UNIVERSITY OF CALIFORNIA, IRVINE

Secure ECG-Based Biometric Authentication in Body Area Sensor Networks

THESIS

submitted in partial satisfaction of the requirements for the degree of

MASTERS OF SCIENCE

in Computer Science

by

Farshad Momtaz

Dissertation Committee:
Professor Tony Givargis, Chair
Professor Isaac D. Scherson
Assistant Professor Amir AghaKouchak

2016
DEDICATION

I dedicate this work to my parents who have supported me and my dreams and made me who I am today.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>viii</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td>ix</td>
</tr>
<tr>
<td>ABSTRACT OF THE DISSERTATION</td>
<td>x</td>
</tr>
</tbody>
</table>

1 Introduction  1

2 Body Area Networks  4
  2.1 Applications  4
  2.2 Challenges  6
  2.3 Hardware  7
    2.3.1 Node Architecture  8
  2.4 BAN Communication  9
    2.4.1 Communication Architecture  10
  2.5 Security and Privacy  11
    2.5.1 ECG Based Authentication  12

3 Problem Statement and Methodology  16

4 Sensor Hardware Interface  18
  4.1 Sensors  18
  4.2 Sensor board  19
  4.3 Results  20

5 Digital Signal Processing  21
  5.1 Digital Filtering  21
  5.2 ECG Feature Detection  22
  5.3 Model-based Data Validation  23
  5.4 Matlab Implementation  25
    5.4.1 Setup  25
    5.4.2 Results  26
# Embedded System Implementation

6.1 Target Platform ......................................................... 28
6.2 Simulink ................................................................. 29
6.3 C-Implementation ....................................................... 31
   6.3.1 Filter implementation in C ................................... 31
   6.3.2 Sampling Rate .................................................... 32
   6.3.3 Results and Evaluation ......................................... 33

# Protocol

7.1 Authentication Protocol ............................................... 36
7.2 Parameters of the Authentication Protocol .......................... 38
7.3 Analysis ......................................................................... 40
   7.3.1 False Negatives Authentications (FNA) ....................... 41
   7.3.2 False Positive Authentication (FPA) ........................... 42
7.4 Parameter Selection and Sensitivity Analysis ....................... 43

# Conclusion

8. Conclusion ..................................................................... 47

# Bibliography

Bibliography ....................................................................... 49
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A Body Area network comprises sensors attached to a human body. ECG data can ensure that sensors, attached to the same body (A,B,C) trust each other, but do not trust sensors (E) and devices (D) that are not attached to the same body.</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>Examples of Body Area Network devices on a human body</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>BAN Node Architecture [9]</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>Two-tier BAN communication architecture. (A) presents usage of an ad hoc network and (B) presents usage of preexisting network infrastructure for extra-body communication</td>
<td>10</td>
</tr>
<tr>
<td>2.4</td>
<td>Characteristics of the heart signal and Inter-Pulse-Intervals (IPI) between peaks.</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Development flow.</td>
<td>17</td>
</tr>
<tr>
<td>4.1</td>
<td>Schematic of our sensor processing board: The instrumentation amplifier obtains the difference of the sensor inputs, before the single signal is filtered and amplified.</td>
<td>19</td>
</tr>
<tr>
<td>4.2</td>
<td>5 Seconds of gathered sensor data after processing on our sensor board. Visible are the ECG characteristics, but also some residing high frequency artifacts.</td>
<td>20</td>
</tr>
<tr>
<td>5.1</td>
<td>Steps of the digital processing: the ECG needs to be filtered, peaks are detected and validated. Output is a table of QRS values.</td>
<td>22</td>
</tr>
<tr>
<td>5.2</td>
<td>Signal and detected QRS peaks for two QRS complexes after PTA. Note that a third erroneous QRS complex is detected at time stamp 4.8s.</td>
<td>24</td>
</tr>
<tr>
<td>6.1</td>
<td>Photo of our experimental setup: Electrodes with sensor board in front, the RaspberryPi board at the back.</td>
<td>29</td>
</tr>
<tr>
<td>6.2</td>
<td>Block diagram of the Simulink Model. (A) Data acquisition, (B) Data conversion, (C) Output, (D) Lowpass filter, (E) Pan-Tompkins QRS detection, and F Time tracker.</td>
<td>30</td>
</tr>
<tr>
<td>6.3</td>
<td>Sampling rate trade-offs: (A) memory to error rate, and (B) processing time.</td>
<td>32</td>
</tr>
<tr>
<td>6.4</td>
<td>Distribution of (A) measured IPIs, (B) difference between two adjacent measurements, and (C) measurement errors between sensors</td>
<td>34</td>
</tr>
</tbody>
</table>
7.1 Message sequence chart for the biometric authentication protocol between Sensor nodes $S_1$ and $S_2$, with processing steps for $S_1$. The authentication is successful if the last two steps succeed. 

7.2 Rejected authentications (false negatives) for measured truthful authentication attempts, (A) for different number of samples, and (B) for different dynamic ranges.

7.3 Accepted authentications (false positives) for forged authentication attempts, (A) for different number of samples, and (B) for different dynamic ranges.

7.4 Impact of increased variance (e.g. more noise) in the basis measurements on the allowed thresholds to limit false positives and false negatives.
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Example of BAN sensors [9]</td>
<td>8</td>
</tr>
<tr>
<td>5.1</td>
<td>Results for two sensors of the Matlab implementation on the PC.</td>
<td>26</td>
</tr>
<tr>
<td>6.1</td>
<td>Results for two sensors of the Simulink implementation on the BAN.</td>
<td>31</td>
</tr>
<tr>
<td>6.2</td>
<td>Results for two sensors of the C implementation on the BAN.</td>
<td>34</td>
</tr>
<tr>
<td>7.1</td>
<td>Quantization per sample</td>
<td>39</td>
</tr>
<tr>
<td>7.2</td>
<td>Example signature sizes</td>
<td>39</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to express my deepest appreciation for my committee chair, Professor Tony Givargis, who has helped me throughout my undergraduate and graduate programs and throughout the preparation of this thesis. Without his guidance, this thesis would not be possible. I am grateful to my committee members, Professor Isaac D. Scherson and Professor Amir AghaKouchak, for the time and work they devoted to reviewing this thesis. I am thankful to Steffen Peter for his help and guidance throughout my research and preparation of this work. I would also like to thank Bhanu Pratap Reddy for his contributions to this research.

I thank MDPI journal for their permission to share the copyrighted material of the paper *Design of Secure ECG-Based Biometric Authentication in Body Area Sensor Networks*. This work was supported in part by the National Science Foundation under NSF grant number 1136146.
CURRICULUM VITAE

Farshad Momtaz

EDUCATION

Masters of Science in Computer Science  
University of California, Irvine  
Irvine, CA  
June 2016

Bachelor of Science in Computer Science and Engineering  
University of California, Irvine  
Irvine, CA  
June 2015

Associate in Arts in Natural Sciences and Mathematics  
Irvine Valley College  
Irvine, CA  
June 2012

RESEARCH EXPERIENCE

Graduate Research Assistant  
University of California, Irvine  
Irvine, California  
September 2015 – June 2016

Undergraduate Research Assistant  
University of California, Irvine  
Irvine, California  
September 2013 – June 2015

WORK EXPERIENCE

Software Engineering Intern  
UnitedHealth Group - Optum Technology  
Irvine, California  
June 2015 – August 2015

Teaching Assistant  
University of California, Irvine  
Irvine, California  
March 2015 – June 2015

Software Engineering Intern  
Toshiba America Electronic Components  
Irvine, California  
April 2014 – September 2014

Microsoft Student Partner  
Microsoft  
Irvine, California  
September 2013 – June 2014

Software Engineering Intern  
Jet Propulsion Laboratory - NASA  
Pasadena, California  
June 2013 – September 2013
ABSTRACT OF THE DISSERTATION

