.
23. H. Y. Shahir, S. Daneshpajouh, and R. Ramsin, "Improvement strategies for agile processes: A SWOT analysis approach," *in Software Engineering Research, Management and Applications, 2008. SERA '08. Sixth International Conference*, 2008, pp. 221-228.
24. Wall, Larry, Tom Christiansen, and Jon Orwant. *Programming Perl*. O'Reilly Media, Incorporated, 2000.
25. Sobell, Mark G. *Practical Guide to Ubuntu Linux (Versions 8.10 and 8.04)*. Prentice Hall, 2008.
26. Wavecom, AT commands User Guide for WaveCom IP 5.01,Wavecom,Sep.[2008], [online]. Available http://www.embeddedarm.com/documentation/third-party/ts-gsm1-AT_Commands_User_Guide_Wavecom_IP_v5.01.pdf [Accessessed : 23 Jan. 2010].
27. Hua, C. H. E. N. "MVC design pattern." *Computer Knowledge and Technology (Academic Exchange)* 14 (2007): 051.
28. Leff, Avraham, and James T. Rayfield. "Web-application development using the model/view/controller design pattern." *Enterprise Distributed Object Computing Conference, 2001. EDOC'01. Proceedings. Fifth IEEE International*. IEEE, 2001.
29. MySQL, A. B. "*MySQL*." (2001).
30. Batini, Carlo, Stefano Ceri, and Shamkant B. Navathe. *Conceptual database design: an Entity-relationship approach. Vol. 116*. Redwood City, California: Benjamin/Cummings, 1992.

31. Ramakrishnan, Raghu, and Johannes Gehrke. *Database Management Systems*. Osborne/McGraw-Hill, 2000.
32. Welling, Luke, and Laura Thomson. *PHP and MySQL Web Development*. Sams Publishing, 2003.
33. Lee, James B., and Brent Ware. *Open Source Web Development with LAMP: Using Linux, Apache, MySQL, Perl, and PHP*. Addison-Wesley Professional, 2003.
34. Rubin, Jeffrey, and Dana Chisnell. *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. Wiley, 2008.
35. Myers, Glenford J., Corey Sandler, and Tom Badgett. *The Art of Software Testing*. Wiley, 2011.

Chapter 4 **Non-invasive Glucose Measurement**

4.1 Introduction

As discussed in chapter 3, WithCare+ with continuous feedback from the diabetic team could motivate patients to do glucose level measurements 3 to 4 times a day as needed. However, there remains a problem with the current method of glucose measurement. The finger prick method for glucose level monitoring on a daily basis for many years is clearly inconvenient and results in many patients not doing enough blood tests, which in turn results in less control of the blood glucose level. This can lead to hypoglycaemia, hyperglycaemia or later life health complications. The difficulties involved with invasive measurement can clearly be addressed by a non-invasive method. In order to make non-invasive glucose measurement a reality, different types of spectroscopic and optical methods, such as near-infrared (NIR) spectroscopy, Raman spectroscopy, fluorescence spectroscopy and polarization have been extensively researched [1-5]. NIR spectroscopy has become an established technology due to its non-invasive, non-destructive, and painless nature. In addition, it requires little or no sample preparation and provides a reasonable signal to noise ratio [6-8].

However, ascertaining the concentration of glucose from NIR spectra in the presence of other signals, associated spectral variations and the underlying spectral noise remain an issue. This challenge can be handled by employing sophisticated multivariate data-analysis algorithms to build robust regression models for the calibration and prediction of glucose concentration from NIR spectra. Principal Component Regression (PCR), and Partial Least Squares Regression (PLSR) are popular methods traditionally utilized in the data-analysis of NIR spectra [9-13], but to date these techniques have not been

able to address the calibration challenges associated with non-invasive glucose measurement.

In this chapter, locally weighted partial least squares regression (LW-PLSR) is proposed for the determination of glucose concentration from NIR data. LW-PLSR is developed by combining the locally weighted regression (LWR) technique with PLSR. The efficacy of the proposed model is validated in the determination of concentration of glucose from samples of an aqueous solution composed of glucose, urea and triacetin in different concentrations. The experiments were carried out in a non-controlled environment or sample conditions.

The next section gives an overview of NIR spectroscopy, Section 4.3 gives an overview of pre-processing techniques, Section 4.4 explains the theory of multivariate calibration techniques and Section 4.5 describes experimental work carried out for collecting the spectral data. Section 4.6 explains the development of the proposed novel calibration model and Section 4.7 concludes the chapter.

4.2 NIR Spectroscopy

Spectroscopic techniques are used to determine the concentration of a substance by measuring how it interacts with light. Absorbance is a logarithmic ratio of the radiation falling upon a material, to the radiation transmitted through a material. When the amount of absorption is plotted against wavelength, the resulting curve is referred as a spectrum. Each material shows a specific and reasonably unique spectrum, depending on its chemical structure, physical state, and temperature.

NIR spectroscopy utilizes the near infrared region of the electromagnetic spectrum (2.5-0.78 μm). This type of spectroscopy is based on molecular overtone and

combination vibrations. The NIR spectrum of glucose possesses three absorbance bands in the combination (2-2.5μm) and the overtone (1.54-1.82μm) regions of the NIR spectrum centered at 2.115, 2.273, 2.326 μm and at 1.73, 1.69 and 1.61 μm respectively. The absorbtion of glucose in the short-wave region (0.7-1.33μm) is low. Most recent glucose analysis studies used the combination region where absorbtion of water is low and the glucose absorbance band is narrower (particularly at 2.273 μm) [14].

Near infrared (NIR) has become popular due to its advantages such as:

- It is capable of providing non-destructive analysis i.e. the samples can still be used after analysis.
- NIR spectroscopy is painless.
- It requires little or no sample preparation for analysis of sample.
- NIR spectroscopy is faster in comparison to other techniques with a reasonable signal-to-noise-ratio (SNR).
- The cost per sample is often much lower as compared to other analytical methods.

NIR spectroscopy received a big boost with the advent of the Fourier transform infrared (FTIR) spectrometer in the early 1970s. The FTIR spectrometer provides high SNR, short scanning time, high resolution, high precision and improved data processing as compared to conventional infrared systems [15]. The instrument obtains infra-red spectra by collecting the interferogram of a sample signal using a Michelson interferometer.

4.2.1 NIR for Non-invasive Glucose Measurement

NIR spectroscopy is based on collecting reflectance or absorption spectra of the tissue with a spectrometer. When a spot on the skin is illuminated with near-infrared light, the light is partially absorbed and scattered due to its interaction with chemical components in the illuminated tissue. The rest of the light which is either transmitted or reflected out of the tissue is received by optical detectors. An analysis of changes in the intensity of the light combined with multivariate calibration techniques [29] allows the extraction of the tissue's chemical components, including glucose.

Glucose produces one of the weakest NIR absorption signals per concentration unit of the body's major components. NIR spectroscopy measurement enables investigation of tissue depths in the range of 1 to 100 millimeters with a general decrease in penetration depth as the wavelength value is increased. NIR transmission through a finger web, ear lobe, and finger cuticle or reflected from the skin of the forearm has been attempted and NIR diffuse reflectance measurements performed on the finger cuticle have shown good correlation with blood glucose [30].

NIR transmission spectroscopy with multivariate calibration method was used by Robinson [31] to measure blood/tissue glucose concentrations in diabetic subjects. Burmeister and Arnold [32], employed near-infrared transmission spectroscopy to study non-invasive blood glucose sensing in different measurement sites like nasal septum, lip, cheek, and tongue. The results demonstrated that the tongue provides spectra with the highest signal-to-noise ratio as it contains the least amount of fat. The transmission spectra of the tongue in the 1400 -2000 nm range was collected and results showed a standard error of prediction (SEP) in excess of 54 mg/dl for all diabetic subjects [33].

Hiese and Marbach [34] have studied determination of glucose concentration through a diffuse reflectance measurement in the oral mucosa membrane in the spectral range of 1111 - 1835 nm and reported the best SEP value of 43 mg/dl.

Jagemann et al. [30] used a fibre optic probe to study diffuse reflectance over the 800 - 1350 nm range on the middle finger of the right hand. The blood glucose concentrations of the test persons were perturbed using carbohydrate loading. The results were evaluated using PLSR and radial basis function neural networks. In these tests, the mean root square prediction error was 3.6 mg/dl. This method has the advantage that no reagents are required for the measurements and that fibre optical components can be used. As a consequence only insulators are in direct contact with the skin. Furthermore the spectrometer can be constructed without moving parts, leading to a rugged design.

Diffuse reflectance studies of the inner lip also have shown good correlation with blood glucose and indicated a time lag of 10 minutes between blood glucose and the measurement signal [35].

In spite of all these studies in the area of non-invasive glucose measurement no successful device has been translated into clinic. A device of reasonable size, acceptable accuracy and at reasonable cost would be an instant medical and commercial success.

However, the NIR spectra of glucose in many samples in practice are weak in intensity, broad and overlapped. Furthermore, the quality of the collected spectra is influenced by several factors such as:

1. The high background spectra of water.

2. Baseline variations resulting from the instrumental and ambient variations.

3. High frequency noise, such as the detector noise.

4. Highly overlapping of the spectra of other components because the structure of the glucose molecules and bonds are similar to many other constituents of samples.

5. The light scattering which is anisotropic, inhomogeneous, and nonlinear resulting from the surfaces and other constituents in a sample. The scattering reduces the SNR due to the degradation of the optical signal.

6. The low concentration of glucose in the blood serum as compared to other constituents.

NIR spectroscopy has attracted interest from healthcare [16], pharmaceuticals [17] and the food industry [18], although it is acknowledged that certain issues such as specificity and calibration still remain a challenge.

These challenges, which arise in quantitative analysis of glucose using NIR, can be overcome by using multivariate calibration methods coupled with pre-processing techniques. The ability of the multivariate model to forecast the concentration of the analyte of interest in the measurements is susceptible to the quality of the collected spectra. The practical data, even though collected from an instrument of high signal-to-noise ratio, is affected by negative factors such as baseline variations, noise, and scattering. The quality of the raw spectra is thus improved by using pre-processing methods before developing the multivariate model. The next section gives an overview of pre-processing techniques.

4.3 Pre-processing Techniques

The goal of pre-processing is to prevent unwanted variations from influencing the prediction capability of the multivariate model by removing the high frequency noise

and baseline variations, enhancing the resolution of the collected spectra, removing the high background, selecting the appropriate spectral range, and compensating for the negative influence of temperature variation. Derivatives and digital bandpass filtering are some of the techniques which are commonly employed in pre-processing of NIR spectra.

4.3.1 Derivatives

Derivatives of spectra have been known and used in NIR spectroscopy for a long time. NIR spectra often suffer from baseline shift, background noise, light scattering and multi-collinearity among variables, depending on the type of sample and instrument. The elimination of these unwanted variations is important for the development of a robust and accurate calibration model. It is possible to minimize the impact of these unwanted variations by applying suitable signal processing (pre-processing) methods. One of the common mathematical pre-processing methods used in chemometrics is taking derivatives to improve spectral resolution. The first derivative can eliminate baseline offset variations within spectra. However, the signal-to-noise ratio (SNR) deteriorates due to derivatives as they increase the noise. The higher the degree of differentiation used, the higher is the decrease in SNR. This problem is resolved by smoothing the data before differentiation and using a low degree of differentiation. The common smoothing and differentiation method in chemometrics is the one used by Savitzky and Golay [19] , which is based on a moving window averaging method.

The simple difference method, the Moving Window Polynomial Least Squares Fitting (MWPLSF) or the frequency domain techniques may be used for obtaining derivatives

of the spectrum. The first derivative of a discrete absorbance spectrum in a simple difference method can be expressed as:

$$a_i' = \frac{a_{i+1} - a_i}{\lambda_{i+1} - \lambda_i}$$

where a_i' is the first derivative of the spectrum, a_i is the absorbance spectrum at the wavelength λ_i, i=1…m and *m* is the number of variables (wavelengths). For equal differences between the sampling wavelengths, the first derivative can be simplified as follows:

$$a_i' = a_{i+1} - a_i$$

Despite its simplicity, the simple difference method has the main drawback that the peaks, maximum or minimum, are shifted relative to the raw spectra. This method is suitable for spectra with high resolution.

The MWPLSF method is based on the Savitzky-Golay smoothing filter. In this method, the *n-th* derivative is obtained by performing the smoothing operation against the *n-th* ordinary derivative of the polynomial, instead of using the polynomial directly as in the smoothing. Once the weighting coefficients are obtained, they are used to obtain the differentiated spectra. The MWPLSF method will not introduce shifting in the peaks of the differentiated spectra but will set the two sides of the differentiated spectrum to zero.

The most practical method used to compute the derivative of the spectra is that using the Fourier transform. Compared to the MWPLSF method, the Fourier differentiator is

simple, direct, fast and does not require complex tables. Furthermore, it does not produce a shifting in the peaks or suppress part of the spectrum ends. For example, the first derivative can be computed by multiplying the Fourier transform of the spectrum by a linear ramp. Then the inverse Fourier transform of the multiplier output is computed to produce the differentiated spectrum in the wavelength domain.

4.4 Multivariate Calibration Methods

Multivariate regression algorithms are widely used in quantitative analysis of NIR spectra to extract the concentration of the analyte of interest from the collected raw spectra of the samples[20]. This goal is achieved by generating a calibration model that is related to the response and predictor variables, which can be used later on to estimate the response to new data. The model development involves two phases:

1. Calibration phase

2. Validation phase

Spectral data are divided randomly into the training dataset and test dataset. In the calibration phase, a training dataset is used to generate the calibration model, and a test dataset is used to test the generalization capability of a calibration model to predict the concentration.

Generally, the performance of the calibration model can be evaluated by computing the coefficient of determination (R^2), the root means square error of calibration (RMSEC), the root mean square error of cross validation (RMSECV) and the root mean square error of prediction (RMSEP).

R^2 is a measure of how well the regression line fits the data. This statistic determines the amount of variation in the data that is adequately modelled by the calibration equation as a fraction of the total variation. In other words, it denotes the strength of the linear association between dependent values and independent values. The value of R^2 represents the percent of the data that is the closest to the line of best fit. This means if the regression line passes through every point on the scatter plot, then all of the variation in dependent variable is explained by the independent variable and R^2 will be equal to 1. Conversely, if the regression line is away from the points on the scatter plot, the value of R^2 will decrease. So R^2 values approaching 1.0 are attempted when developing calibration models. R^2 is calculated as below:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_p} (c_i - \hat{c}_i)^2}{\sum_{i=1}^{n_p} (c_i - \bar{c})^2} \quad (1)$$

where c_i is the actual concentration of the training data, $\hat{c}_i$ is the predicted concentration of the training data, $\bar{c}$ is the mean concentration and n_p is the number of samples.