Secure ECG-Based Biometric Authentication in Body Area Sensor Networks

By

Farshad Momtaz

Masters of Science in Computer Science
University of California, Irvine, 2016

Professor Tony Givargis, Chair

Body Area Sensor Networks (BANs) utilize wireless communicating sensor nodes attached to a human body for convenience, safety, and health applications. Physiological characteristics of the body, such as the heart rate or ECG signals, are promising means to simplify the setup process and to improve security of BANs. This thesis describes the design and implementation steps required to realize an ECG-based authentication protocol to identify sensor nodes attached to the same human body. This thesis starts by reviewing the latest literature in the field of Body Area Network, and looks at possible applications, challenges, and overall architecture of BANs. It also addresses the design of a body-area sensor system, including the hardware setup, analogue and digital signal processing, and required ECG feature detection techniques. A model-based design flow is applied, and strengths and limitations of each design step are discussed. Real-world measured data originating from the implemented sensor system then are used to set up and parametrize a novel physiological authentication protocol for BANs. The authentication protocol utilizes statistical properties of expected and detected deviations to limit the number of false positive and false negative authentication attempts. The result of the described holistic design effort is the first practical implementation of biometric authentication in BANs that reflects timing and data uncertainties in the physical and cyber parts of the system.
Chapter 1

Introduction

Technological Advancements in communication technologies, low power sensors, and chip design have led to an increase in the development of wearable and implantable technologies. The need to retrieve the collected information and communicate between these devices has resulted in the development of Body Area Networks (BANs) or Body Sensor Networks (BSNs) [9]. BAN devices have vast applications in health care, fitness tracking, vehicle-area-networks, and military operations. Such medical and safety related BAN applications call for a high level of access control and data protection [24, 46, 14, 44]. However, the goal of good security in BANs is challenged by the capabilities of typical body area sensor nodes. For economical and practical reasons the nodes are small and resource-constrained, providing only limited computation power and memory.

While security protocols and implementations exist to protect data on severely constrained devices [38, 23], the question that remains is, how devices that belong to the same body area identify and trust each other. Figure 1.1 illustrates the problem: Sensors that are attached to one person (Sensor A, B, and C) should know and trust each other, while sensors attached to other persons (E) or entirely forged data (D) are not trusted. Solutions like pre-deployed
Figure 1.1: A Body Area network comprises sensors attached to a human body. ECG data can ensure that sensors, attached to the same body (A,B,C) trust each other, but do not trust sensors (E) and devices (D) that are not attached to the same body.

keys [50] or manual setups are cumbersome and error-prone – in particular in environments with several interfering BANs.

The work in this thesis addresses the challenge of identifying nodes that are physically attached to the same human body. This mechanism can be used:

- for fast and convenient setup of a BAN, for instance for fitness trackers, chest sensors, and smart wrist bands,
- for the setup of a trusted and secure body area environment with a shared key, and
- as a second authentication factor in BANs with critical implanted medical equipment, to prevent accidental or malicious erroneous access to the medical devices.

As physiological characteristic in this thesis, we use Electrocardiography (ECG) data. ECG records the electrical activity of the heart and is characteristic of a person at a given time. ECG and related heart data can be obtained from sensors that are attached to the body, even locally as shown in [16, 48, 3, 47]. Literature has already discussed ECG-based authentication and key-agreement protocols [57, 55, 52, 40] without, however, considering practical
implications of the low cost sensors and resource-constrained BAN platforms. Instead, existing work used clinical ECG data obtained from medical databases and processed the data on PCs, ignoring uncertainties originating from sensors and processing.

In this thesis we present a biometric authentication protocol that intrinsically reflects the statistical properties of the uncertainties, to systematically balance the risk of false rejected authentications and false accepted attempts. We address these issues in two stages:

1. Design and implementation of a sensor platform (Chapter 4) including suitable data processing and feature detection methods. We apply a model-based design flow [21], starting with an analytical model in Matlab, test it in real-time models on a PC (Chapter 5), and finally translate it to the embedded system platform, taking into account the limited resources of a BAN in Chapter 6.

2. Based on empirical data gathered from the implemented sensor node, we design and parametrize a secure session establishment protocol in Chapter 7. We show that statistical properties of the uncertainties of the system can be harnessed to improve the confidence within the authentication process.

The contribution of our work is the connection of the implementation results and the parametrization of the security protocol. We show that uncertainty in the measurements can be handled, but needs to be reflected in the security protocol to avoid a high number of false rejections or false authentications.

The result is the first biometric authentication protocol that works on actual BAN nodes. The presented system shows 100% correct authentications with a probability for a successful attack of less than 0.1%. We present the hardware design, the algorithm, and software implementation, and discuss the selected security protocol parameters.
Chapter 2

Body Area Networks

Body Area Sensor Networks are a promising technology for convenience, safety, and health applications [34]. BANs consist of small wearable or in-body implanted devices that use sensors and actuators to monitor and interact with the human body. Examples of BANs include fitness trackers, smart glasses [10], vital tracking of emergency response teams [39], and medical implantable devices such as heart pacemakers and insulin pumps. Figure 2.1 shows examples of BAN devices on a human body. In this chapter, we will explore applications, challenges, hardware and network architecture, and privacy and security aspects of Body Area Networks.

2.1 Applications

Due to the desirable properties of BANs, such as portability, low-power needs, and ease of use, they have a wide variety of applications. Health care, fitness tracking, and military operations are just some of the applications of BANs.

One of the biggest applications of BANs is their usage in the health and medical industry.
With the aging population and the growing cost of health care, BANs are a great alternative for the next generation medical devices. MHealth is defined as mobile health care practice that relies on electronic processes and communication [20]. This generation of medical devices can be implanted in or on the patients’ body, giving medical professionals a more complete overview of patients’ health. Due to extreme portability of BANs, patients have the freedom to carry out their daily tasks without the need to be hospitalized. The patients’ data can be analyzed locally using automated algorithms or by sharing it with medical professionals over the internet.

Fitness trackers are commercially available BAN devices in the market. The recent increase in wearable fitness trackers has encouraged companies such as Microsoft, Apple, Fitbit, and Garmin to manufacture powerful sensor rich devices with a long battery life. These devices can be equipped with heart rate, accelerometer, gyrometer, GPS, skin temperature, UV, and microphone sensors. By collecting and analyzing information about users’ daily activities, these devices are able to provide users with an overview of their activities, such as steps taken, calories burnt, and hours slept. Fitness trackers usually rely on their mobile companion application to analyze data and interact with the user.
Another application of BANs is military operations. BAN devices can use GPS, camera, and physical sensors to collect information about the current state of soldiers. The main role of BAN devices in military is “(a) ensuring that adequate water is delivered and consumed, (b) reducing the likelihood of body harm attributed to harsh environmental conditions, such as heat stroke, and (c) improving the quality of medical care in the event of an injury” [9]. Information plays an important role in battlefields. The data collected from each individual soldier is transferred through a mesh network, and then transmitted to Technical Operations Center (TOC) or medical assistance unit only after it reaches a device with access to the Wireless Local Area Network (WLAN) or cellular network [17]. The transmitted information can notify the authorities about possible hazards and injuries.

2.2 Challenges

Traditional Wireless Sensor Networks (WSNs) have been well studied and numerous protocols and algorithms have been proposed [2]; however, protocols and algorithms are not well studied for BAN devices most of which have limited resources, such as power, computation, and networking. Compared to WSNs, BANs face the following challenges:

- Most BAN devices have limited power resources because of their small form factors. These devices are often smaller than $1cm^3$ [15], which limits the space required for batteries. This limitation requires BANs to conserve and harvest energy through their surroundings.

- The devices are placed on the human body for a specific application; therefore redundant nodes are not available.

- Unlike WSN, the BAN nodes are not stationary and they require to be light and portable. Moreover, they have the be resistant to movement and ever changing network
topology.

- In medical applications, timing is an important factor; therefore, the nodes should communicate with low latency. Moreover, the devices are all equally important and no device should starve from lack of communication.

- To minimize interference and comply with health requirements [19], an extremely low transmit power should be used for communication.

- The small form factor and power limitations have left nodes with limited computation resources. In order to cope with this problem, algorithms and protocols used should be well analyzed and optimized.

### 2.3 Hardware

Body Area Network Nodes can be divided into three different categories [25]:

- **Sensor Nodes**: The main function of the sensor nodes is to sample and transmit information. The sampled data is usually analyzed locally and a summary of the readings is transmitted using the BAN [1]. Temperature, ECG, and blood pressure are some of the examples of the sensor nodes.

- **Actuator Nodes**: These nodes are in charge of taking action based on the sampled data or interaction with the user. Examples of such nodes are insulin pumps and medicine dispensers.

- **Personal Devices**: Unlike sensor and actuator nodes, personal devices usually have powerful computing unit and access to the WLAN and cellular network. These devices are in charge of collecting information. The collected information is either locally analyzed and stored or stored and analyzed in a remote location using the internet.
Sensor Data Rates

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Data Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accelerometer/gyroscope</td>
<td>High</td>
</tr>
<tr>
<td>Blood glucose</td>
<td>High</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Low</td>
</tr>
<tr>
<td>CO₂ gas sensor</td>
<td>Very low</td>
</tr>
<tr>
<td>ECG sensor</td>
<td>High</td>
</tr>
<tr>
<td>EEG sensor</td>
<td>High</td>
</tr>
<tr>
<td>EMG sensor</td>
<td>Very High</td>
</tr>
<tr>
<td>Pulse oximetry</td>
<td>Low</td>
</tr>
<tr>
<td>Humidity</td>
<td>Very low</td>
</tr>
<tr>
<td>Temperature</td>
<td>Very low</td>
</tr>
<tr>
<td>Image/video sensor</td>
<td>Very high</td>
</tr>
</tbody>
</table>

Table 2.1: Example of BAN sensors [9]

2.3.1 Node Architecture

Sensor and actuator nodes have a common architecture, which is shown in Figure 2.2. The sensor node architecture has four main modules: microprocessor, Radio Frequency (RF), sensor, and memory module, which are all connected together using a bus. The sensor module consists of a sensor, a filter, and an Analog-to-Digital Converter (ADC). The sensor transforms physical energy into analog electricity which is later filtered and sampled using the built-in ADC. Table 2.1 presents a list of commonly used sensors and their data rates. The sampled data is sent to the microprocessor module, which further analyzes it locally; then the data is either sent to the RF module to be transmitted or to the memory module to be stored. The memory module consists of Random Access Memory (RAM), Read Only Memory (ROM), and Flash. The RF module is in charge of transmitting data to the Wireless Body Area Network (WBAN) or the Wireless Personal Area Network (WPAN).
2.4 BAN Communication

Communication plays an important role in BANs; it provides nodes with a pathway to transmit and receive data and instructions. Therefore, continuous communication and quality of service is an important factor of Media Access Control (MAC) Layer protocols. Moreover, the proposed protocols have to be power efficient in order to save energy. One common technique used in WSNs is to constantly monitor radio transmissions to receive data or instructions. This method would insure constant communication and high quality of service but would result into high power usage. Another approach used in T-MAC [51], S-MAC [56], and TRAMA [42] suggests periodically toggling the radio on and off to conserve energy. In this approach the nodes would turn their radio on and only transmit information during a synchronized transmission schedule. In order to synchronize the transmission schedule between the BAN nodes, H-MAC [26] suggests using the heartbeat rhythm, which is the same throughout the body. On the other hand, B-MAC and WiseMAC use the Low-Power Listening (LPL) approach and use channel pulling to check whether a node needs to wake up in order to transmit or receive information. This method conserves energy by reducing the idle receiving time.
Figure 2.3: Two-tier BAN communication architecture. (A) presents usage of an ad hoc network and (B) presents usage of preexisting network infrastructure for extra-body communication.

### 2.4.1 Communication Architecture

BAN communication architecture can be separated into two tiers: Intra-Body Communication and Extra-Body Communication [25]. Figure 2.3 shows an example of this two-tier architecture.

The Intra-Body Communication consists of a network of nodes connected by wires or wirelessly. Using wires as a communication medium results in a higher quality of service and better conservation of energy. However, a wireless network is more cost efficient and enables a larger array of applications [11]. The wireless connection usually has a coverage range of 2 meters around the human body; However, one-hop communication between the nodes on the body cannot always be guaranteed. In [45], researchers document a lack of communication between nodes placed on the chest and back of a patient when using TelosB mote, an open source research platform for low power computation within WSN. The placement
of sensors and the distance of the nodes’ antenna from the body can highly affect the range of communication [43]. Multi-hop communication between nodes can be used to increase the quality and reliability of communication. Moreover, using a hybrid approach, the data can be sent to an intermediate processor node using wired or wireless connections for collection and pre-processing. The data collected by devices within Intra-Body Communication is usually transferred to a personal or centralized storage database using the Extra-Body Communication.

The main function of Extra-Body Communication is to collect data from the Intra-Body network, transfer it to be store it, and shared with medical professionals or other individuals. This tier enables a wide array of E-healthcare applications. The data collected by nodes in the Intra-Body Communication network can be collected by the Extra-Body Communication network using either the preexisting network infrastructure or an ad hoc network [9]. The data can be transferred to a central database using a personal device that is connected to the cellular network or WLAN. On the other hand, the ad hoc network allows for direct data collection without the need to retrieve data from a personal or central storage database. This network consists of receiver nodes that pass information from Intra-Body nodes to the local infrastructure. For example, the ad hoc network can be deployed in emergency rooms to collect node information in case of an emergency. The ad hoc network would decrease the retrieval time of the information but, unlike the WLAN or cellular network, has much smaller coverage.

2.5 Security and Privacy

The privacy and security of information play an important role in BANs. Data security is defined as secure transfer and storage of the collected data, and data privacy is defined as the sharing of collected data with the devices and people who have access permission
Privacy and security breaches occur when the data is created, transferred, stored, and processed. During each of these steps the data security and privacy should be insured. One of the main challenges of providing data security and privacy is balancing between security, privacy, safety, ease of use, usability, and latency. In other words, increase in security and privacy, such as use of encryption, would reduce usability and increase latency. For example, an unconscious patient’s life can be endangered if he/she is in a time sensitive emergency and is unable to grant permission to the medical professionals to access the information collected by his/her BAN.

2.5.1 ECG Based Authentication

A sub domain of BAN security and privacy is the application of biometric and physiological body properties as means to establish authentication or to generate keys in such BANs. ECG basics are explained in the following subsection.

2.5.1.1 ECG Basics

Electrocardiography (ECG) is defined as the process of recording the electrical activity of the heart over a period of time using electrodes placed on a person’s body. These electrodes detect the tiny electrical changes on the skin that arise from the heart muscle depolarizing.
during each heartbeat [34].

A schematic trace of two heart beats is shown in Figure 2.4. Dominant characteristics of the signal are the five peaks, called P, Q, R, S, and T, while the most significant peak is the R peak. All the information provided by ECG exists mainly between 0.05 Hz to 100Hz, since the length of a QRS complex is typically between 0.06 and 0.1s [22]. While characteristics like the Q-to-R or R-to-S timing might be applied for authentication purposes, in this paper we focus on the Inter-Pulse-Intervals (IPI). As shown in Figure 2.4, the IPI can be measured between two adjacent Q peaks (Q-IPI), R peaks (R-IPI) or S peaks (S-IPI). Since the typical heart-rate (HR) of a human varies between 30 and 240 beats per minute (bpm) [22], the IPI ranges between 250ms and 2000ms (IPI=60sec/HR).

While the occurrence of the heart-beats is not simultaneous at each location at the body, the IPIs for a person are approximately equivalent regardless of the measured location. Therefore the IPI is well suited for an application in BANs. The IPI is particularly interesting as it might be measured not only by ECG sensors but by optical [48], acoustic [3] or tissue pressure [47] sensors and other BAN devices.

2.5.1.2 Related Work

ECG- and IPI- based authentication received significant research attention [55, 57, 52, 18, 54]. For instance, ESKE [55] is a noise-tolerant key generation scheme that works without pre-deployment of key material. ESKE applies wavelet filtering and requires a sample size of more than 30s for a single authentication, which is unsuitable for most BAN scenarios.

Another key agreement scheme [57] is based on the Fuzzy Vault Scheme [13]. The method allows a receiver to reconstruct a message if most coefficients of the encryption polynomial are known. A similar approach is pursued in the physiological signal based key agreement (PSKA) scheme [52], which creates a session key from the frequency information of the ECG
data. However, PSKA still needs the receiver and sender to share a set of exact equivalent key values, which cannot be guaranteed in most BANs. These works are all executed on a PC and frameworks like Matlab. In addition, the methods require long sampling periods to avoid offline guessing of the key. Our work instead separates key agreement and authentication. This way we apply traditional, well established key agreement methods, such as Diffie-Hellman as basis, and use physiological authentication as additional step.

OPFKA [18] is a physiological-feature-based key agreement protocol that was actually implemented on a BAN platform. However, OPFKA also has only been tested with medical databases, such as the MIT-BIH Arrhythmia database [33], as source for the reference ECG signals. These works do not discuss the impact of sensing, measurement, and timing errors. Our work clearly shows the importance of coping and harnessing the measurement uncertainties. In fact, none of the presented schemes worked with the real data we gathered from our implemented sensor node.