The RMSEC describes the degree of agreement between the calibration model estimated concentration values for the calibration samples and the accepted true values for the calibration samples used to obtain the model parameters in Equation (2). Typically, RMSEC provides overly optimistic estimates of a calibration model's predictive ability for samples measured in the future. This is because a portion of the noise in the standards is inadvertently modeled by the estimated parameters. A better estimate of the calibration model's predictive ability may be obtained by cross-

validation with the calibration samples or from a separate set of validation samples. RMSEC is calculated as below:

$$RMSEC = \left(\sum_{i=1}^{n_p} (c_i - \hat{c}_i)^2 / n_p \right)^{1/2} \quad (2)$$

where c_i is the actual concentration of the training data, $\hat{c}_i$ is the predicted concentration of the training data, and n_p is the number of training samples.

The RMSEP measures the ability of the calibration model to predict the concentration of the analyte of interest for new data. The RMSEP is computed as below:

$$RMSEP = \left(\sum_{i=1}^{n} (t_i - \hat{t}_i)^2 / n \right)^{1/2} \quad (3)$$

where t_i is the actual concentration of the test data, is $\hat{t}_i$ is the predicted concentration of the test data, and n is the number of test samples.

Cross-validation is a statistical technique of estimating and comparing calibration models by dividing data into two segments: one used to train a model and other used to test the model. In cross-validation the training and test sets must cross-over in successive rounds so that each data point has a chance of being validated against. K-fold cross validation is the basic form of cross-validation while the other forms are special cases of k-fold.

In k-fold cross-validation the dataset is first partitioned into k equally sized data segments. Then k iterations of training and testing are performed such that different fold

of data is used for testing while the remaining k-1 folds are used for training. The main advantage of k-fold cross validation is the accurate performance estimation of the calibration model. The flow chart below depicts the process of k-fold cross validation. In chemometrics, data mining and machine learning 10-fold cross-validation (k=10) is the most common.

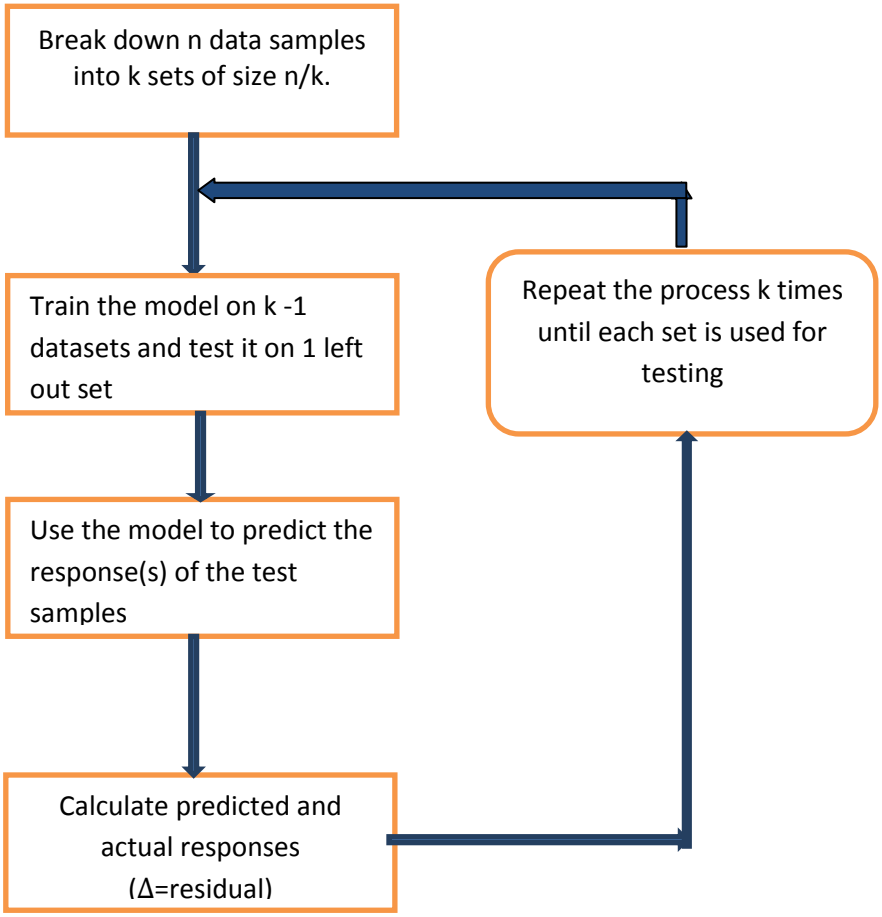


Figure 4.1 Flow Chart of K-Fold Cross Validation

Other common method of cross-validation is leave-one-out cross-validation (LOOCV), which is a special case of k-fold cross-validation. In LOOCV all the data except for a single observation are used for training and the model is tested on the left out observation. This process of leaving a sample out is repeated until all the calibration

samples are left out. LOOCV provides unbiased estimate but its variance can lead to unreliable estimates.

4.4.1 Multiple linear Regression (MLR)

A multiple linear regression (MLR) analysis is carried out to predict the values of a dependent variable given a set of explanatory variables. In spectroscopy, MLR is used to model the relation between the concentration and the absorbance spectra based on the inverse Beer-Lambert law [28] as:

Concentration= Bias + (Regression coefficient 1) * (Absorbance at Wavelength 1) +
(Regression coefficient 2) * (Absorbance at Wavelength 2) +
+…+ (Regression coefficient N) * (Absorbance at Wavelength N)

This can be written as:

$$C = \beta_0 + \beta_1 A_{\lambda 1} + \ldots.. + \beta_N A_{\lambda N} \quad (4)$$

where C is the estimated concentration of the analyte of interest, β_N is known as the regression coefficients vector, $A_{\lambda N}$ is the absorbance at each considered wavelength. The regression coefficients β_N are computed based on the Least Squares method [21,22].

$$\beta_N = (A^T A)^{-1} A^T C \quad (5)$$

MLR is one of the oldest methods used in multivariate analysis but has some drawbacks. The main problem with MLR is that the number of samples should be greater than or equal to the wavenumbers. However, in practice it is common to measure the NIR spectra with a number of wavenumbers greater than the number of

samples. Furthermore, increasing the number of samples may produce collinearity and nonlinearity, which cause the MLR calibration model to be unstable since it does not work well with collinearity and nonlinearity. To overcome these problems, principal component regression (PCR) has been developed based on principal component analysis (PCA) and MLR.

4.4.2 Principal Component Analysis (PCA)

PCA is a mathematical technique which was proposed by Pearson [23] and Hotelling [24], whilst the best modern reference is from Joliffe [21,22]. This technique takes advantage of an orthogonal linear transformation to convert a set of correlated independent variables into uncorrelated variables known as principal components. In this transformation, the first principal component accounts for most of the variance in the data and each successive component has the highest variance as compared to the other principal components with the constraint of being orthogonal to the preceding components. Thus, PCA transforms the data to a new coordinate system such that the first principal component lies on the first coordinate and the second principal component lies on the second coordinate and so on. Usually only the first few principal components are used and the final components are ignored as they are assumed to model noise. This is how the PCA reduces the dimensionality of the multivariate data without losing useful information, and this reduction in dimensionality helps to find the patterns in the multivariate data, which are otherwise hard to find in high dimensions.

In PCA, the original data matrix X, consisting of n spectra (rows) containing p datapoints, is described by the scores S matrix and loadings L matrix as shown in eq (6). The matrices S and L can be calculated by the nonlinear iterative partial least squares (NIPALS) algorithm [21,22]. The matrix E consists of the difference between the values

in the data matrix and the values obtained from the multiplication of S and L^T. It is called a residual matrix (or error matrix). Each datapoint is described by eq (7).

$$X = SL^T + E \tag{6}$$

$$x_{ij} = \sum_{n=1}^{p} s_{nj} l_{in} + e_{ij} \tag{7}$$

where i is the sample number and j is the spectral wavelength.

4.4.3 Principal Component Regression (PCR)

PCR is one of the most popular regression techniques for the development of calibration models for NIR data. It is a regression analysis technique which utilizes principal component analysis (PCA) when estimating regression coefficients.

It is implemented in two steps. In the first step the scores of the data matrix are computed using PCA; then the scores are regressed against the concentration of the analyte of interest to generate the MLR calibration model instead of the original absorbance spectra.

The first step of principal component analysis (also referred as eigen analysis) can be performed by singular value decomposition (SVD) on the data matrix X. This operation decomposes X into three matrices which are termed as the left singular values (LSV) matrix (or the U matrix); the singular values matrix; and the right singular values matrix (RSV) (or the V matrix), which is referred to as the P matrix in PCA terminology. It is now easy to find the scores matrix and the Loadings matrix from these decomposed matrices.

The scores matrix is calculated as:

S= X * V where S is the scores matrix, X the data matrix and V the loadings matrix.

The S and V can be used to calculate X as below:

$$X \text{ (estimated)} = S * V^T \quad (8)$$

If c represents the concentration matrix, then the set of regression coefficients is calculated as below:

$$b = V * S^{-1} * U^T * c \quad (9)$$

The predicted or estimated values of c are computed as

$$c \text{ (estimated)} = (S * V^T) * b \quad (10)$$

The PCR calibration model can be realized regardless of the selected wavelengths, and it is reliable against collinearity. The PCR model is robust and stable against noise if the number of observations is high, since it can eliminate loading factors that have low variance and low influence on the change of the absorbance spectra. On the other hand, PCR has some limitations that may degrade the performance of the calibration model. For instance, sometimes the analyte of interest only provides a small contribution to the variation of the spectra A and the information related to it may be eliminated during the decomposition process, especially if the signal to noise ratio of the measured spectra is low. Hence, the quality of the PCR model is sensitive to the number of loading factors that are used to generate the calibration model. Secondly, the PCR algorithm decomposes the absorbance spectra matrix A only, without using prior information about the analyte concentrations. Thirdly, in the presence of outlier data the principal

components may not represent the maximum variation of the information correctly because the covariance matrix is sensitive to outlier data [21,22].

4.4.4 Partial Least Squares Regression (PLSR)

Partial Least Squares Regression (PLSR) has become the most widely used technique for quantitative analysis of NIR spectra as it overcomes most of the serious drawbacks of other multivariate techniques. It decomposes the concentration and the absorbance spectra simultaneously, and uses prior information about concentrations in the decomposition of the spectra, and vice versa.

Compared to PCR regression, PLSR is more robust, faster, and requires a lower number of factors to establish the calibration model [21,22]. In contrast to PCR, the PLSR model extracts and sorts the loading factors according to their covariance with the scores of the absorbance matrix and the concentration vector. Thus, information related to the concentration of analyte of interest will be enhanced even if it has a low contribution in the absorbance spectra, because the chosen factors weight the spectra according to their correlation with the concentration.

PLSR utilizes a special case of singular value decomposition (SVD) which uses both the data matrix and concentration vector and decomposes them to their scores and loading matrices. This special case of SVD is sometimes referred as PLSSVD to distinguish it from SVD which is used in the case of PCR as discussed in previous section.

If X represents a data matrix, c is the concentration vector and the PLSSVD operation is performed on these matrices, the result is three matrices. These matrices are known as the left singular values (LSV) matrix (or the U matrix); the singular values matrix; and the right singular values matrix (or the V matrix). These decomposed matrices make it

easy to find the Scores matrix and Loadings matrix. The Loadings matrix is simply the V matrix; this matrix is referred to as P matrix in PLS terminology. The PLS Scores matrix is calculated as:

S = X * V where S is PLS Scores matrix, X is the data matrix and V is the PLS loadings matrix. Now S and V matrices can be used to calculate the following:

$$X \text{ (estimated)} = S * V^T \quad (11)$$

The set of regression coefficients is calculated as follows:

$$b = V * S^{-1} * U^T * c \quad (12)$$

The predicted or estimated values of c are computed as below:

$$c \text{ (estimated)} = (S * V^T) * b \quad (13)$$

The PLS model can extract a number of factors; however, the first few factors have information about the analyte of interest while the last factors may be the result of noise. Furthermore, increasing the number of factors may produce overfitting, which may lead to degradation of the performance of the calibration model. Therefore, the factors that have information about the analyte of interest should be determined to prevent overfitting and allow a high capability of prediction.

4.4.5 Locally Weighted Regression (LWR)

Linear models are sometimes limited due to the chemical properties of a measuring object, which have an intricate effect on NIR spectra. Another issue in building robust models is how to manage the variations in the process characteristics, which is vital in the chemical industry. Thus, maintenance of the models is an important issue to

consider in soft-sensors [9]. Traditional linear models generally discard the data after the training phase, losing potentially valuable extra information during the prediction phase. This is why "memory-based" methods, such as locally weighted regression (LWR), retain the training data and use it for each prediction, has been advocated for dynamic processes. LWR is a technique for non-parametric regression, which performs regression around a point of interest, using only training data that are "local" to that point [25].

In LWR, the goal is to fit θ_τ to minimize $\sum_i w_i(\mathbf{y_i} - \theta_\tau \mathbf{x_i})^2$ where $\mathbf{x_i}$ is the i-th row of the input data matrix, $\mathbf{y_i}$ is the output vector, and the w_is are non-negative valued weights. For a particular value of i, if w_i is large then θ is chosen in such a way as to make $(\mathbf{y_i} - \theta_\tau \mathbf{x_i})$ small. However, the $(\mathbf{y_i} - \theta_\tau \mathbf{x_i})$ error term will be ignored in the fit if w_i is small.

The weights depend on the particular point $\mathbf{x_i}$ Moreover, if $|\mathbf{x_i} - \mathbf{x}|$ is large, then w_i is small and if $|\mathbf{x}_i - \mathbf{x}|$ is small then w_i is close to 1. Thus, the value of θ is selected in such a way as to give a higher "weight" to the errors in training examples close to the query point $\mathbf{x}$. τ is called the bandwidth parameter and it controls how quickly the weight of a training example falls off with the distance of its $\mathbf{x}_i$ from the query point $\mathbf{x}$.