The feasibility of IPI-based authentication for BSNs was shown in [8] and [40]. Poon [40] demonstrated the applicability of IPI-based authentication even for different types of sensors. Their experiments showed the interoperability between ECG and pulse oximeter (PPG) sensors. The work also demonstrated the suitability of the IPI authentication for older and less healthy persons. Contrary to our work, [8] and [40] apply binary codes that requires a high number of IPI values (>30) for an authentication. They also do not consider the impact of design decisions in the signal processing and embedded systems design.

The application of physiological characteristics has also been discussed for permanent user authentication, for instance for persistent storage in the cloud [7]. The idea is to utilize encryption keys that are based on permanent characteristics of the human physiology. While our work does not directly aim at permanent authentication, the implementation and design flow we describe can be applied to support permanent identification on BAN nodes in the future.
A range of alternative approaches have been presented to set up secure BANs and provide access and privacy protection [12]. As example, [30] applies near field communication to ensure close proximity of BAN nodes in medical applications. Other approaches include password-authenticated key exchange and agreement [50], wristbands, proximity sensors, or using the body as a shared communication medium [35]. We consider these approaches as possible second factors in a secure authentication process for BANs.
Chapter 3

Problem Statement and Methodology

The general idea of the biometric authentication protocol we introduce in Chapter 7 is to measure the IPIs for a certain time on different nodes. The nodes are considered to be on the same body if the measured IPIs are equivalent or very similar. The major challenges in this process are:

- uncertainty of the underlying physical (biological) phenomenon,
- timing uncertainty and jitter in the cyber-parts (sensors, interfaces, processing), and
- parametrization of the authentication protocol, to omit rejection of valid sensor pairings, but reduce the probability of invalid pairing attempts to succeed due to a high tolerance of deviation.

To address these challenges, in this paper we pursue a model-based design flow that gravitates around real sensor data, instead of an ideal library of data. The methodology and outline of our work is shown in Figure 3.1. As the first step we select hardware, including the sensors and their interfaces. The sensor access includes the design of a sensor board for analog filtering and signal pre-processing. Then, we select and implement the digital signal
processing steps and the IPI detection algorithm that works with the gathered data. For first practical testing, then we use the sensor board from the Matlab implementation of the signal processing steps, running on a PC, as discussed in Chapter 5. For the actual BAN implementation, we apply automatic generation of the system code from the Matlab/Simulink environment, and compared the performance to a manual implementation in C. The benefit of the Simulink approach is the seamless model-based design flow, while the main benefit of the C-implementation is its superior performance. Finally, we apply the sensor data, gathered on our prototype, to parametrize a secure biometric BAN authentication protocol. The result is a protocol that utilizes the properties of the measured sensor data to improve the confidence in the authentication.
Chapter 4

Sensor Hardware Interface

Since the goal of this work is to use actual sensor data for the authentication, as a first step we have to select suitable sensors, and interface the sensors to the computation nodes. One criteria at this point was not to chose clinical sensor and filter systems, which indeed are available [5] but due to cost and size are not applicable to BANs. Instead we used low-cost sensors and designed a sensor processing board, which then can be connected to the BAN nodes.

4.1 Sensors

As sensors we use conventional wet cloth electrodes with repositionable conductive adhesive hydrogel to measure the electrical activity from skin surface [31]. One main challenge of these sensors is their relatively low DC offset voltage, with raw signal amplitude below 0.5mV. Additional noise originates internally from the device but also by the environment. As a result, initial experiments with a plain A/D conversion failed due to the low voltage level, a fast fading signal, and the high noise.
Figure 4.1: Schematic of our sensor processing board: The instrumentation amplifier obtains the difference of the sensor inputs, before the single signal is filtered and amplified.

4.2 Sensor board

To extract, stabilize, and clean the signal we designed a sensor board that amplifies and filters the signals. We applied a standard difference amplifier approach [5]. The difference amplifier is a suitable solution since the basic ECG data is obtained as an output of the difference of two leads placed on the body. The block diagram of our sensor board is shown in Figure 4.1. The circuit consists of three parts: the differential amplifier, a filter, and a post amplifier. As differential amplifier we use an INA121 instrumentation amplifier, due to its high precision and high noise rejection and its sensitivity to the ECG input range. The capacitor (C1) at the output of the instrumentation amplifier stabilizes the signal by removing DC shifting. The output of the instrumentation amplifier is still noisy and contains many unwanted frequency components. Therefore we apply an operational amplifier (LM358) to filter the ECG-characteristic frequencies. After the filter stage, signal post amplification is carried out to match the input requirements of the processing device. In other words, the range of the output signal is determined by the next device that is going to use and process the output signal. For instance, the target system we introduce in Chapter 6 requires input levels in the range of 1mV to 10mV. Based on this requirement we use a simple inverting amplifier in which R4 and R5 are chosen appropriately for the required gain, given as \( G = -\frac{R_5}{R_4} \). The parameters can be further adapted to reflect the input requirements of the processing platform.
Figure 4.2: 5 Seconds of gathered sensor data after processing on our sensor board. Visible are the ECG characteristics, but also some residing high frequency artifacts.

4.3 Results

A trace for data gathered with the sensor board is shown in Figure 4.2. The board requires $6.0\mu W$ and has an output impedance of $1.2\, M\Omega$. In Figure 4.2 we can see that the output levels are clearly distinguishable and can be processed by the subsequent embedded processing system. In the trace, the PQRST wave peaks can be identified clearly. Notable is also the steadiness of the signal level, which is important for further processing. However, we still see some high frequency components as noise. That is why in the next step we investigate digital signal processing and available feature detection algorithms for their suitability to work with the gathered sensor data.
Chapter 5

Digital Signal Processing

This chapter discusses the digital processing of the gathered ECG signal. The process is shown in Figure 5.1. The input is the continuous sampled ECG signal, and the output is a list of R-, Q-, and S-IPIs. To process the data, the following general steps have to be executed:

1. digital low-pass filtering to clean the ECG signal,

2. detection of the ECG features (QRS peaks), and

3. validation and correction of obtained values, based on biometric model properties.

We describe details of the three steps in the following sections.

5.1 Digital Filtering

Even though the analog filters on our sensor board already filtered the signal, several high-frequency artifacts from wires, hardware and interfaces remain. As introduced in Section
Figure 5.1: Steps of the digital processing: the ECG needs to be filtered, peaks are detected and validated. Output is a table of QRS values.

2.5.1.1, all important information of the ECG signal are located between 0.05Hz to 100Hz. Such fixed cut-off frequencies can be efficiently digitally filtered using Infinite Impulse Response (IIR) butterworth filters [41]. The basic transfer function of an IIR filter is given as:

\[ H(z) = \frac{\sum_{i=0}^{P} b_i z^{-i}}{1 + \sum_{j=1}^{Q} a_j z^{-j}}, \]  

(5.1)

while \( a \) and \( b \) are multi-dimensional characteristic coefficients of the filter, and \( z \) is the signal to be filtered. The generation of the coefficients for the given the cutoff frequencies and the order of the filter is well described in related work [41]. The interested reader can find the coefficients and their generation in the Matlab and C code at [6]. The actual implementation of the filter depends on the capabilities of the underlying hardware and is suspect to quality-to-resource trade-offs as we will discuss in Section 6.3.

5.2 ECG Feature Detection

The filtered signal is processed further to extract the features of the ECG signal. These features, which are the timing of the Q, R, and S peaks, are the core identifiers for the intended authentication scheme described in Chapter 7. To extract ECG features a range of approaches have been proposed in related work [32], [36]. One implementation option is the detection of the large R-peaks. However, due to the lack of redundant information, the simple R-detection leads to many unrecoverable errors and large timing uncertainties.
We apply the Pan-Tomkins real-time QRS detection algorithm (PTA) [36]. PTA extracts the QRS complex from a given ECG signal and is suitable for resource-constrained devices. PTA is also considered as robust in presence of abnormal ECGs, such as arrhythmias [37]. PTA performs a sequence of filtering and comparison steps, including:

- a five-point derivative filtering to provide the slope information of the QRS complex, using the transfer function $H(z) = \frac{1}{5}(-z^{-2} - 2z^{-1} + 2z^{-1} + z^{2})$,
- squaring of the signal, to obtain all positive signal values and nonlinear amplification to emphasize the characteristic higher ECG frequencies,
- fixed moving window integration to obtain waveform feature information in addition to the slope of the R wave, and
- a comparison step to identify the largest peaks in a window to locate Q, R, and S.

The output of the PTA is a table of identified Q, R, and S time indices. The benefit of PTA is that each step can easily be implemented even on severely constrained embedded devices. However, one disadvantage of PTA is that the simplified computation model leads to erroneously detected peaks if applied to non-ideal input signals. One example of an erroneous QRS complex is shown after 4.8s in Figure 5.2. In this example PTA falsely identified another QRS complex just in the slope after the correct S peak. In fact, practical measurements showed errors in about 6% of our measured QRS complexes, which have to be addressed in the following model-based data validation step.

### 5.3 Model-based Data Validation

To reduce the impact of erroneously detected peaks, we exploit data redundancy and knowledge about the typical heart beat to identify and fix these errors. The general idea is to vali-
Algorithm 1 Model-based Validation on detected QRS peaks

Input:
arrays Q, R, S time and values \( \triangleright \) size of arrays might differ, may contain invalid values
Window Size = 0.1s

Output:
arrays \( Q', R', S' \) time and values \( \triangleright \) size(\( Q \))=size(\( R \))=size(\( S \))

1: \( \text{EndWindowTime} = 0; \)
2: \text{for all } \( r_{\text{peak}} \in R : \text{time}(r_{\text{peak}}) > \text{EndWindowTime} \text{ do} \)
3: \( \text{StartWindowTime} = \text{time}(r_{\text{peak}}) - \text{WindowSize}/2 \)
4: \( \text{EndWindowTime} = \text{time}(r_{\text{peak}}) + \text{WindowSize}/2 \)
5: \( \text{id}_R = \text{id}_Q = \text{id}_S = 0 \) \( \triangleright \) pointer to identified Q,R,S
6: \text{for all } \( r \in R : \text{StartWindowTime} \leq \text{time}(r) \leq \text{EndWindowTime} \text{ do} \)
7: \( \text{if } \text{id}_R = 0 \lor \text{value}(r) > \text{value}(\text{id}_R) \text{ then } \text{id}_R = r \)
8: \text{for all } \( q \in Q : \text{StartWindowTime} \leq \text{time}(q) \leq \text{EndWindowTime} \text{ do} \)
9: \( \text{if } \text{id}_Q = 0 \lor \text{value}(q) < \text{value}(\text{id}_Q) \text{ then } \text{id}_Q = q \)
10: \text{for all } \( s \in S : \text{time}(\text{id}_R) \leq \text{time}(s) \leq \text{EndWindowTime} \text{ do} \)
11: \( \text{if } \text{id}_S = 0 \lor \text{value}(s) < \text{value}(\text{id}_S) \text{ then } \text{id}_S = s \)
12: \( \text{if } \text{id}_Q \neq 0 \land \text{id}_R \neq 0 \land \text{id}_S \neq 0 \text{ then} \)
13: \( Q' \leftarrow = \text{id}_Q; R' \leftarrow = \text{id}_R; S' \leftarrow = \text{id}_S; \)
14: \text{return } Q', R', S'

date that the detected QRS values are in the expected QRS order, and that the magnitudes and time differences of the detected Q,R and S peaks are within the expected theoretical range for a normalized ECG signal. Since the duration of a QRS complex is expected to be between 0.06 and 0.1s, the distance between two consecutive Q, R, or S peaks must be greater than 0.06s as well. These basic rules lead to Algorithm 1.

The inputs to the validation algorithm are the Q, R, and S locations (time, value) delivered by PTA. The algorithm iterates through the locations of detected peaks, and copies valid QRS locations into the output arrays \( Q', R', \) and \( S' \). To identify a valid complex, a window of

Figure 5.2: Signal and detected QRS peaks for two QRS complexes after PTA. Note that a third erroneous QRS complex is detected at time stamp 4.8s.
0.1s centered at R-peaks \(r_{peak}\) (line 2-4) is moved over the input data. Within each window we identify the highest R (line 6-7), lowest Q (line 8-9), and lowest S (line 10-11). Only if valid Q, R, and S locations are found with a window range, they are added to the output arrays (line 12-13), otherwise the peaks are discarded. The result is an array of valid Q, R, and S time values from which the IPI values can be computed.

Applying the presented algorithm, the false peaks in the example in Figure 5.2 are removed because the false R appears within the 0.1s window after the correct R. The false R and its adjacent S then are discarded because their peaks are smaller than the correct ones, and the Q is discarded since it does not follow the expected pattern. In our tests, which we present next, the validation algorithm could fix 100% of the erroneously detected peaks.

### 5.4 Matlab Implementation

In this section we describe the setup and the results of a first prototype system that works with the presented sensor board but performs the digital processing steps in Matlab on a PC.

#### 5.4.1 Setup

For the Matlab implementation two steps need to be addressed: First, how to interface the sensor board, and second, how to execute the digital processing, discussed in the previous section.

To interface the sensor device with the PC we require an analog to digital conversion that is easily accessible from Matlab. One implementation option is an external AD converter. However, for our experiments we sampled the data via the microphone port of the PC, via
Table 5.1: Results for two sensors of the Matlab implementation on the PC.

<table>
<thead>
<tr>
<th>Sensor 1</th>
<th>1st IPI</th>
<th>2nd IPI</th>
<th>3rd IPI</th>
<th>4th IPI</th>
<th>5th IPI</th>
<th>6th IPI</th>
<th>7th IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0.927</td>
<td>0.908</td>
<td>0.880</td>
<td>0.864</td>
<td>0.799</td>
<td>0.828</td>
<td>0.774</td>
</tr>
<tr>
<td>Sensor 2</td>
<td>0.927</td>
<td>0.908</td>
<td>0.880</td>
<td>0.865</td>
<td>0.799</td>
<td>0.828</td>
<td>0.774</td>
</tr>
<tr>
<td>Sensor 1</td>
<td>0.926</td>
<td>0.909</td>
<td>0.880</td>
<td>0.865</td>
<td>0.798</td>
<td>0.829</td>
<td>0.774</td>
</tr>
<tr>
<td>Q</td>
<td>0.926</td>
<td>0.909</td>
<td>0.879</td>
<td>0.865</td>
<td>0.799</td>
<td>0.828</td>
<td>0.775</td>
</tr>
<tr>
<td>Sensor 2</td>
<td>0.926</td>
<td>0.908</td>
<td>0.880</td>
<td>0.865</td>
<td>0.799</td>
<td>0.828</td>
<td>0.774</td>
</tr>
<tr>
<td>Sensor 1</td>
<td>0.927</td>
<td>0.908</td>
<td>0.879</td>
<td>0.865</td>
<td>0.799</td>
<td>0.827</td>
<td>0.775</td>
</tr>
<tr>
<td>S</td>
<td>0.927</td>
<td>0.908</td>
<td>0.880</td>
<td>0.865</td>
<td>0.799</td>
<td>0.827</td>
<td>0.775</td>
</tr>
</tbody>
</table>

a standard audio jack. The Matlab function AudioRecorder provides direct and convenient access to the analog input data, which performs continuous sampling between 1kHz and 96kHz with a precision of 8bit to 24bit. For our experiment we used the settings 3kHz and 16bit.

Since Matlab has native support for designing filters, butterworth filters can easily be designed using in-built functions such as butter(). The function generates the coefficients that are utilized to generate the impulse response of the filter using impz(). With the generated filter response we can perform the filtering on any given signal using filtfilt(). Similarly, the steps of the Pan Tompkins algorithm are implemented applying the Matlab functions for convolutions and signal processing. Finally the data validation algorithm is directly implemented as the pseudo code shown in Algorithm 1.

5.4.2 Results

To assess the processing steps and the algorithms we connected the sensor board to a PC (i5, 4GB RAM) running Matlab. The results for two sensor systems tracking the Q-, R-, and S IPIs are shown in Table 5.1. In Table 5.1 we see that both sensor systems obtain approximately the same IPIs. Also the Q, R, and S IPIs match each other for a given time index. The results indicate the suitability of the sensor board and the data processing steps to be applied for the intended physiological authentication of sensor nodes attached to the same person. The computation time for processing the eight IPIs in Matlab on the PC is on
average 0.14s. The memory consumption is 28MB. To reduce the computation and memory footprint, we translate the digital processing steps from the PC to a BAN node in the next chapter.
Chapter 6

Embedded System Implementation

In this chapter we further follow the model-based design methodology and translate the digital ECG signal processing implementation from the PC-Matlab environment to an embedded system platform. The core challenge in this part is to cope with the limited interfaces and processing capabilities. We describe and compare the results for an automatically generated implementation from the Matlab-Simulink environment, and a manual implementation in C. Due to its superior performance, the latter one is the basis for the security protocol we describe in the next chapter.