The following equation can generalize appropriate choice of weights for a vector $\mathbf{x}$ and bandwidth parameter τ:

$$w_i = \exp\left[-\frac{(\mathbf{x}_i - \mathbf{x})^T(\mathbf{x}_i - \mathbf{x})}{2\tau^2}\right] \quad (14)$$

LWR has been used in agriculture and the food industry [18]. Kim et al have used LWR for estimation of active pharmaceutical ingredients[26]. However, as far as can be

determined, no attempt has been made to date to use LWR for predicting glucose concentrations from NIR spectra.

In this work, a local linear regression model using LW-PLSR is developed for the quantitative analysis of glucose using NIR spectra, for the first time. It is also shown, using practical data in a non-controlled environment, the proposed LW-PLSR technique performs better than the conventional linear regression techniques.

4.5 Experimental Work

In this study, samples were prepared by mixing glucose, urea and triacetin in a phosphate buffer solution. Urea and triacetin are used to model the urea and triglycerides in blood, respectively. Thirty samples were prepared with different concentrations in each sample ranging from 20 mg/dL to 500 mg/dL for glucose, 0 to 50 mg/dL for urea, and 10 to 190 mg/dL for triacetin to reflect their concentrations in blood. The buffer solution was prepared by dissolving 3.4023 g of potassium di-hydrogen phosphate and 3.0495 g of sodium mono-hydrogen phosphate in distilled water. 5-Fluorouracil was added as a preservative. The aqueous solutions of glucose and urea were prepared by dissolving dry solutes in the buffer solution, whereas, triacetin solution was prepared by diluting it with the buffer solution.

A Fourier transform spectrometer (FTIR Cary 5000 version 1.09) was used for collecting the spectrum from the samples, in the wavelength region of 2100nm to 2400nm with a spectral resolution of 1 nm.

The samples were placed on an infrared quartz cuvette with a fixed pathlength of 1mm. The spectrum of each sample was collected 3 times without removing the sample from the spectrometer and in this manner a total of 90 spectra were collected. The absorbance spectrum of a buffer solution was used as the reference.

Many previous studies have been carried out in a controlled environment to compensate for the effect of baseline variations. However, in this study the experiments were carried out in a non-controlled environment in order to validate the ability of the proposed models, to deal with the uncompensated variations, and determine the concentration of glucose from the NIR spectra of the mixture solutions.

4.6 Model Development, Results and Discussion

The collected spectra, spanning the whole range of concentrations, were divided randomly into two datasets. The training dataset consisted of 60 spectra and test dataset consisted of 30 spectra. The training dataset was used for the calibration and the test dataset for validation of the models respectively.

4.6.1 First Derivative as a Pre-processing Technique

PCR, PLSR, and LW-PLSR were initially developed without pre-processing. All the models were developed following the same basic procedure. The models were built by utilizing Matlab version R2010a.

The "10-fold" cross validation [27] was performed to find the optimal number of principal components (PCs) and latent variables (LVs) in case of PCR and PLSR respectively. In the cases of PCR and PLSR, most of the variance is explained by 6 factors as illustrated in figure 4.2 and figure 4.3 respectively.

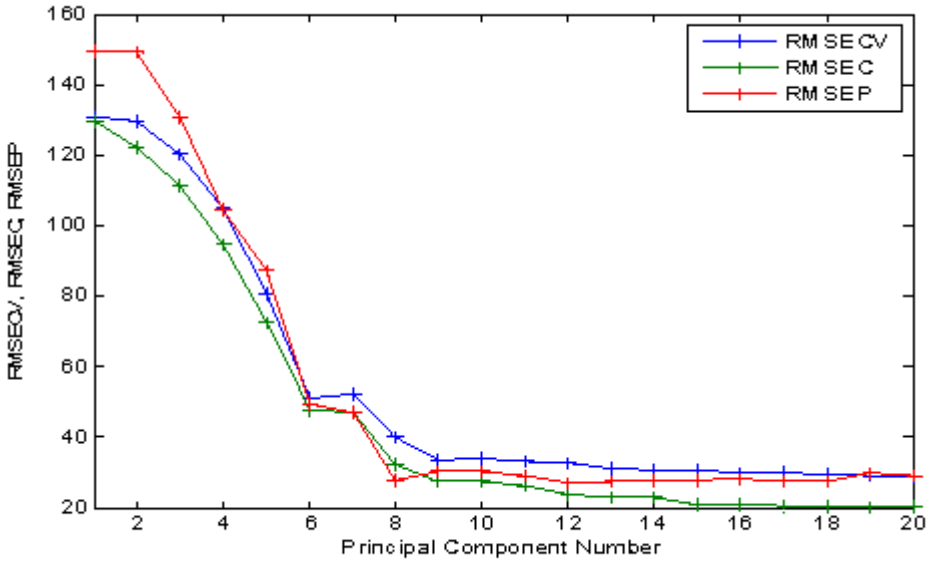


Figure 4.2 Variation Captured in the PCR Model

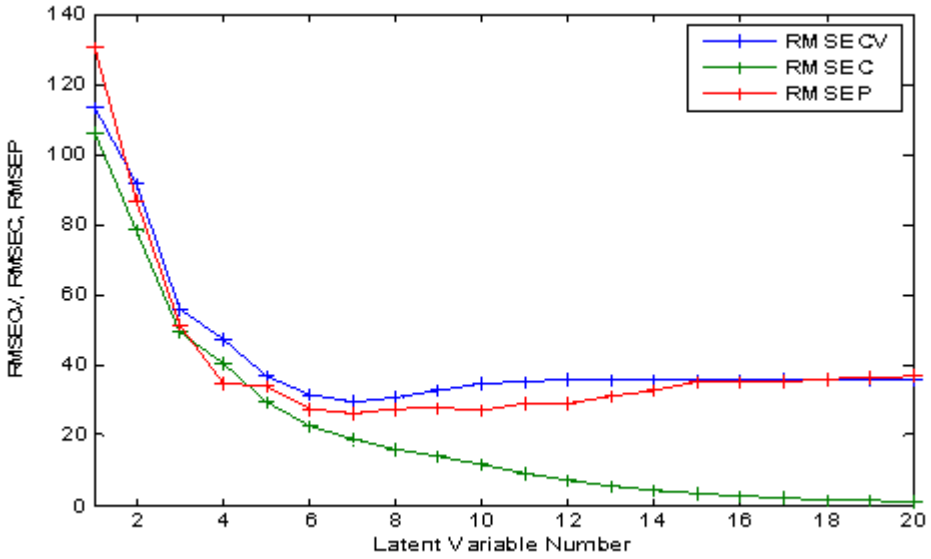


Figure 4.3 Variation Captured in the PLSR Model

The motivation for using the locally weighted regression was to compare the ability of LWR-PLSR to predict the concentration of glucose in the mixture solution of glucose, triacetin and urea in comparison to PCR and PLSR.

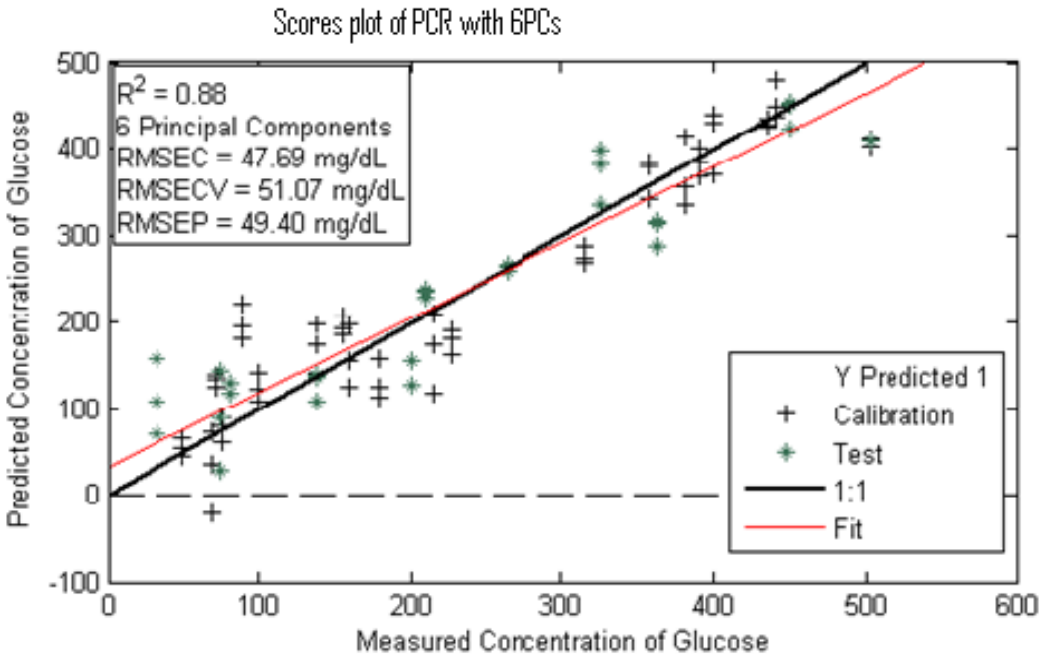


Figure 4.4 Scores Plot of the PCR Model

Figure 4.4 above shows the scores plot generated using PCR. As shown in the plot, R^2 is 0.88, RMSEC is 47.69 mg/dL, RMSECV is 51.07 mg/dL and RMSEP is 49.4 mg/dL.. Next, a calibration model was developed using the PLSR. The scores plot generated using PLSR is shown in figure 4.5 where R^2 , RMSEC, RMSECV and RMSEP were improved to 0.97, 22.54 mg/dL, 31.59 mg/dL and 27.56 mg/dL, respectively. However, as can be seen the RMSEP is still high.

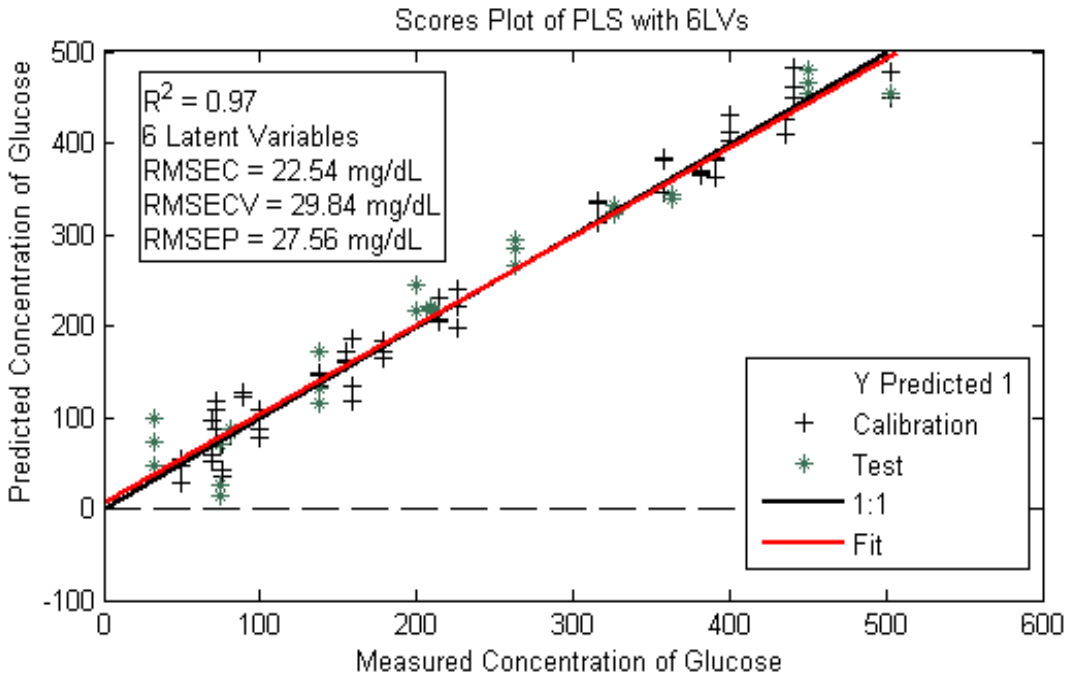


Figure 4.5 Scores Plot of PLSR Model

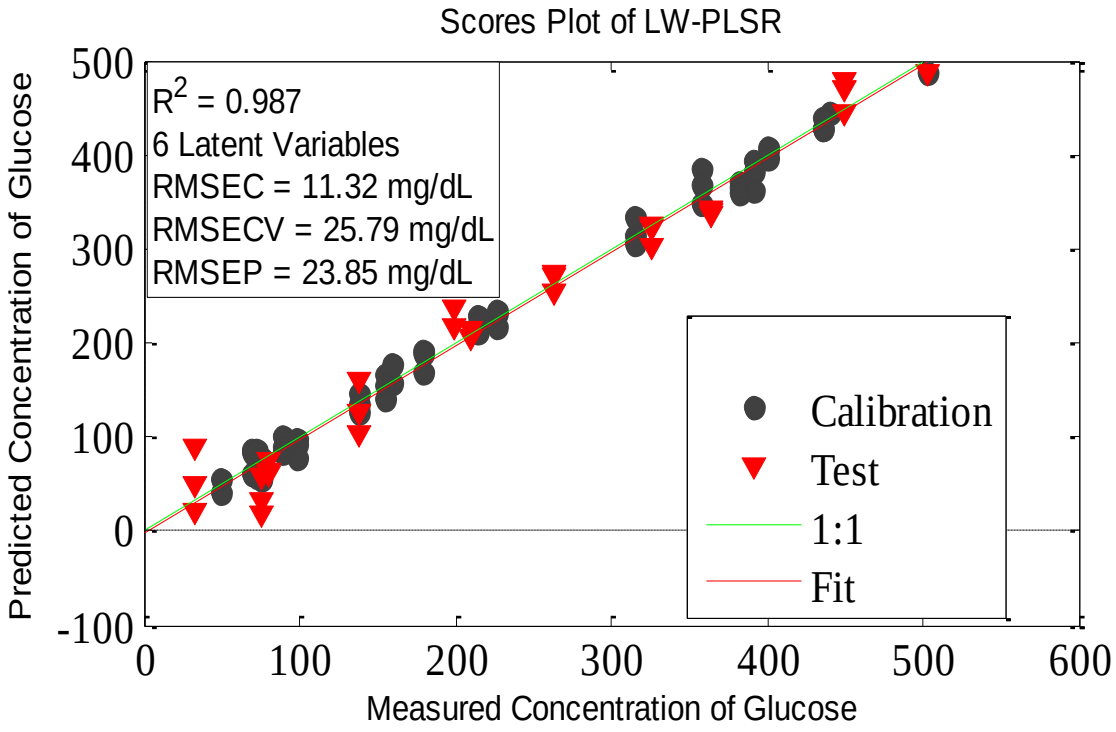


Figure 4.6 Glucose Prediction Performance using the LW- PLSR

Then, a predictive model was developed utilizing the LW-PLSR technique. As shown in figure 4.6, R^2 , RMSEC, RMSECV and RMSEP have improved to 0.98, 11.32 mg/dL, 25.79 mg/dL and 23.85 mg/dL respectively.