6.1 Target Platform

As embedded target platform in this chapter we chose a Raspberry Pi (RPi). The RPi is a low power (0.7-1.2W) single-board computer with a 700MHz ARM processor and 512 MB of RAM. The RPi is one of the more powerful BAN nodes, and has already been applied in a range of practical BAN and e-health applications [49, 53]. One particular benefit of the RPi is it is supported by Matlab, Simulink and a range of design tools, which facilitates easy and
fast prototyping.

For our prototype we used the RPi and connected our sensor board via an external 16bit ADC (ADS1115) to the RPi’s GPIOs, resulting in an effective measurement resolution of $9 \mu V$, which proved to be sufficient for tracking the ECG signal. The setup of the RPi, the sensor board, and the sensors is shown in Figure 6.1. In the following we describe two approaches to implement the software on the system: first, automatically generated from Matlab/Simulink, and second, manually implemented in C.

6.2 Simulink

For the first implementation we further followed the model-based design methodology [21], by using the Matlab system to automatically generate C-code that can be compiled for the target system. We applied the Simulink coder [29], which is an industry grade tool that generates C and C++ code from a Simulink model. The Simulink coder helps the system designer by evaluating the model and block parameters, propagating signal widths and sample times, and determination of the execution order of blocks within the model.

However, the Simulink coder still requires the manual conversion of the Matlab model to a
block dataflow models in Simulink. Most parts of our system could be directly translated, since the applied filter functions and signal processing blocks are available in Simulink. The major challenge concerns the time measurement. The algorithm implemented on Simulink calculated QRS time indices only in its respective frames and not for the continuous recording over a period of time. Hence a timer was required within the model to track the signals in real time. In the Simulink system diagram, shown in Figure 6.2, the timing counter is highlighted as block F. The other parts of the design correspond to the data processing steps discussed in the previous section and are: A Data acquisition, B Data conversion, C Output, D Lowpass filter, E Pan-Tompkins QRS detection. The model is available from our project webpage [6].

For the generated code on our target platform we obtained the following results. The Q and R IPI results of the implementation for two sensor systems are shown in Table 6.1. Like the Matlab implementation (Table 5.1) we expected similar values for all measured IPIs. However, in Table 6.1 we see significant variations and errors. The main reason for the IPI mismatches are delays in data acquisition and over-utilization of the computation resources. In fact, due to the limited computation resources we had to disable the tracking of S peaks, in order to obtain at lest R and Q-IPIs. This result shows the existing limitation of model-based
design. The automatically translated Simulink implementation of the processing model is not suitable for the BAN node – even on the relatively powerful RPi node. Hence we need to translate the data processing steps into a lower level implementation language like C, manually, to address the limited resources of a BAN.

6.3 C-Implementation

The starting point of the manual ANSI C implementation is the generated Matlab code. Due to the use of model based design we were able produce C code with modest effort. The model based design helped us translate the logic implemented in Matlab and Simulink to C code. Nevertheless we had to reduce the processing overhead and memory consumption of the embedded Matlab functions with tailored lightweight C code. The two most taxing functions of the signal processing steps are filtering and data acquisition. Therefore, we discuss the filter implementation and the sampling rate trade-off in the following paragraphs.

6.3.1 Filter implementation in C

One of the most utilized functions in the data process flow is Matlab’s filter function (\texttt{filtfilt}). \texttt{filtfilt} is used for digital filtering of the input data and for various steps of the PTA. The function utilizes the defined filter characteristic given by the impulse response to shape the data. However, to achieve the same functionality we only need a convolution \((y(t) = \sum h(u)x(n-u))\) of the input signal \(x\) with the fixed impulse response \(h\). The fil-

<table>
<thead>
<tr>
<th>IPI in [s]</th>
<th>1st IPI</th>
<th>2nd IPI</th>
<th>3rd IPI</th>
<th>4th IPI</th>
<th>5th IPI</th>
<th>6th IPI</th>
<th>7th IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor 1 R</td>
<td>0.927</td>
<td>0.908</td>
<td>0.880</td>
<td>0.864</td>
<td>1.299</td>
<td>0.828</td>
<td>0.774</td>
</tr>
<tr>
<td>Sensor 2 R</td>
<td>0.99</td>
<td>0.7737</td>
<td>0.73</td>
<td>0.908</td>
<td>0.8617</td>
<td>1.053</td>
<td>0.977</td>
</tr>
<tr>
<td>Sensor 1 Q</td>
<td>0.926</td>
<td>0.909</td>
<td>0.880</td>
<td>0.865</td>
<td>0.798</td>
<td>0.829</td>
<td>0.774</td>
</tr>
<tr>
<td>Sensor 2 Q</td>
<td>1.1723</td>
<td>0.828</td>
<td>1.0</td>
<td>1.086</td>
<td>0.9137</td>
<td>0.994</td>
<td>0.9873</td>
</tr>
</tbody>
</table>
Figure 6.3: Sampling rate trade-offs: (A) memory to error rate, and (B) processing time.

The data processing steps and the PTA require the input signal for a fixed time of several seconds. Accordingly, the samples for this time period have to be stored before the data can be processed. The memory required for storing the data is determined by the sampling rate of the signal, and it affects the total memory consumption significantly. For example, for one second of ECG signal data with a sample rate of 330, an input buffer of at least 1320 bytes is needed, if double precision data types are used. Moreover since the input signal has to be manipulated and filtered all through the algorithm the sampling rate directly affect the processing time and power. However, a low sampling rate might affect the processing quality negatively. In order to identify a preferable sampling rate and investigate the effect on memory consumption and the processing error we applied the algorithm on signals with different sampling rates, ranging from 150 to 2000 samples/sec. The error factor in the experiment is the squared timing error of the identified Q, R, and S peaks.
The result for processing seven seconds of ECG data is shown in Figure 6.3 (A) and (B). The results shown in Figure 6.3 (A) confirm that the error factor decreases with an increase in sampling rate, while the memory usage increases. While the memory consumption grows linearly with an increased sampling rate, the error rate decreases slower with increased rates. As a result, increasing the sampling rate beyond 500 Hz only improves the processing marginally. Hence for our implementation we used a sampling rate of 800 Hz and required a data memory amount of approximately 10 kB per second. As we will discuss in Section 7.4, the sampling rate of 800 Hz is preferable for the proposed authentication protocol due to its low deviation of sensor readings.

Figure 6.3 (B) shows the computation efforts on our target platform for the different sampling rates for processing the data. It can be seen that the processing efforts increase about linearly with higher sampling rates, from 23 ms at 150 Hz sampling rate to 280 ms at 2000 Hz. As we will see in Chapter 7 the signal processing is the dominant processing consumer in our system, so that the 113 ms required for the preferred 800 Hz sampling rate is well in line with the system requirements.

### 6.3.3 Results and Evaluation

Using the mentioned techniques we were able to reduce the memory usage from 28 MB required for the generated Simulink implementation to 16 KB for the C-based implementation. The average CPU utilization of the RPi dropped from more than 70% required for Simulink to less than 5% for the C-based implementation.

In this section we describe the measurement results for the C-implementation on the RPi with the sensor board (see Figure 6.1). For our tests we used two sensor systems to gather the data required for the authentication process. During the tests we validated the general functionality and quality of the measurements. The two independent sensors were attached
to the left and right wrist respectively, with separate reference signals at the chest area. For the tests, the data was synchronized by time stamps, stored in text files that could be analyzed offline. In total we gathered 800 IPI samples.

A snippet of measured data for two nodes is shown in Table 6.2. We can see that both sensors deliver very similar Q, R, and S peaks, with only minor deviations. A histogram of the obtained IPIs, expressed as heart beats per minute is shown in Figure 6.4 (A). All measurements are between 60 and 100bpm, without any uncharacteristic outliers.

As underlying statistics for the authentication protocol we further tracked how close the IPI values for adjacent heartbeats are, and how much timing error we observe between two nodes. The results for the two studies are shown in Figure 6.4 (B) and (C) respectively. We see that adjacent heartbeats are related to each other, with a standard deviation of more than 60ms. In contrast, our measurement error between the boards shows a standard deviation of

![Histograms](image)

Figure 6.4: Distribution of (A) measured IPIs, (B) difference between two adjacent measurements, and (C) measurement errors between sensors
\( \sigma_s = 0.74 \text{ms} \). Two orders of magnitude difference between measurement errors and natural uncertainty of the underlying phenomenon generally indicate the applicability of our system for a secure physiological authentication. The measurements were taken for a sampling rate of \( f = 800 \text{Hz} \). The uncertainty of the measurement error increases with lower sampling rates. For a sampling rate of \( f = 666 \text{Hz} \) we measured a standard deviation of \( \sigma_s = 6.7 \text{ms} \), and \( f = 500 \text{Hz} \) resulted in \( \sigma_s = 110 \text{ms} \). We could not observe an improvement of the measurement precision for sampling rates above 800Hz.
Chapter 7

Protocol

In this chapter we apply the results and measurements we gained for the system, presented in the previous chapter, to engineer and evaluate a secure and robust biometric authentication protocol. The major challenge is that different sensor nodes measure similar but not exactly the same values, so that a margin of error has to be accepted. In the authentication protocol we exploit the fact, that the natural distribution of IPIs (see Figure 6.4 (B)) is larger than the measured uncertainty of our system (Figure 6.4 (C)). The hypothesis is that we can decide whether deviations for a set of sampling points from different sensors is caused by technical uncertainties or by a node that is not attached to the same body. Furthermore, our investigations address the question, how the number of samples and their resolution affect the quality and reliability of the authentication process.

7.1 Authentication Protocol

The authentication protocol we discuss in this section ensures that two nodes ($S_1$ and $S_2$) agree that they are attached to the same body, that is, they sense the same ECG data. We
do not assume any specific network topology or hierarchy between $S_1$ and $S_2$. We further assume that $S_1$ and $S_2$ already agreed on a shared session key, which is used throughout the communication. That might be possible with Diffie-Hellman or other light-weight key agreement protocols [4].

A message sequence diagram of the physiological authentication protocol with its 5 steps is shown in Figure 7.1. The protocol starts with the session establishment and measurement of the IPIs. The key ideas of the protocol are:

1. The IPIs of the two peers are compared based on statistical properties in the $IPI_1 \approx IPI_2$ operation in step 5. The comparison operation ($\approx$) compares the standard deviation of the differences between the two nodes and depends on a range of uncertainties and parameters that we discuss in the next section.

2. Each peer sends a hashed value of its measured IPIs before sending the actual IPI data (step 3). Receiving the peer’s hash value (e.g. $H_2$) before sending the own IPIs ($IPI_1^*$), prevents the peer node from forging its IPI ($IPI_2^*$) after receiving the authentic data. If the received IPI does not match the received hash value (step 4), the authentication is invalid.

3. Possible replay attacks and simultaneous connection attempts are additionally hindered by enforcing the use of a unique random nonce $N_i$, which in our case is a 32 bit random integer number. The nonce is generated in step 2, and has to be used by the peer for the hash generation (step 3). Therefore, $H_i$ is the hash value of the concatenated received nonce and the array of measured IPI values.

A peer node is authenticated as a member of the BAN if the last two steps (4 and 5) succeed.
1. Get IPI Measurements \((IPI_1^*)\)
2. Generate Random Unique Nonce \(N_1\)
3. Compute \(H_1 = \text{hash}(N_2, IPI_1^*)\)
4. Compare \(H_2 = \text{hash}(N_1, IPI_2^*)\)
5. Compare \(IPI_1 \approx IPI_2\)

Figure 7.1: Message sequence chart for the biometric authentication protocol between Sensor nodes \(S_1\) and \(S_2\), with processing steps for \(S_1\). The authentication is successful if the last two steps succeed.

### 7.2 Parameters of the Authentication Protocol

Since participated sensor nodes might sense or process different IPI values, the comparison function \(\approx (IPI_1^*, IPI_2^*)\) over the two IPI arrays of equal size has to accept a degree of uncertainty in the data values. Thereby, it is the goal to minimize the number of false positive authentications (FPA) and false negative authentications (FNA).

**False negative authentications (NFA)** is the percentage of legitimate authentication attempts that are rejected due to a high level of measurement or processing errors.

**False positive authentications (NPA)** is the percentage of non-legitimate authentication attempts that are accepted by a node due to a high tolerance for measurement or processing errors.

The comparison function \(\approx (IPI_1^*, IPI_2^*)\) compares the statistical properties of the two IPI vectors. The function and its acceptance properties and its footprint can be tailored by three parameters:
Dynamic range \( r \), measured in dB, expresses the resolution of the sampled IPI values. \( r \) is based on the distance \( \Delta \) between two adjacent IPI values in ms. Since the underlying physical phenomenon is in a range of \( R = [0, 2000]\)ms, we can compute the required number of bits per value by \( b = \log_2 \frac{|R|}{\Delta} \), and the dynamic range is \( r = 20 \log_{10} \frac{|R|}{\Delta} \). Table 7.1 shows example values for \( \Delta \), \( b \), and \( r \). We assume that a lower dynamic range will reduce the number of false positives but increases the number of false negatives.

It should be noted that changing the resolution also changes the statistical properties of the reference model. For instance the results in Figure 6.4 was taken with a resolution of \( r = 0.5\)ms and resulted in a standard deviation of \( \sigma_{0.5ms} = 1.28 \), while a resolution of 2ms results in \( \sigma_{2ms} = 0.45 \).

Number of samples \( s \) defines the number of IPI data points we use for one authentication process. The number of samples influences the time for taking measurements, as well as the message size. Table 7.2 shows the message sizes for a range of configurations. From quality perspective we can expect that more samples compensate for outlier data points, and increase confidence in the positive or negative decision.

Allowed deviation \( d \), measured in number of standard deviations \( (\sigma) \), determines how much deviation is acceptable to distinguish legitimate measurements from data not originating from the same body. Therefore, \( d \) can be considered as similarity factor that decides whether a node is trusted or not. The value of \( \sigma \) is based on the actual measurements (see Figure 6.4) and the underlying dynamic range. Using the error factor \( e = \frac{\sigma_e}{\sigma} \), while \( \sigma \) is the standard deviation of the expected measurement error (see Figure 6.4 (C)), and \( \sigma_e \) is the error distribution of the difference between \( IPI_1 \) values.

### Table 7.1: Quantization per sample

<table>
<thead>
<tr>
<th>distance ( \Delta ) [ms]</th>
<th>req. bits ( b )</th>
<th>dyn. range ( r ) [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>66</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>53</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>40</td>
</tr>
</tbody>
</table>

### Table 7.2: Example signature sizes

<table>
<thead>
<tr>
<th>distance ( \Delta ) [ms]</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2 20 40 80 160</td>
</tr>
<tr>
<td>4</td>
<td>4 18 36 72 144</td>
</tr>
<tr>
<td>8</td>
<td>8 16 32 64 128</td>
</tr>
<tr>
<td>20</td>
<td>20 14 28 56 112</td>
</tr>
</tbody>
</table>
and $IPI_2$:

$$\sigma_e = \sqrt{\frac{1}{s} \sum_{i=1}^{s} (IPI_1[i] - IPI_2[i])}$$  \hfill (7.1)

we define the comparison function as:

$$\approx (IPI_1^*, IPI_2^*) = \begin{cases} 
true & \text{iff } d \leq e \\
false & \text{iff } d > e 
\end{cases}$$ \hfill (7.2)

It can be assumed that a larger $d$ reduces the number of false negatives, but increases the number of false positives.

The number samples ($s$) and the dynamic range ($r$) have a minor effect on the total computation time of the protocol. The protocol requires two hashing operations (of the nonce and the samples), and the computation of one standard deviation (of the differences between the measurements of the two peers). Depending on $s$ and $r$, the total time for these operations varies between $130\mu s$ and $160\mu s$ when we apply MD5 as hash function, and between $250\mu s$ and $290\mu s$ when SHA256 is applied. These figures are negligible compared to the signal processing efforts reported in Section 6.3 and therefore do not affect the search for preferable parameter combinations.

### 7.3 Analysis

To identify superior combinations of parameters in this multi-dimensional space – and to validate the parameter assumptions expressed in the previous section – we executed a variety of tests with different variable and invariant parameters. All the experiments were executed in a Matlab environment using the authentication protocol proposed in Section 7.1. The
expected reference error distribution and its standard deviation $\sigma$ is provided by our practical sensor measurements (Figure 6.4). In the following section we investigate the impact of the parameters on the FNA first, then study impact on the FPA, and conclude with the combined parameter selection and a sensitivity analysis.