Table 4.1 below summarises the comparative results of the three calibration models developed. The data was pre-processed using the first derivative in all these models. As is evident from the table, LW-PLSR performed better in comparison to PCR and PLSR.

Table 4.1 comparison between PCR, PLSR and LW-PLSR

	R^2	RMSEC (mg/dL)	RMSECV (mg/dL)	RMSEP (mg/dL)	Optimal Parameters	Pre-processing
PCR	0.93	47.69	67.59	35.29	6PCs	None
PCR	0.88	21.18	51.07	49.40	6PCs	First Derivative
PLSR	0.94	10	34.07	32.81	6LVs	None
PLSR	0.97	22.54	31.59	27.56	6LVs	First Derivative
LW-PLSR	0.95	15.15	30.17	28.35	6LVs	None
LW-PLSR	0.98	11.32	25.79	23.85	6LVs	First Derivative

4.7 Conclusions

In this chapter, LW-PLS has been investigated and validated for the quantitative analysis of glucose using NIR spectroscopy. The proposed novel model has been developed and applied to predict the glucose concentration in a mixture composed of triacetin, urea and glucose in a non-controlled environment.

The predicted results of the proposed model have been compared with models developed using PCR and PLSR on the same data under the same pre-processing conditions. The proposed model has been shown to yield improved performance in

terms of R^2 , RMSEC, RMSECV and RMSEP. This improvement in prediction is encouraging and may lead to the possibility of more specialists in the area of chemometrics using the LW-PLSR.

References:

1. J. Chaiken, W. Finney, P. E. Knudson, R. S. Weinstock, M. Khan, R. J. Bussjager, *et al.*, "Effect of hemoglobin concentration variation on the accuracy and precision of glucose analysis using tissue modulated, noninvasive, in vivo Raman spectroscopy of human blood: a small clinical study," *Journal of Biomedical Optics*, vol. 10, pp. 031111-031111, 2005.
2. D. Rohleder, W. Kiefer, and W. Petrich, "Quantitative analysis of serum and serum ultrafiltrate by means of Raman spectroscopy," *Analyst*, vol. 129, pp. 906-911, 2004.
3. B. D. Cameron, H. W. Gorde, B. Satheesan, and G. L. Cote, "The use of polarized laser light through the eye for noninvasive glucose monitoring," *Diabetes Technology & Therapeutics*, vol. 1, pp. 135-143, 1999.
4. I. Barman, G. P. Singh, R. R. Dasari, and M. S. Feld, "Turbidity-corrected Raman spectroscopy for blood analyte detection," *Analytical Chemistry*, vol. 81, pp. 4233-4240, 2009/06/01 2009.
5. M. A. Arnold, "Non-invasive glucose monitoring," *Current Opinion in Biotechnology*, vol. 7, pp. 46-49, 1996.
6. M. Ferrari, L. Mottola, and V. Quaresima, "Principles, techniques, and limitations of near infrared spectroscopy," *Canadian Journal of Applied Physiology*, vol. 29, pp. 463-487, 2004.
7. M. Blanco and I. Villarroya, "NIR spectroscopy: A rapid-response analytical tool," *TrAC Trends in Analytical Chemistry*, vol. 21, pp. 240-250, 2002.

8. C. Pasquini, "Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications," *Journal of the Brazilian Chemical Society*, vol. 14, pp. 198-219, 2003.
9. H. Martens and T. Naes, *Multivariate calibration*: John Wiley & Sons Inc, 1992.
10. T. Næs and H. Martens, "Principal component regression in NIR analysis: Viewpoints, background details and selection of components," *Journal of Chemometrics*, vol. 2, pp. 155-167, 1988.
11. H. Martens, T. Karstang, and T. Næs, "Improved selectivity in spectroscopy by multivariate calibration," *Journal of Chemometrics*, vol. 1, pp. 201-219, 1987.
12. H. Mark and J. Workman, *Chemometrics in Spectroscopy*: Academic Press, 2007.
13. Brereton, Richard G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*. Wiley, 2003.
14. Al-Mbaideen, Amneh, and Mohammed Benaissa. "Frequency self deconvolution in the quantitative analysis of near infrared spectra." *Analytica Chimica Acta*705.1 (2011): 135-147.
15. "Fourier transform infrared spectroscopy. Part II. Advantages of FT-IR." *Journal of Chemical Education* 64, no. 11 (1987): A269.
16. M. R. Robinson, R. Eaton, D. Haaland, G. Koepp, E. Thomas, B. Stallard, and P. Robinson, "Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation," *Clinical Chemistry*, vol. 38, p. 1618, 1992.
17. Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, and N. Jent, "A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 44, pp. 683-700, 2007.

18. Y. Ozaki, W. F. McClure, and A. A. Christy, *Near-infrared Spectroscopy in Food Science and Technology*: John Wiley and Sons, 2007.
19. A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry,* vol. 36, pp. 1627-1639, 1964.
20. M. J. Wabomba, G. W. Small, and M. A. Arnold, "Evaluation of selectivity and robustness of near-infrared glucose measurements based on short-scan Fourier transform infrared interferograms," *Analytica Chimica Acta,* vol. 490, pp. 325-340, 2003.
21. Adams, Mike J. *Chemometrics in Analytical Spectroscopy*. Royal Society of Chemistry, 2004.
22. Kramer, Richard. *Chemometric Techniques for Quantitative Analysis*. CRC Press, 1998.
23. Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2, 559-572.
24. Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417-441.
25. Cleveland, W.S., and S.J. Devlin. 1988. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of American Statistical Association*. 83: 596-610
26. Kim, S., Kano, M., Nakagawa, H., & Hasebe, S. (2011). Estimation of active pharmaceutical ingredients content using locally weighted partial least squares and statistical wavelength selection. *International Journal of Pharmaceutics*, *421*(2), 269-274.

27. Kohavi, Ron. "A study of cross-validation and bootstrap for accuracy estimation and model selection." *International Joint Conference on Artificial Intelligence*. Vol. 14. Lawrence Erlbaum Associates Ltd, 1995.
28. Swinehart D.F. "The Beer-Lambert law." *Journal of Chemical Education* 39, no. 7 (1962): 333.
29. Heise HM, Marbach R, Janatsch G & Krse-Jarres JD. Mulitivariate determination of glucose in whole blood by attenuated total reflection infrared spectroscopy. *Analysis Chemistry* 1989; 61: 2009 - 2015.
30. Jagemann KU, Fischbacher C, Danzer K et. al. Application of near-infrared spectroscopy for non-invasive determination of blood/tissue glucose using neural networks. *Zeitschrift fur Physikalische Chemie* 1995; 191:179 - 90.
31. Robinson MR, Eaton RP, Haaland DM, Koepp GW, Thomas EV, Stallard BR & Robinson PL. Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation. *Clinical Chemistry* 1992; 38(9): 1618-22.
32. Burmeister J & Arnold MA. Evaluation of measurement site for non-invasive blood glucose sensing with near-infrared transmission spectroscopy. *Clinical Chemistry* 1999; 45(9): 1621 - 1627.
33. Blank TB, Ruchti TL, Malin SF & Monfre SL. The use of near-infrared diffuse reflectance for the non-invasive prediction of blood glucose levels. *IEEE Laser Electro-Optics Society Newsletter* 1999; 13(5): 9-12.
34. Heise HM, Marbach R, Koschinsky TH & Gries HM. Non-invasive blood glucose sensors based on near-infrared spectroscopy. *Artificial Organs* 1994; 18: 439 – 447.
35. Marbach R, Koschinsky TH, Gries FA & Heise HM. Non-invasive blood glucose assay by near-infrared diffuse reflectance spectroscopy of the human inner lip. *Applied Spectroscopy*. 1993; 47(7): 875 - 881.

5.1 Introduction

As seen in chapter 4, when ascertaining the concentration of glucose in the presence of other signals from spectra of their mixtures, associated spectral variations and the underlying spectral noise remain an issue, which can be overcome by implementing sophisticated multivariate data-analysis algorithms to build robust regression models for calibration and prediction of glucose concentration from NIR spectra. Principal Component Regression (PCR), and Partial Least Squares Regression (PLSR) are popular methods traditionally utilized in the data-analysis of NIR spectra [1-4], but these techniques may not be the best choice with non-linear data. Non-linearity in NIR spectra may arise from various factors, namely: deviations from the Beer- Lambert law, which are typical of highly absorbing samples; non-linear detector responses; interactions between analytes, etc.

Multilayer perceptron-based neural network models have been reported [5-7] with relative success due to their strong non-linear modelling capabilities. The drawback of neural networks is the existence of more than one local minimum. Recently, support vector machines (SVM) [8-11] have demonstrated promising results similar to those of the neural-networks for non-linear data but without its pitfalls. SVM, a supervised learning algorithm, was initially developed to build the classification models and was later extended to build the regression models. SVM, when used to build the regression models, is known as Support vector regression (SVR). It has recently attracted growing research interest in chemometrics due to its good generalization ability, unique and global optimal solution [8-11].

Information about an NIR spectrum is present in the frequency components of its Fourier transformation which reside in the mid-band range, while the noise and baseline variations reside in the high and the low frequency range of the Fourier domain respectively. The quality of the raw spectra can potentially be improved by using an optimized digital bandpass (DBP) filter to enhance performance of the calibration model. In recent research, a time domain filtering technique has been used as a pre-processing technique on input data while building the PLS regression model [12].

The digital band pass filter was effective in filtering out the low frequency baseline variations and high frequency noise [13]. In particular, Chebyshev and Gaussian filters have been reported [17] to improve the results of PCR model for determination of glucose concentration but no attempt has been made to couple these filters with support vector regression. In this chapter, a novel model that couples SVR with a DBP filter is presented. The performance of the model is investigated using both time domain (Chebyshev) and frequency domain (Gaussian) filters for the quantitative analysis of glucose from NIR spectra. The model is validated using experimental data collected from the Fourier transform near-infrared (FT-NIR) spectrometer.

The next section gives an overview of digital bandpass filtering, Section 5.3 explains the theory of SVR and Section 5.4 describes experimental work carried out for collecting the spectral data. Section 5.5 explains the development of proposed novel calibration models and Section 5.6 concludes the chapter.

5.2 Support Vector Machines Theory

SVM is firmly grounded in the framework of statistical learning theory which was developed by Vapnik from 1960-1990 [9,10] and the research was further extended by Smola and his co-workers to its current form [11] . In SVM the model represents a

training sample set as vector mapped space so that the samples from the different classes are separated by a distinct gap which is as large as possible. The model is then validated on a test set for prediction accuracy. SVMs demonstrate high performance in practical applications when solving complex classification problems [9,10]. The SVMs were initially developed as a binary classification tool and were extended to solve regression problems [11] .

Support vector machines (SVM) are based on the structural risk minimization (SRM) principle, which is superior to the traditional empirical risk minimization (ERM) principle [9-11] . SRM seeks to minimize an upper bound of the generalization error; on the other hand, ERM only minimizes the training error. Thus, SRM usually achieves a higher generalization performance than ERM based neural networks in solving many machine learning problems. In SVM, the solution to the problem is only dependent on support vectors, which are a subset of the training data points. SVM is a predictive classification and regression tool, which uses machine learning theory to maximize the accuracy of prediction and automatically avoids over fitting of data. In this study, however, the aim is to use SVM for regression (SVR).

The commonly used variations of SVR are epsilon-SVR and nu-SVR [10]. The important parameters for building the SVR model are gamma, cost, and epsilon or nu.

Given a training dataset $(\mathbf{x}_i, \mathbf{y}_i)$ (where $\mathbf{y}_i \in R$, $\mathbf{x}_i \in R^n$, and $i=1\cdots m$, Where m is the number of training samples) $\mathbf{x}$ is the vector of predictors and $\mathbf{y}$ is the vector of responses. SVR approximates the predictive function f as follows [11]:

$$f(x)=\boldsymbol{\omega}\cdot\phi(x)+b \qquad (1)$$

where $\boldsymbol{\omega}$ is the weight vector in primal weight space; (.) denotes the dot product; ϕ(x) represents the high-dimensional feature space which is nonlinearly mapped from the input space $\mathbf{x}_i$ and $\mathbf{b}$ is the bias term. The coefficients $\boldsymbol{\omega}$ and $\boldsymbol{b}$ are estimated by minimizing the regularized risk function.

The optimization problem can be expressed as:

$$\frac{1}{2}||\omega||^2 + C\frac{1}{m}\sum_{i=1}^{m} L_\varepsilon\,(y_i, f(\mathrm{x}_i)) \qquad (2)$$

The $||\omega||^2$ is referred as the regularized term and minimizing $||\boldsymbol{\omega}||^2$ will ensure the function is as smooth as possible, thus acting as the function capacity controller. The term $L_\varepsilon(\boldsymbol{y}_i, f(\mathbf{x}_i))$ is the empirical error measured by the ε-insensitive loss function, as defined below :

$$L_\varepsilon(\boldsymbol{y}_i, f(\mathbf{x}_i)) = \begin{cases} |\boldsymbol{y}_i - f(\mathbf{x}_i)| - \varepsilon, & if\ |\boldsymbol{y}_i - f(\mathbf{x}_i)| \geq \varepsilon \\ 0\,, & otherwise \end{cases} \qquad (3)$$

The equation (3) defines ε- tube [12] and if the predicted value is within the tube the loss is zero, while if the predicted point is outside the tube, the loss is the magnitude of the difference between the predicted value and the radius ε of the tube. C is called the regularization constant.

To obtain $\boldsymbol{\omega}$ and $\boldsymbol{b}$, Eq. (2) is transformed to the primal objective function (4) by introducing the positive slack variables ξ_i, ξ_i^* as below:

$$\text{Minimize } \frac{1}{2}||\boldsymbol{\omega}||^2 + C\frac{1}{m}\sum_{i=1}^{m}(\xi_i + \xi_i^*) \qquad (4)$$

Subject to $\quad y_i - \omega.\phi(x_i) - b \leq \epsilon + \xi_i$ and $\omega.\phi(x_i) + b - y_i \leq \epsilon + \xi_i^*$

In most cases the optimization problem (4) can be solved more easily in its dual formulation

$$L = \frac{1}{2}||\boldsymbol{\omega}||^2 + C\frac{1}{m}\sum_{i=1}^{m}(\xi_i+\xi_i^*) + C\frac{1}{m}\sum_{i=1}^{m}(\eta_i\xi_i+\eta_i^*\xi_i^*)$$

$$-\sum_{i=1}^{m} a_i(\epsilon + \xi_i - y_i + \boldsymbol{\omega}.\phi(x_i) + \boldsymbol{b})$$

$$-\sum_{i=1}^{m} a_i^*(\epsilon + \xi_i^* + y_i - \boldsymbol{\omega}.\phi(x_i) - \boldsymbol{b}) \quad (5)$$

In the equation above, L is the Lagrangian and $\eta_i, \eta_i^*, a_i, a_i^*$ are Lagrange multipliers. Hence, the dual variables in (5) have to satisfy positive constraints,

$$\eta_i^*, a_i^* \geq 0 \quad (6)$$

It follows from the saddle point condition that the partial derivatives of L with respect to the primal variables are as defined in equations (7), (8), (9) and (10).

$$\begin{cases} \frac{\partial L}{\partial \omega} = \boldsymbol{\omega} - \sum_{i=1}^{m}(a_i - a_i^*)\phi(\boldsymbol{x}_i) = 0 & (7) \\ \frac{\partial L}{\partial b} = \sum_{i=1}^{m}(a_i^* - a_i) = 0 & (8) \\ \frac{\partial L}{\partial_{\xi_i}} = C - a_i - \eta_i = 0 & (9) \\ \frac{\partial L}{\partial_{\xi_i^*}} = C - a_i^* - \eta^*{}_i = 0 & (10) \end{cases}$$

Substituting (7), (8), (9) and (10) in (5) yields the dual optimization problem as shown in equation (11).

Maximize W (a_i, a_i^*)

$$=\sum_{i=1}^{m} \boldsymbol{y}_i(a_i - a_i^*) - \epsilon \sum_{i=1}^{m}(a_i + a_i^*) - \frac{1}{2}\sum_{i,j=1}^{m}(a_i - a_i^*)(a_j - a_j^*)\phi(\boldsymbol{x}_i.\boldsymbol{x}_j) \quad (11)$$

Subject to $\sum_{i=1}^{m}(a_i - a_i^*)$=0 and $a_i, a_i^* \in [0, C]$

In the derivation of equation (11), $\eta_i, \eta^*{}_i$ are eliminated through equation (10) which can be rewritten as $\eta_i^* = C - a_i^*$.

Equation (7) can be written as $\boldsymbol{\omega} = \sum_{i=1}^{m} (a_i - a_i^*)\phi(\boldsymbol{x}_i)$ and thus

$$f(x) = \sum_{i=1}^{m}(a_i - a_i^*)\ (\phi(\boldsymbol{x}_i).\phi(\boldsymbol{x})) + \boldsymbol{b} \tag{12}$$

This is known as support vector machines regression expansion, i.e. $\boldsymbol{\omega}$ can be completely described as a linear combination of training patterns x_i . As is clear from equation (12), the complete algorthim can be described in terms of the dot products between the data.

The parameter b can be calculated by exploiting the Karush–Kuhn–Tucker (KKT) conditions, which state that at optimal solution the product between dual variables and constraints has to vanish. In SVR this means:

$$\begin{cases} a_i[\in + \xi_i - \boldsymbol{y}_i + \boldsymbol{\omega}.\phi_i + b] = 0 \\ a_i^*[\in + \xi_i^* + \boldsymbol{y}_i - \boldsymbol{\omega}.\phi_i - b] = 0 \end{cases} \tag{13}$$

$$\begin{cases} \eta_i\xi_i = (c - a_i)\xi_i = 0 \\ \eta_i^*\xi_i^* = (c - a_i^*)\xi_i^* = 0 \end{cases} \tag{14}$$

From the equation (13) it follows that only for $|y_i - f(x_i)| \geq \varepsilon$ the Lagrange multipliers may be nonzero. So one gets a sparse expansion of ω, and samples that come with non-vanishing coefficients are called Support Vectors (SVs). From equation (13) and (14), b can be computed as :

$$\begin{aligned} \mathrm{b} &= \boldsymbol{y}_i - \boldsymbol{\omega}.\phi_i - \epsilon \ \ for\ a_i\ \epsilon\ (0, \mathrm{C}) \\ \mathrm{b} &= \boldsymbol{y}_i - \boldsymbol{\omega}.\phi_i + \epsilon\ for\ a_i^*\ \epsilon\ (0, \mathrm{C}) \end{aligned} \tag{15}$$

The inner product can be defined through a kernel function $K(x,x_i)$ according to Mercer's condition. So the equation (12) can be rewritten as follows:

$$f(x) = \sum_{i=1}^{m}(a_i - a_i^*)\ K(\mathbf{x}, \mathbf{x}_i) + \boldsymbol{b} \quad (16)$$

By applying kernels, all necessary computations can be performed directly in input space, without mapping to feature space $\phi(x)$. Some of the popular kernel functions are

(1) Linear kernel: $K(x_i,x_j) = < x_i, x_j >$

(2) Polynomial kernel: $K(x_i,x_j) = (< x_i, x_j > + p)^d$, $p > 0$, $d \in N$

(3) Multilayer Perceptron Kernel:

$K(x_i,x_j) = \tanh(\varphi < x_i, x_j > \theta)$, $\varphi > 0$, $\theta > 0$

(4) Gaussian radial-basis function (RBF) kernel

$K(x_i,x_j) = \exp(-||x_i - x_j||^2 / \sigma^2)$, where σ is a kernel parameter.

Using different kernel functions, it is possible to construct different learning machines with arbitrary types of decision surfaces.

First, based on the above theoretical statement of SVR, it is possible to estimate the regression using a set of linear functions that are defined in a high-dimensional feature space, while the inputs have non-linear performance. Second, SVR carries out the regression estimation by risk minimization, based on statistical learning theory, where the risk is measured using Vapnik's ϵ-insensitive loss function [11]. Finally, SVR implements the SRM principle, which minimizes the risk function consisting of the empirical error and the value of the confidence level.

5.3 Experimental Work

This section describes how the same collected spectra as discussed in chapter 4 were used to generate the DBPF-SVR, DBPF-PLSR and DBPF-PCR calibration models. In

this study, samples were prepared by mixing glucose, urea and triacetin in a phosphate buffer solution. Thirty samples were prepared with different concentrations in each sample ranging from 20 mg/dL to 500 mg/dL for glucose, 0 to 50 mg/dL for urea, and 10 to 190 mg/dL for triacetin to reflect their concentrations in blood.

A Fourier transform spectrometer (FTIR Cary 5000 version 1.09) was used for collecting the spectrum from the samples, in the wavelength region of 2100nm to 2400nm with a spectral resolution of 1 nm. The spectra for each sample were collected 3 times without removing the sample from the spectrometer and in this way a total of 90 spectra were collected. The absorbance spectra of a buffer solution were collected at the beginning of the experiments and used as the reference. The experiments were carried out in a non-controlled environment in order to validate the ability of the proposed model to determine the concentration of glucose from the mixture solutions.

5.4 Model Development, Results and Discussion

The collected spectra, spanning through the whole range of concentration, were divided randomly into two datasets. The training dataset consisted of 60 spectra from 20 samples and the test dataset consisted of 30 spectra from remaining 10 samples. The training dataset was used for the calibration and the test dataset for validation of the models respectively.

5.4.1 First Derivative as a Pre-processing Technique

PCR, PLSR and SVR were initially developed without pre-processing. All the models were developed following the same basic procedure. The models were built by utilizing Matlab version R2010a and LIBSVM [14], a library for support vector machines.

The “10-fold” cross validation was performed to find the optimal number of PCs and LVs in case of PCR and PLSR respectively as explained in chapter 4 section 4.6.1.

In the case of SVR, the choice of the kernel function is a key step. The radial basis function (RBF) kernel nonlinearly maps samples onto a higher dimensional space, and unlike the linear kernel, it can handle cases when the relation between class labels and attributes are non-linear. Furthermore, the linear kernel is a special case of RBF as reported by Keerthi and Lin [15]. The sigmoid kernel behaves similarly to RBF in certain cases [16] . The RBF kernel has fewer parameters than the polynomial kernel which reduces the model complexity. Hence, the RBF kernel exhibits fewer numerical difficulties in contrast to polynomial kernels whose values may go to infinity or zero.

In this study, the RBF was used as the kernel function because of its compactness as compared to other feasible kernel functions for developing the SVR model. In addition to σ, there are two other important parameters when using RBF kernels: the cost parameter C and epsilon ε. Kernel parameter σ controls the shape of the separating hyperplane and C represents the penalty associated with errors larger than epsilon. Increasing γ usually increases the number of support vectors. According to Eq. (4), the parameter *C* determines the trade-off between the model complexity and the degree to which deviations larger than ε are tolerated in the optimization formulation (4). In addition, referring to Eq. (11), *C* decides the range of values $0 \leq (a_i, a_i^*) \leq C$, i=1,...,m, assumed by dual variables used as linear coefficients in the SVR solution. Hence, a 'good' value for *C* can be chosen equal to the range of output values of the training data. However, the selection of parameter *C* also depends on the domain of application knowledge. In training the regression function, there is no penalty associated with points which are predicted within distance ε from the actual value. Decreasing ε forces closer fitting to the training data. In order to prevent the over-fitting problem, the cross-validation approach is exploited to find the optimal parameters. In K-fold cross-

validation, the training set is divided into K-subsets of equal size. Sequentially, one subset is tested using the regressor trained on the remaining (K − 1) subsets.

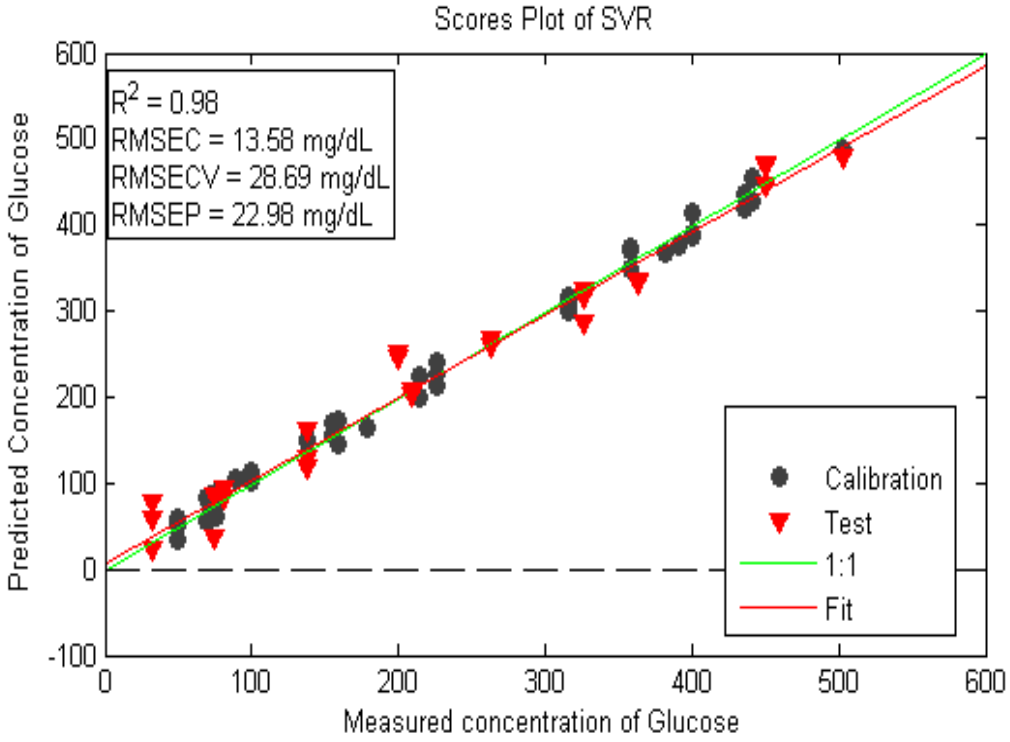


Figure 5.1 Scores Plot of SVR Model

The predictive performance of the models was evaluated by the following standards: correlation coefficient (R^2), root mean square error of calibration (RMSEC), root mean square error of cross validation (RMSECV) and root mean square error of prediction (RMSEP). A good model should have a high correlation coefficient, low RMSEC, low RMSECV, and low RMSEP.

Table 5.1 below summarises the comparative results of the three calibration models developed. The data was pre-processed using the first derivative and digital band pass filters in all these models. As is evident from the table, SVR coupled with digital band pass filters performed better in comparison to PCR and PLSR.

5.4.2 Digital Bandpass (DBP) Filtering as a Pre-processing Technique

The purpose of evaluating SVM for regression coupled with digital band pass filter was to compare its ability to predict the concentration of glucose in the mixture solution of glucose, triacetin and urea with the SVR pre-processed with derivatives. The results were further compared with the traditional linear methods such as PCR and PLSR coupled with DBP filters. Figure 5.2, 5.3 and 5.4 show the flow charts of selecting optimized parameters when coupling PCR, PLSR and SVR with DBP filter respectively.