### 7.3.1 False Negatives Authentications (FNA)

Figure 7.2 (A) and Figure 7.2 (B) show the rate of FNAs for different parameter settings. We applied the measured data for two sensors and applied different settings of sampling rate, dynamic range, and allowed deviation. The goal is to identify the sensitivity of the parameters to the legitimate authentication attempts.

Figure 7.2 (A) shows the impact of the sample sizes for different allowed standard deviations, averaged for a invariant set of dynamic ranges. The results from this picture are:

- For all cases but $d = 1\sigma$, an increased number of samples reduces the number of false negatives. The reason is that a higher number of samples help to reduce the effect of possible outliers.

- Allowing only a deviation of $d = 1\sigma$ is too strict for all settings. In fact we see that higher number of samples increases the probability of a rejected authentication.

- With a deviation of $d = 4\sigma$ and more, and $s = 4$ samples and more, we practically cannot identify any false negatives anymore, i.e. all legitimate authentications are identified correctly.

Figure 7.2 (B) shows the impact of the dynamic range for the same standard deviations. The results are similar to the previous results in a way that:
Figure 7.2: Rejected authentications (false negatives) for measured truthful authentication attempts, (A) for different number of samples, and (B) for different dynamic ranges.

- Allowed deviation of $d = 1\sigma$ and $2\sigma$ results in too many errors regardless of the dynamic range.

- With a deviation of $d = 4\sigma$ and more, and a dynamic range of $r = 50$dB (8bit) and more, we practically cannot identify any false negatives anymore, i.e. authentication attempts are assessed correctly.

While the latter observation is most important for our parameter selection, it is worth to note that increasing the dynamic range does not generally reduce the FNR. One reason for this behavior could be that very high and very low resolutions might over-emphasize outliers, and therefore reject valid authentications.

7.3.2 False Positive Authentication (FPA)

In this section we investigate the impact of the parameters to non-legitimate authentication attempts. We assume the threat of similarity attacks [28] in which an adversary can exploit that values in an equivalence class are distinct but semantically similar. In other words, if an adversary can generate or guess similar IPI values, the presented protocol can be
compromised. To generate non-legitimate IPIs we assume a powerful attacker model in which an attacker knows the last legitimate sensor reading and is aware of the inter-IPI distribution (Fig. 6.4 (B)). Other models, such as simple guessing, or first order approximations did not result in observable successful authentications.

Figure 7.3 (A) and (B) show the FPA for different numbers of samples and dynamic ranges, respectively. The results in this case are:

- Increasing the number of samples or increasing the dynamic range always decreases the chance for a successful attack.

- Very high allowed deviations ($d > 10\sigma$) improve the chances for a successful attack, while small allowed deviations ($\leq 8\sigma$) prohibit attacks.

### 7.4 Parameter Selection and Sensitivity Analysis

Based on the practical and analytical experimental results, the parameters of the comparison function $\approx$ can be set, so that the FPA and FNA rate is reduced, and the required resources
(packet size) and authentication time (number of required samples) are considered. As a result, for our analysis of the data gathered with the presented BAN setup, we decided on a sample size of 8, with 8bit sample encoding (48dB) and an allowed deviation of $8\sigma$. In our experiments, this setup led to 0% false positives and 0% false negatives. In fact for the given setting, the allowed deviation $d$ could be freely selected between 2 and 14 to obtain the same positive result.

To investigate the sensitivity of the parameter selection, in cases where the nodes process the data with a higher timing uncertainty than in our experiment, we ran tests with an increased noise level. Specifically, we studied the sensitivity of the accepted deviation $d$ in environments with higher uncertainty. Possible reasons for higher deviations are less precise sensors and signal processing, timing-uncertainties in the embedded system, but also older or less healthy persons [40]. We were interested in the required thresholds of $d$ to limit the FNR to 1%, and to limit the FPR to the levels of 1% and 5%.

The results are shown in Figure 7.4. Figure 7.4 (A) shows the thresholds for our selected settings (8 samples, 48dB). We see that the acceptable threshold for the FNR stays between $1\sigma$ and $2\sigma$ and is not affected by increased variation of the underlying measurements. The reason is that $\sigma$ is a factor in the acceptance equation 7.2, and therefore an increased noise level in the underlying distribution increases the acceptance deviation as well.

However, Figures 7.4 (A) and (B) show that the thresholds for FPR decrease with an increased deviation in the reference error distribution. That effect is plausible, because if a high level of noise has to be tolerated, it is more challenging to differentiate between deviations due to attacks and natural noise.

Our default settings of the protocol, shown in Figure 7.4 (A), can tolerate measurement uncertainty up to $\sigma = 4ms$, if we require $1\sigma$ between the FPA and FNA thresholds. Higher noise levels would increase the FPAs, since attacks cannot be clearly distinguished from
Figure 7.4: Impact of increased variance (e.g. more noise) in the basis measurements on the allowed thresholds to limit false positives and false negatives.

Larger uncertainty than $\sigma = 4$ ms requires an adaptation of the system parameters. Figure 7.4 (B) shows the thresholds for 16 samples and 60 dB (11 bit) dynamic range. The plot shows that in this case a 20 times higher basis deviation still provides sufficient space to differentiate between authentic and forged attempts. However the cost for this improved confidence is the extended sample time (16 instead of 8 heart beats) and the larger packet sizes (160 instead of 64 byte).

To validate the simulation results, we applied practical measurements gathered with different sampling rates. The effective ranges for a 99% confidence of FPA and FNA are highlighted as the rectangles in Figures 7.4 (A) and (B). We only show the 800 Hz and 666 Hz sampling rate, because higher rates were not distinguishable from 800 Hz. Lower rates such as 500 Hz result in a standard deviation of $\sigma > 100$ ms, which is unpractical. The measurements show that the practical FNA threshold is about $1\sigma$ higher than estimated in the simulations, while the FPA threshold corresponds to the simulations. As a result shown in Figure 7.4 (A), the data gathered at 666 Hz cannot be successfully distinguished, because with $\sigma = 6.8$ the practical threshold to prevent attacks is lower than the threshold to identify legitimate attempts.
With the extended setting (Figure 7.4 (B)) we still have a valid range for $d$ between $2.5\sigma$ and $4\sigma$. The results show that protocol parameters with a preferable quality-to-resource trade-off can be found if the uncertainty of the implemented system are known.
Chapter 8

Conclusion

The combination of real sensors, sophisticated protocols and system parameterization is never easy. This thesis has shown how a physiological authentication protocol for body-area networks can be designed, implemented, and parametrized to work with the real-world uncertainties of low-cost body-area sensor nodes. The key for the presented authentication protocol is the statistical analysis of actual sensor measurements, which allows a designer to adapt system parameters in accordance with properties of real-world BAN deployments.

As reproducible basis of our work, in the first part of this thesis we described how to design and implement the ECG sensor board and its processing system. In this part it turned out that noise and signal quality had to be addressed in all design and processing steps, starting from analog pre-processing to the model-based validation step of the detected ECG peak signals. During the design we further observed that model-based design flows are helpful in early design stages. However, to obtain a sufficient system performance, manual design efforts were required, in both the analog and the digital part of the system. The resulting sensor system is the first reported BAN sensor system to facilitate real-time tracking of ECG IPI data for inter-node authentication.
Practical measurements with the presented systems are the key for the design and the parametrization of the actual authentication protocol. With the statistical data of the practical behavior, we could define the allowable deviation margins so that honest authentications were permitted and false attempts could be prevented. We exploited the fact that the deviations of the underlying biometric property is higher than the deviations of sensing and processing data to optimize and tailor the parameters of the authentication function and to reduce sensing and processing overhead. The recommended system configuration requires 8 samples of 8 bit integer precision each, resulting in 100% correct authentications and an infeasible probability for a false positive authentication.

While the results of this first real-world biometric BAN system are promising, we also identified a range of possible future works. One future work is the extended validation of the protocol for more people in different situations and possible abnormal ECG properties. The goal would be a combination of biological sources of uncertainty as discussed in [40] and the technological aspects, discussed in this thesis, to one framework to determine preferable system configurations to minimize the probability of errors. Another open research question is how can we make model-based design more effective, without the need of reimplementation of algorithms in a lower level implementation languages. The goal is the adaptation of the presented setup to an even small smaller 8 or 16 bit computation platform. To reach that goal, extending the presented authentication scheme to generate secure session keys, instead of accepting pre-agreed session keys, could further improve the efficiency and applicability of the presented biometric authentication system.
Bibliography