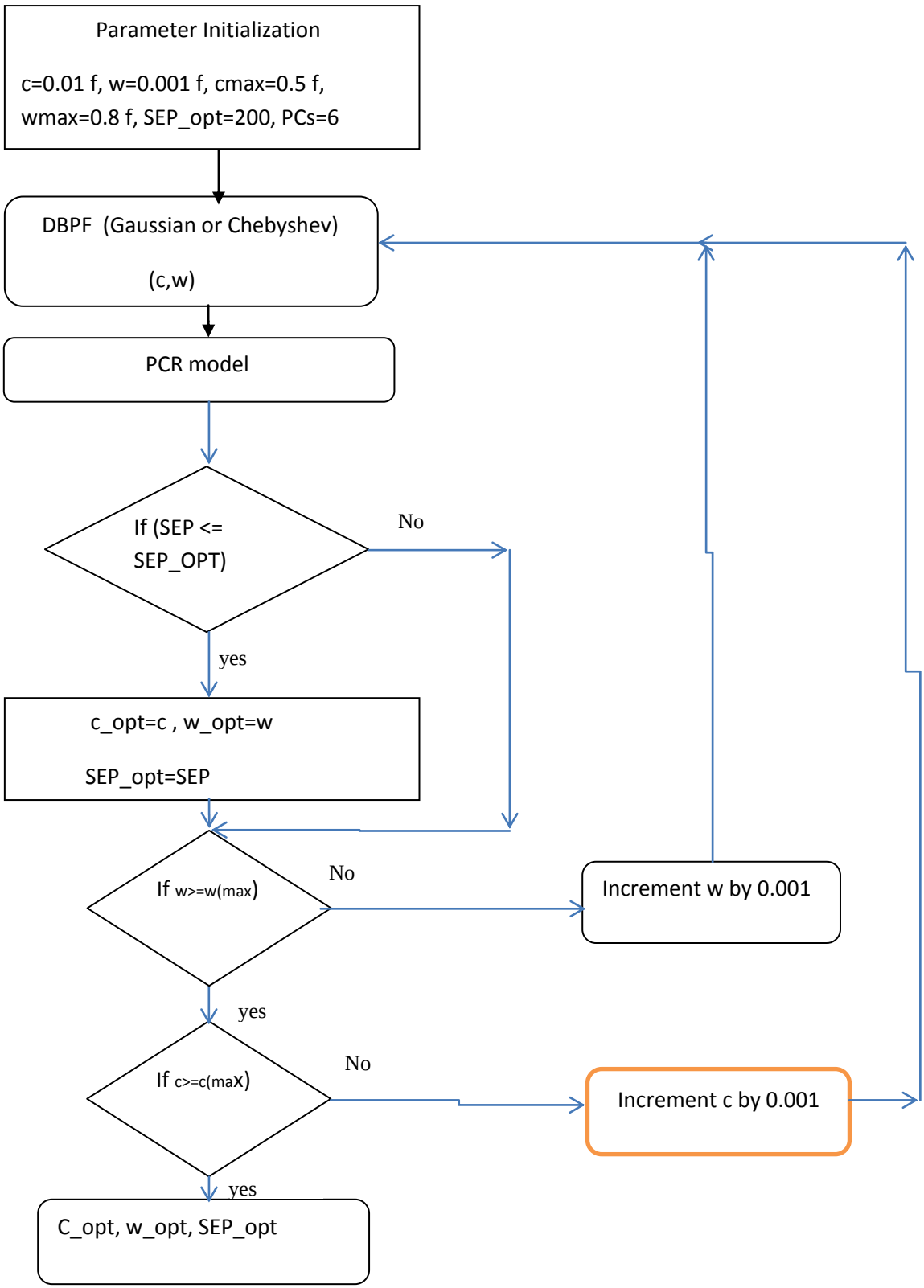


Figure 5.2 Flow chart of Parameter optimization for DBP using PCR Model

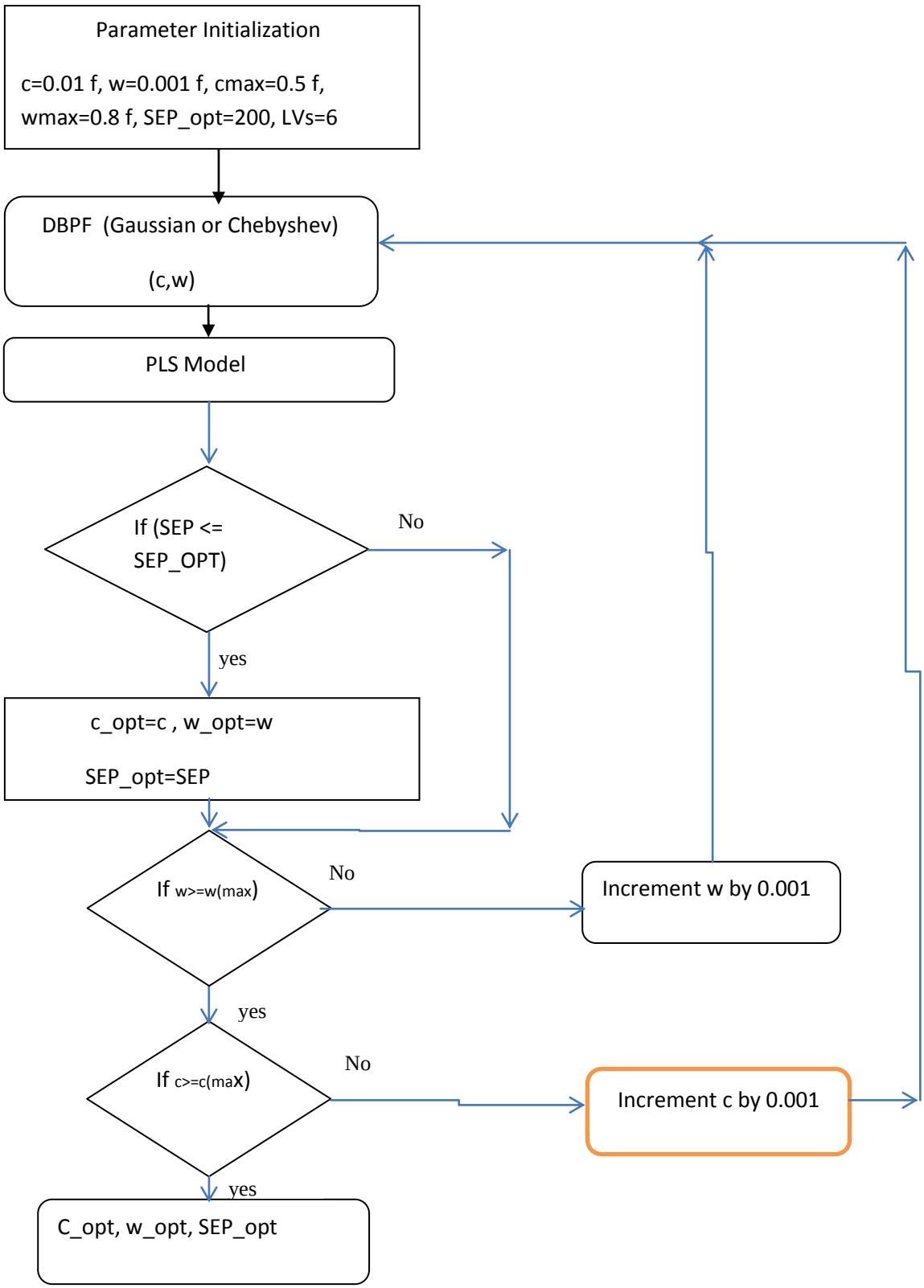


Figure 5.3 Flow chart of Parameter optimization for DBP using PLS Model

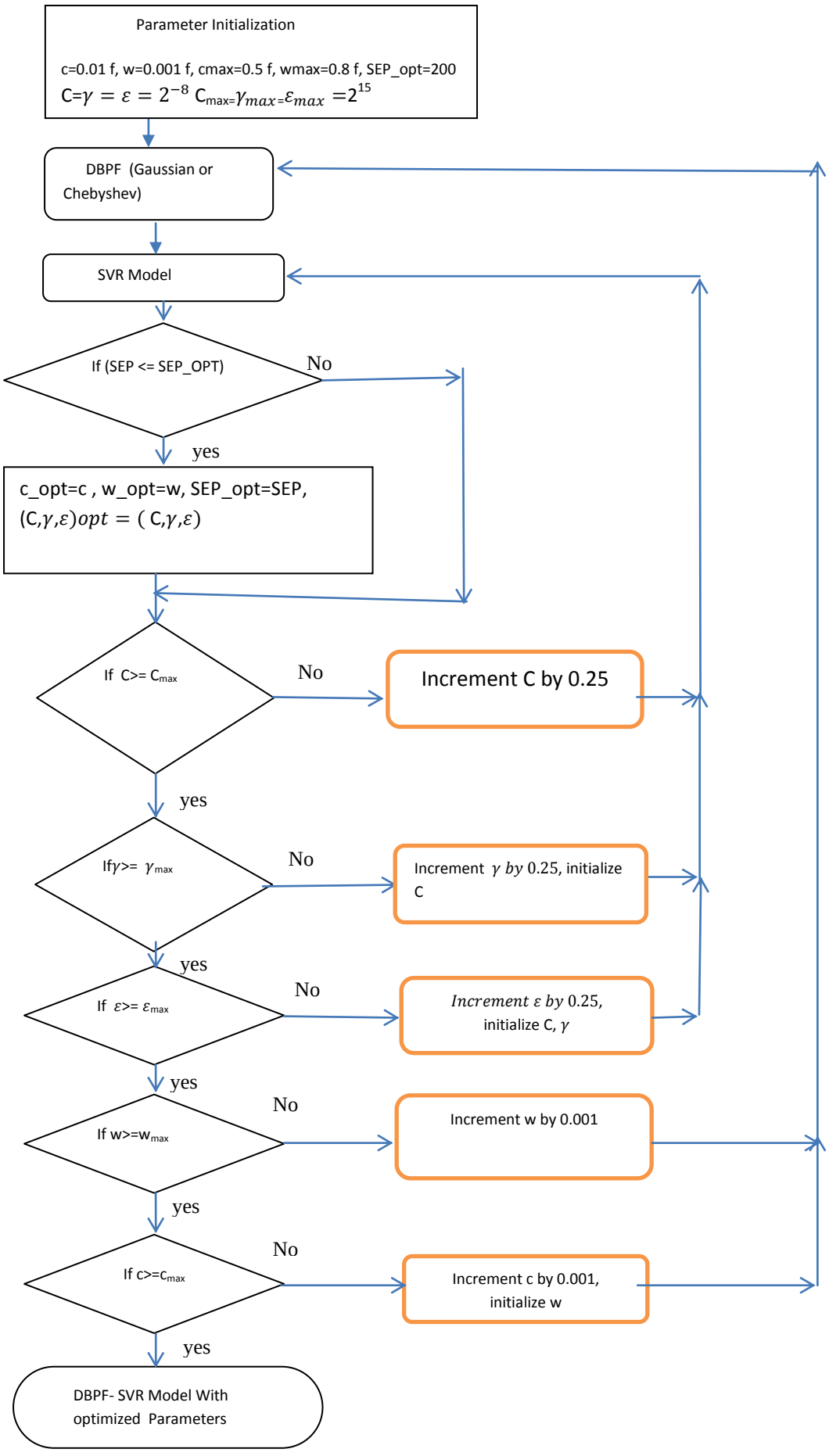


Figure 5.4 Flow chart of Parameter optimization for DBPF-SVR Model

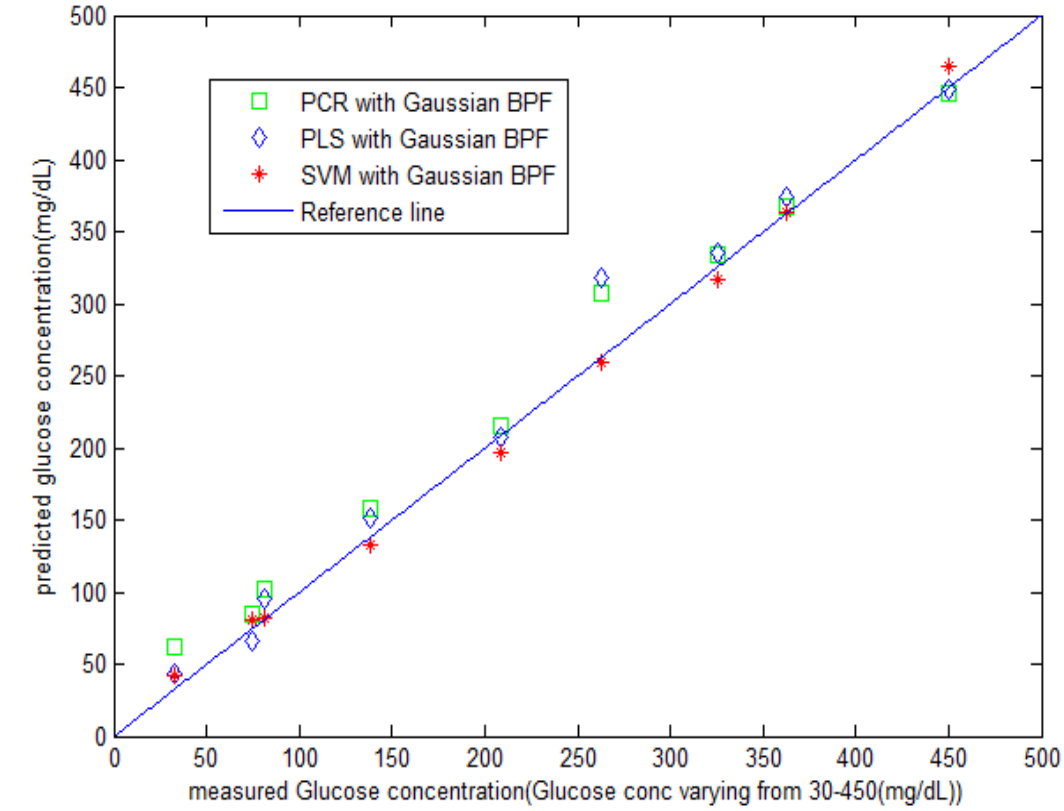


Figure 5.5 Comparison of the PCR, PLSR and SVR models using Gaussian Band Pass filter (BPF)

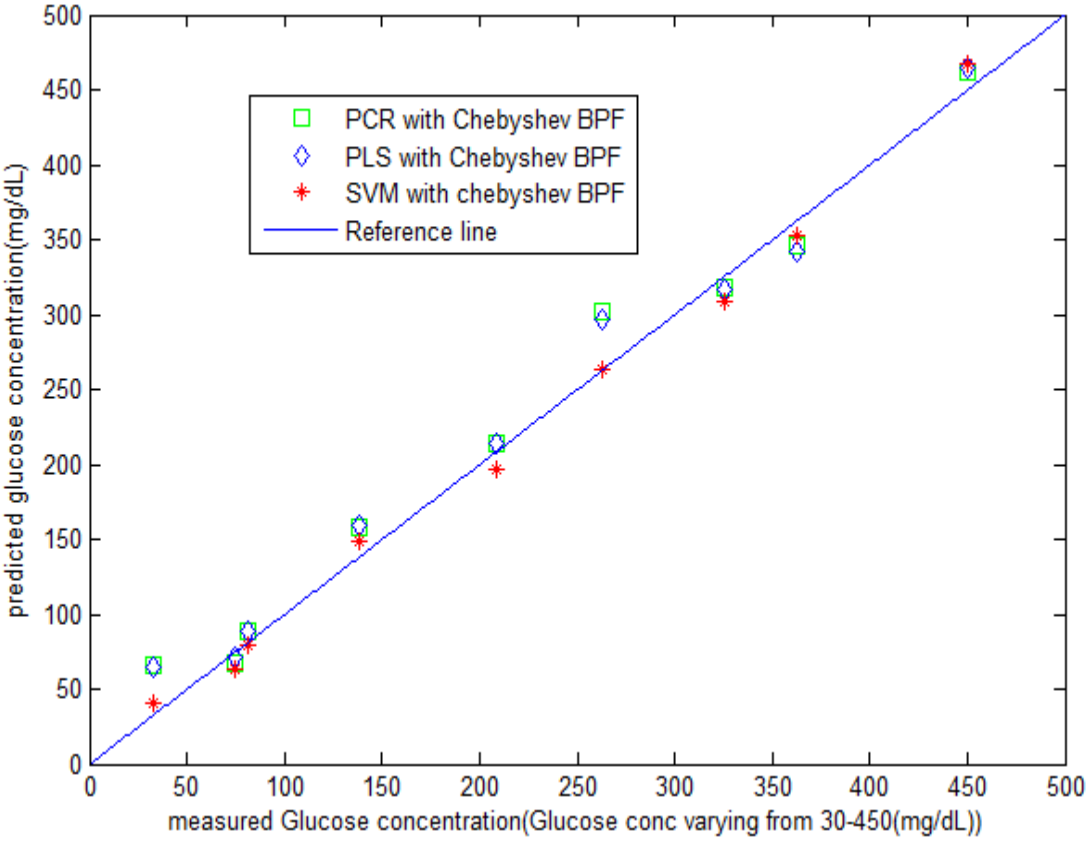


Figure 5.6 Comparison of PCR, PLSR and SVR models using the CBP filter

Figure 5.5 and Figure 5.6 show the comparison of PCR, PLSR and SVR models using the Gaussian and Chebyshev digital band pass filters for pre-processing respectively. The plots show measured glucose concentration on the x-axis and the y-axis represents the predicted glucose concentration for three different models.

Table 5.1 Comparison of SVR with PCR and PLSR

	R^2	RMSEC (mg/dL)	RMSECV (mg/dL)	RMSEP (mg/dL)	Optimal Parameters	Pre-processing
PCR	0.88	21.18	67.59	49.40	6PCs	None
PCR	0.93	47.69	51.07	35.29	6PCs	First Derivative
PCR	0.96	17.54	56.7	24.77	6PCs	GDBPF**
PCR	0.98	15.93	51.23	18.98	6PCs	CDBPF***
PLSR	0.94	10	34.07	32.81	6LVs	None
PLSR	0.97	22.54	31.59	27.56	6LVs	First Derivative
PLSR	0.96	12	38.3	24.59	6LVs	GDBPF**
PLSR	0.98	15.92	28.43	19.06	6LVs	CDBPF***
SVR	0.90	12.5	38.44	42	ε=0.1 γ=0.001 C=0.1*10^6	None
SVR	0.98	13.5	28.98	22.98	ε=0.2 γ=0.001 C=0.2*10^6	First Derivative
SVR	0.98	12.09	28.00	15.17	ε =0.04 γ=0.8 C=0.04*10^6	GDBPF**
SVR	0.98	12.47	27.40	14.59	ε =4.5 γ=1.56 C=4.5*10^6	CDBPF***

GDBPF**= Gaussian digital band pass filter CDBPF*** = Chebyshev digital band pass filter

5.5 Discussion

NIR spectra often suffer from baseline shift, background noise, light scattering and multi-collinearity among variables, depending on the type of sample and instrument. The elimination of these unwanted variations is important for the development of a robust and accurate calibration model. It is possible to minimize the impact of these unwanted variations by applying suitable signal processing (pre-

processing) methods. One of the common mathematical pre-processing methods used in chemometrics is taking derivatives to improve spectral resolution. The first derivative can eliminate baseline offset variations within spectra. In this work Savitzky and Golay [19] , which is based on a moving window averaging method, was employed for differentiation. As shown in table 1, employing the first derivative improves prediction capability of filter. However, the signal-to-noise ratio (SNR) deteriorates due to derivatives as they increase the noise. The higher the degree of differentiation used, the greater is the decrease in SNR.

Hence, in this work DBPF has been employed to compare their results with the first derivatives. Due to its simplicity, the DBP filtering has generated a great interest in different applications of signal processing. DBP filter, in context of NIR spectroscopy, is a set of mathematical techniques for improving the accuracy and reliability of raw spectra by filtering out undesired signals. This filter can be controlled and fine-tuned to obtain maximum information about the analyte of interest. It can be defined completely by its centre frequency and bandwidth. The DBP filter can be implemented either in the frequency or time domain. In time-domain filtering, the modified profile of raw spectra is acquired by convolving the raw spectra with the impulse response function of the desired filter. In frequency-domain filtering, fast Fourier transformation (FFT) of the raw spectra is computed, and the result of this transformation is then multiplied by the frequency response function of the digital filter. Finally, an inverse fast Fourier transform (IFFT) is computed on the result of the previous step to get the filtered spectra [13].

Chebyshev filters exhibit an optimal trade-off between a steeper roll off and passband ripple as compared to other time-domain filters. The filter is defined by its transfer

function and filtering operations are performed in the time domain. The order of the filter, magnitude and location of the ripple control the characteristics of the filter.

Gaussian filters are designed to give no overshoot to a step function input while minimizing the rise and fall time. A Gaussian filter can be implemented either in the time domain or frequency domain. The Gaussian function has the same profile in the time domain and frequency domain and it can be represented completely by its standard deviation and mean, which are identical to band width and centre frequency respectively [13]. Hence, in this work the Gaussian filter has been implemented in the frequency domain.

In the optimization process, the DBP filters were designed by varying the centre frequency(c) from 0.01f to 0.5 f and the bandwidth (w) from 0.001 f to 0.8 f with an interval of 0.001 f, where c and w are normalized between 0 and 1. In each iteration, the designed DBP filter is coupled to the calibration model and the RMSEP value is calculated. The calibration model with the minimum RMSEP value corresponds to the optimized DBP filter. The optimum parameters for the Gaussian filter were c=0.019 f, w=0.012 f and for the Chebyshev filter were c=0.03 f, w=0.10 f, where f is the normalized frequency between 0 and 1, c is the central frequency and w is the bandwidth of the filter.

5.6 Conclusions

In this chapter, DBPF-SVR have been investigated and advocated for the quantitative analysis of glucose using NIR spectroscopy. The novel model has been developed and applied to predict the glucose concentration in a mixture composed of triacetin, urea and glucose in a non-controlled environment. The prediction results of the proposed models have been compared with models developed using DBPF-PCR and DBPF-PLSR on the

same data under the same pre-processing conditions. The proposed model has been shown to yield improved performance in terms of R^2 , RMSEC, RMSECV and RMSEP.

This improvement in prediction is encouraging and may lead to the possibility of more specialists in the area of chemometrics using DBPF-SVR

References:

1. H. Martens and T. Naes, *Multivariate Calibration*: John Wiley & Sons Inc, 1992.
2. T. Næs and H. Martens, "Principal component regression in NIR analysis: Viewpoints, background details and selection of components," *Journal of Chemometrics*, vol. 2, pp. 155-167, 1988.
3. H. Martens, T. Karstang, and T. Næs, "Improved selectivity in spectroscopy by multivariate calibration," *Journal of Chemometrics*, vol. 1, pp. 201-219, 1987.
4. H. Mark and J. Workman, *Chemometrics in Spectroscopy*: Academic Press, 2007.
5. F. Despagne and D. L. Massart, "Neural networks in multivariate calibration," *Analyst*, vol. 123, pp. 157R-178R, 1998.
6. F. Marini, R. Bucci, A. Magrě, and A. Magrě, "Artificial neural networks in chemometrics: History, examples and perspectives," *Microchemical Journal*, vol. 88, pp. 178-185, 2008.
7. N. B. Karayiannis and A. N. Venetsanopoulos, *Artificial Neural Networks: Learning Algorithms, Performance Evaluation, and Applications* vol. 209: Springer, 1993.
8. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.
9. V. N. Vapnik, *The Nature of Statistical Learning Theory*: Springer Verlag, 2000.
10. C. M. Bishop, "Pattern Recognition and Machine Learning (Information Science and Statistics). 2006. corr. 2nd printing edn," ed: Springer (October 2007), 2007.
11. A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199-222, 2004.
12. Ham, F. M., Kostanic, I. N., Cohen, G. M., & Gooch, B. R., Determination of glucose concentrations in an aqueous matrix from NIR spectra using optimal time-

domain filtering and partial least-squares regression. *Biomedical Engineering, IEEE Transactions on, 44*(6), 475-485,1997.

13. Smith, S. W. *Digital Signal Processing: A Practical Guide for Engineers and Scientists*. Newnes, 2003.
14. C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST),* vol. 2, p. 27, 2011.
15. S. S. Keerthi and C. J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," *Neural computation,* vol. 15, pp. 1667-1689, 2003.
16. H.-T. Lin and C.-J. Lin. A study on sigmoid kernels for SVM and the training of non PSD kernels by SMO-type methods. Technical report, Department of Computer Science, National Taiwan University, 2003. URL http://www.csie.ntu.edu.tw/ ~cjlin/papers/tanh.pdf.
17. Al-Mbaideen, A.A.; Rahman, T.; Benaissa, M., "Determination of glucose concentration from near-infrared spectra using principal component regression coupled with digital bandpass filter," *Signal Processing Systems (SIPS), 2010 IEEE Workshop on* , vol., no., pp.243,248, 6-8 Oct. 2010 doi: 10.1109/SIPS.2010.5624795.

Chapter 6 **Conclusions and Future Work**

This chapter presents a summary of the work carried out for this dissertation and recommends future work for this research topic.

6.1 Summary and Conclusion

The research work presented in this thesis has focussed on developing a telemedicine system which can facilitate different stakeholders in a diabetic team. The members of a diabetic team are the patient, the parents or caregivers, the clinician, nurses and the dietician. The primary focus of WithCare+ is to improve the communication between different members of a type 1 diabetic team. The proposed system has been validated and accepted by healthcare professionals at the Children's Hospital, Sheffield. The clinical trial, which started in March 2013, will run for 18 months.

Even though healthcare providers can engage patients by providing continuous feedback with the help of WithCare+, invasive glucose measurement still remains a challenge, as it is demotivating for the patients to measure their blood glucose levels on a daily basis. This difficulty can be addressed by non-invasive glucose measurement. NIR spectroscopy is one of the most promising techniques that has been identified for non-invasive glucose measurement due to its non-destructive, non-invasive and painless nature. However, the clinical translation of the technology still remains a challenge due to the calibration of instruments for the determination of glucose concentration from the NIR spectra of blood serum or plasma. This research presents two novel calibration techniques which were developed by exploiting existing signal processing techniques and machine learning algorithms. The research has demonstrated that these techniques can successfully determine and predict the concentration of glucose from the NIR spectra of mixtures composed from glucose, triacetin and urea in a phosphate buffer solution. The concentration of components is selected to be within their physiological range in the blood. The experiments were carried out in a non-controlled environment

to show that the proposed calibration models can effectively suppress most of the experimental variation.

6.2 Future Work

Future directions of this work could focus on the following issues:

The performance of WithCare+ has been validated by healthcare professionals at Sheffield Children's Hospital and the system has passed the user acceptance test (UAT). The system is currently being validated in a clinical trial in a phased manner for 18 months. In future, the system will be validated by running a main trial with more patients for longer duration.

The system could also be extended to monitor other related parameters of type 1 diabetic patients such as insulin and calorie intake. This would enable healthcare professionals to advise patients on the optimum amount of insulin intake through an SMS based on the inputs calculated by a learning algorithm. This learning algorithm could be fine-tuned by the healthcare professional's knowledge and experience, which could result in an expert system. This can be further extended to the concept of the artificial pancreas using the principles of a closed loop control system.

The design of the server software and the database of WithCare+ is flexible enough to monitor other chronic diseases. The future work could enhance the WithCare+ patient unit so that the data from other physiological parameters can also be transmitted using a GSM network and the WithCare+ server software and database can be easily extended to support these extra physiological parameters.

For non-invasive WithCare+, the proposed calibration techniques for non-invasive glucose measurement in this study are validated in mixtures composed of three

components in aqueous solutions. In the future work, however, the proposed methods can be validated through more experiments using blood plasma or blood serum and a large number of samples.

Currently, each research group performs different experiments under different conditions. Therefore, it is recommended that a database of the datasets should be built for all research groups involved in non-invasive glucose measurement. Such a database would enable researchers to compare proposed calibration models using the same reference data without reinventing the wheel.

Appendix A List of Commonly used AT Commands

AT Commands	Meaning
AT+CGMI	Get the name of manufacturer
AT+GMI	Get the ID of manufacturer
AT+GMM	Get the Model information
AT+CGMR	Get software version
AT+CPAS	Get mobile phone activity status
AT+CREG	Get mobile network registration status
AT+CSQ	Get radio signal strength
AT+CMGS	Send SMS messages
AT+CMGR	Read SMS messages
AT+CMGW	Write SMS messages
AT+CMGD	Delete SMS messages
AT+CNMI	Notification of newly received SMS
AT+CPBR	Read phonebook entries
AT+CPBW	Write phonebook entries
AT+CPBF	Search phonebook entries
AT+CLCK	Opening or closing facility locks
AT+CPWD	Changing passwords

AT+CMEE=n	Enable certain error messages
AT+CMGF=1	to operate in SMS text mode
AT+CMGF=0	to operate in PDU mode
AT+CMGL="ALL"	List all the messages
AT+CNMA	New message acknowledgement

Appendix B SMS Message Format

SMS to be sent in 8-bit octet format, MAX SMS length thus 140 chars

Header (ASCII printable characters) Updated (New)

ModuleID	/	IV	/	Type	/	MeterModel	/	MeterID	/
5	1	var	1	1	1	var ~6	1	var ~16	1

ModuleID: alphanumeric identifier, 5 characters

IV: empty field for plaintext, hexadecimal number when encrypted

Type: single character

G	Glucose, mg/dl (18*mmol/L)

Meter Model: abbreviated string

AssCtr	Assensia Contour
OneTch	OneTouchUltra
FrMini	FreeStyle Mini
CareSN	CareSens N

MeterID: alphanumeric serial number reported by meter

Readings (binary data)

nReadings	[	DTG	value	N times	] (term)
1	1	4	2	x N	1

Length = 2 + 6N (for N readings)

nReadings: uint8 value containing the number of readings 1..15

DTG: uint32 date-time in seconds since 1Jan1970 (unix time)

value: uint16 reading value (interpretation dependant on format)

term: char; terminating character, ‘]’

endpadding: if necessary message may be padded with ASCII 0x00.

Appendix C

WithCare+ Patient Reference

What this system does

This system comprises an electronic module which you connect to your meter with the cable provided. When the app on the mobile phone is active and Start button is pressed then with the button press on the WithCare+ device it sends your readings via your mobile telephone using Bluetooth™ to send a Short Message Service (SMS) to a remote system which stores these results and allows your clinician to access them.

Benefits

This saves your time in having to record results and phone them in. It also automatically checks the results against limits defined by your doctor both in terms of individual readings and two-weekly average. If anything is abnormal it prompts your clinician who may contact you directly.

Connecting Your Meter

Only connect the WithCare+ module using the cable provided to the specific meter agreed with your clinician. The cables are uniquely configured for each type of meter and must not be used with a different meter model. If you need to change meter model stop using the WithCare+ device and consult your clinician.

For certain meters the module should only be plugged in just before sending your readings and disconnected shortly afterwards. Follow the instructions provided by your clinician.

Pairing with your phone

The WithCare+ module is a Bluetooth™ device and prior to use must be first "Paired" with your mobile telephone. Your clinician may do this for you.

Taking your readings

Take your readings as **instructed by your clinician** as normal and follow their instructions on what to do in the event of high or low values. You may record your readings manually if required. Avoid deleting readings from your meters memory as this will prevent the WithCare+ module from working.

Sending Readings

Each time a reading is taken the WithCare+ module can be requested to send it to the server. Steady amber light indicates the module is on but waiting for the button to be pressed (1-2 seconds). It should indicate flashing amber whilst searching your meter for results and finally a steady green on successful sending or red if the process has failed. If at any time it is wished to abort the sending process, press and hold the button for 12 seconds which will power off the module.

Feedback from clinician

The system gives your clinician a facility to quickly text you a message as feedback from your results if needed so do not be surprised if you receive a text shortly after sending some readings or at the start of the day.

Supported Meters

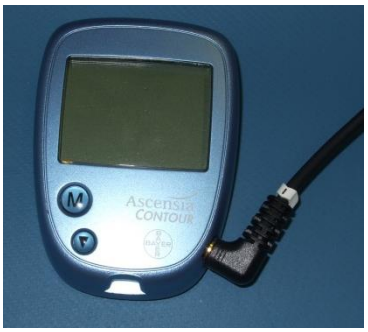
One Touch Ultra Easy

This may be left permanently connected.

CareSens N

This may be left permanently connected.

Bayer Assencia Contour



The correct cable is marked with “1”.

This may be left permanently connected.

Operation / Display

Button Press	***Operation***
short (1-2 seconds)	send readings
long (10 seconds)	Pair with phone
further (long) press	turn off module.

Display showing		***Operation***
☼	None	Module off
🟡	Amber	Waiting for button press
🟡☼	Flashing Amber	Attempting to read meter
🔴🟢	Alternating Red/Green	Attempting Pairing
🟢☼	Flashing Green	Sending data to mobile phone

✹	Green	Success
✹	Red	Failure!

Replacement Battery

The battery should last, in normal use, for more than a year. If problems should occur please consult your clinician. The battery should only be replaced with a similar AAAA battery. Ensure that the orientation is correct (positive + and negative -). Do not attempt to recharge the battery.

Personal Information

Name	
Mobile No.	
Module ID	___ ___ ___ ___ ___
Module PIN	___ ___ ___ ___
Meter Model	
Meter Serial	

- Check that the cable of withcare+ device is attached to glucometer (One Touch UltraEasy in our case).
- Check that Bluetooth is enabled on the phone. If it not ON, please switch Bluetooth ON.
- Start WithCare+ app on the BlackBerry and then press the Start button of withcare+ app on the screen.
- Now press the button on the WithCare+ for 2-3 seconds but not more than 4sec.
- After that the LED on WithCare+ device will start blinking with Orange color.
- After few seconds the gluco-meter will show the message PC. This indicates communication is established and everything is going well. Please continue to wait.
  After some time LED on withcare+ device will start blinking with Green colour. This indicates that the mobile phone received the getting records by Bluetooth from the gluco-meter device.

    - Then finally
        - The LED on the device will remain in Green colour for 5 seconds and will automatically shut-down
        - You should receive message on the mobile phone that message/sms has been sent successfully.

Possible errors:

- If blinking for 5 seconds in Red colour occurs before the shut-down of the WithCare+ device then probably it was never paired before and need to be paired first with some mobile phone.
- If LED on device will remain in Red colour for 5 seconds which indicates the error has occurred in transmission.
    - Then SMS was not sent successfully. In this case check your mobile phone have signal from the GSM network. Or maybe if you are as "Pay as you go customer" check that you have enough credit on the mobile phone.
    - Check that Bluetooth is enabled on your mobile.
    -
    - Check that you have started the WithCare+ app and pressed on Start button.
      If none of the above was the issue, then please contact the nurse. There many other possible error which may indicate that such as low battery on the gluco-meter or on the WithCare+ device.
- If for some reason you press the button for more than 5sec then you will see the blinking of RED and Green colour for 1min. which indicates you accidently enabled the Pairing process. Wait till blinking stops and device shuts down and try again. Nothing bad will happen. Just try to avoid pressing the button for more than 4seconds, because pairing process consume battery resources which are not unlimited.

Appendix D Code for Proof of Concept for Establishing Bluetooth Communication

The following are main steps for establishing Bluetooth Communication between Android phone and a Bluetooth enabled device.

1. Setting up Bluetooth

2. Finding available devices

3. Pairing the devices

4. Connecting the devices

5. Transferring data between the devices.

The classes and interfaces needed for accomplishing the above tasks are available in the 'android.bluetooth' package. These classes extend from 'java.lang.Object' class of Java API and their brief summary is given below:

BluetoothAdapter: This class is the entry-point for all Bluetooth communication and represents the local Bluetooth radio. It can be utilized to find available Bluetooth devices, query a paired device, instantiate a BluetoothDevice, and create a socket connection for listening.

BluetoothDevice: This class implements 'android.os.Parcelable' interface. It emulates the remote Bluetooth device and is used to request a connection to it. This class is used to find the information about the remote device such as its name and address.

BluetoothSocket: This class implements 'java.io.Closeable' interface. It acts as a connection point for data exchange between Bluetooth devices via OutputStream and InputStream.

BluetoothServerSocket: This class also implments ‘java.io.Closeable’ interface. It listens for incoming requests as it represents an open server socket and returns a connected BluetoothSocket on accepting the connection from a remote Bluetooth device.

BluetoothClass: This class implements ‘android.os.Parcelable’ interface. It describes general capabilities of a Bluetooth device but does not describe all services and Bluetooth profiles supported by the device reliably.

Bluetooth Permissions: For any kind of Bluetooth communication between devices it is essential to set up permissions in the manifest file of an application. A constant value represented by ‘android.permission.BLUETOOTH’ must be declared in the manifest file and the application must request Bluetooth permission in order to establish any communication such as requesting or accepting a connection, and transferring data. In order to initiate device discovery or manipulate Bluetooth settings, an application must request ‘android.permission. BLUETOOTH_ADMIN’ and it should be declared in the manifest file of the application as well.

1 Setting up Bluetooth

The first step for an application to set up Bluetooth communication is to ensure that the device supports the Bluetooth and if it does, then it should make sure it is enabled. If the device does not support Bluetooth, in that case the application should gracefully exit. However, if Bluetooth is supported but disabled, then the application should request the user to enable Bluetooth without leaving the application.

This setup is done in two steps, using the getDefaultAdapter() method of BluetoothAdapter class. The code snippet for this operation is shown below:

```
BluetoothAdapter mBluetoothAdapter= BluetoothAdapter.getDefaultAdapter( );

If (mBluetoothAdapter == null) {
```

// if the object value is null, the device does not support Bluetooth.

}

getDefaultAdapter() is a static method and returns a handle to the Bluetooth of the remote device if it exists but it returns a null if the device does not support Bluetooth and in this case the application should gracefully exit.

2 Enable Bluetooth

This step is accomplished by calling isEnabled() method of the BluetoothAdapter handle returned in step 1. If the return value of this method is false, then Bluetooth is not enabled. This can be enabled by calling startActivityForResult() method with the ACTION_REQUEST_ENABLE action Intent. This will request to enable Bluetooth through the system settings without stopping the application. The code snippet for this action is below:

```
if (!mBluetoothAdapter.isEnabled()) {
    Intent enableBtIntent = new
Intent(BluetoothAdapter.ACTION_REQUEST_ENABLE);
    startActivityForResult(enableBtIntent, REQUEST_ENABLE_BT);
}
```

The above code snippet will create a dialog requesting the user to enable Bluetooth as shown in figure 1.

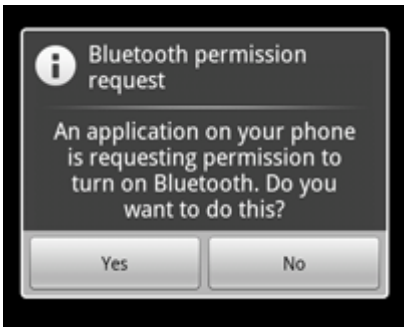


Figure 1 Bluetooth Permission request

If the user clicks on the 'Yes' button, the system will start enabling Bluetooth and focus will return to the application after the process is completed either with success or failure. In the code snippet above, REQUEST_ENABLE_BT (a locally defined constant greater than 0) is passed to startActivityForResult() method , that the system passes back to onActivityResult() callback method as the requestCode parameter. The value of this parameter will be RESULT_OK if the user clicks on the 'Yes' button or RESULT_CANCELLED if the user clicks on the 'No' button.

The application can also detect changes made to the Bluetooth state while it is running by listening to the ACTION_STATE_CHANGED broadcast Intent, which will be broadcast by the system whenever the Bluetooth state changes. There are two fields EXTRA_STATE and EXTRA_PREVIOUS_STATE, which contain the new and old Bluetooth states, respectively. The possible values for these two fields are STATE_TURNING_ON, STATE_ON, STATE_TURNING_OFF and STATE_OFF.

3 Finding Devices

The BluetoothAdapter class can be used through device querying paired devices or the discovery method for finding the available remote Bluetooth devices. In the device discovery method, the application searches the local area for Bluetooth enabled devices and the devices which are currently enabled to be discoverable respond back to the discovery request by sharing some information such as their name and unique media access control (MAC) address. The application can then choose to initiate a connection to the discovered device. If the connection is made with a remote device for the first time, a pairing request is automatically presented to the user. The basic information such as name and MAC address of the remote device is saved by the application after a device is paired and can be retrieved by Bluetooth APIs. The saved MAC address of the remote device can be used by the application

to initiate connection at any time in the future without performing the device discovery step again (assuming the remote device is within the Bluetooth range). There is a difference between being paired and being connected. When two devices are paired that means they are aware of the existence of each other and are ready to establish an encrypted connection with each other, Whereas the connected devices share an RFCOMM channel and can send data to each other. The next section describes how to find already paired devices using the Bluetooth API.

4 Querying Paired Devices

The BluetoothAdapter has a getBondedDevices method which can be used to query the paired devices. This method when called returns a Set of BluetoothDevices representing paired devices. The code snippet below shows how the application can use this method to get a Set of BluetoothDevices representing paired devices and then show the name of each device to the user, using an ArrayAdapter object.

```
Set<BluetoothDevice> pairedDevices = mBluetoothAdapter.getBondedDevices();

// If there are paired devices

if (pairedDevices.size() > 0) {

    // Loop through paired devices

    for (BluetoothDevice device : pairedDevices) {

        // Add the name and address to an array adapter to show in a ListView

        mArrayAdapter.add(device.getName() + "\n" + device.getAddress());

    }

}
```

The MAC address of remote device can be later used to initiate the connection to it, which will be explained in the next section.

5 Connecting Devices

The two devices are considered connected to each other when they each have a connected BluetoothSocket on the same RFCOMM channel. The connection with a remote device can be initiated with a BluetoothDevice object that represents the remote section as explained in the Finding Devices section. There are two steps involved in establishing the connection.

1. Call the method createRfcommSocketToServiceRecord(UUID) of BluetoothDevice to acquire a BluetoothSocket and initiate the connection. The UUID (universal unique identifier) is a string which can be generated online by any of the available engines. The same UUID should be used both on the server and the client for the application.

2. Call connect () method. This call will perform an service discovery protocol (SDP) lookup on the remote device in order to match the UUID. The RFCOMM channel will be shared if the lookup is successful and the remote device accepts the connection and the connect () method will return. This step should always be performed in a thread safe manner as if for any reason the method call fails or times out after 12 seconds, it will throw an exception.

6 Managing a Connection

Now the devices are connected and the data can be transferred between the devices using a BluetoothSocket. The methods of java.io package can be used to transfer data as below:

1. getInputStream and getOutputStream methods can be used to get InputStream and OutPutStream objects to transfer the data.

2. The data is then written using the read (byte [ ]) and write (byte[ ]) methods. The read and write methods should be used in the dedicated threads as they are blocking methods.

The application developed using this procedure was installed on an Android phone and data was written to a PSoC development kit to prove that it is possible to establish the communication over the RFCOMM if the SPP is blocked by the operating system. The following code snippet shows how data can be written to a remote device.

```
/* Call this from the main activity to send data to the remote device */

   public void write(byte[] bytes) {

      try {

         mmOutStream.write(bytes);

      } catch (IOException e) { }

   }
```